Information Design for Social Learning on a Recommendation Platform

Chen Lyu*

This version: Oct 19, 2024

Abstract

A recommendation platform sequentially collects information about a new product revealed from past consumer trials and uses it to better guide later consumers. Because consumers do not internalize the value of information they bring to others, their incentive for trying out the product can be socially insufficient. Given such a challenge, I study how the platform can improve social welfare by designing its recommendation policy. In a model with binary product quality and general trial-generated signals, I find that the optimal design features a U-shaped sequence of recommendation standards over the product's life, and the optimal learning dynamic can involve temporary suspensions following negative consumer feedback when the product is young. Various extensions and comparative statics regarding the optimal recommendation standards are provided. My analysis also illustrates the usefulness of a Lagrangian duality approach for dynamic information design.

Key Words: information design, social learning, recommendation platform, Lagrangian duality

*Peking University HSBC Business School, clyu@phbs.pku.edu.cn. I am grateful to Lones Smith, Marzena Rostek and Daniel Quint for valuable advice and comments. I also thank Emir Kamenica, Benjamin Bernard, Marek Weretka, Matteo Camboni, Nima Haghpanah, Axel Anderson, Miaojun Wang, Zhen Zhou, Teddy Kim, Heski Bar-Isaac as well as various seminar and conference participants at CY Cergy Paris University, the University of Wisconsin-Madison, Zhejiang University, Renmin University, CMGTA' 2024 and ESEM' 2024 for helpful discussions. All mistakes are my own.

1 Introduction

Recommendation platforms are quite popular in our daily lives. For examples, people rely on Goodreads for what to read, Netflix for what to watch, Yelp for where to eat and TripAdvisor for where to travel.[1] To provide better recommendations, a common practice of these platforms is to induce a kind of social learning for new products. Namely, they collect information generated from early consumers' trials of a product (e.g., rating and reviews), and use it to better guide the later consumers. In this process, however, because individual consumers do not internalize the value of information they bring to others, their incentive for trying out the new product can be insufficient. This handicaps learning and can hinder the platforms from making better-informed recommendations.

I study how a platform facing the challenge above should design its dynamic recommendation policy, which can potentially "persuade" consumers towards more socially desirable trials for a new product. In the main model, a sequence of short-lived consumers arrive over the (finite) lifetime of a product with unknown quality, which can be either high or low. Whenever a consumer consumes the product, a signal about its quality will be generated and privately observed by the platform.[2] Unlike some existing studies surveyed later, I allow such consumption-generated signals to be general and non-conclusive. In each period, based on the signals previously received, the platform can guide the current consumer by providing a recommendation message. Knowing the message and her own arrival time, the consumer then makes her consumption decision in a Bayes-rational way. The platform's design problem is to find a dynamic recommendation policy, to which it can commit ex-ante, in order to maximize the total consumer surplus generated on it.[3]

Notice that the platform's recommendations play a dual-role – they both decide how past information is used to guide the current consumer and decide whether new information will be generated for later use.[4] Ideally, the platform should recommend the product for trial as long as this is socially desirable after taking into account the informational value, even if consumption is suboptimal for the current consumer based on the current information. With such a policy, however, the expected quality of some recommendations may be too low for self-interested consumers to follow. An optimal policy therefore must choose when to recommend socially desirable but individually suboptimal consumption

[1]For some non-commercial examples, consider FDA for drug uses and Medicare Advantage Star Ratings for Medicare plan choices in the US.

[2]The platform also receives an initial piece of private signal about the product before any consumption, which reflects its internal research or data about past performance of similar products.

[3]In an extension, I also allow the platform to be biased towards inducing more consumption.

[4]Technically, the design problem is a decentralized bandit problem, where a principle (platform) faces a bandit process but cannot control it directly. Instead, the process is controlled by a sequence of short-lived agents (consumers). The principle privately monitors the process and can influence the agents' decisions only via providing information.

most efficiently, subject to the requirement that the consumer in every period will be willing to follow the recommendation.

The need to convince consumers to follow recommendations induces a sequence of incentive-compatibility (abbr., IC) constraints – one for each period – in the dynamic design problem, which makes it a *constrained Markov decision process*.[5] Solving such a problem is challenging because the standard dynamic programming technique cannot directly handle those constraints. I hence adopt a Lagrangian duality approach. It allows me to partially characterize the shadow values of those IC constraints and thereafter convert the optimization into an unconstrained one. To the best of my knowledge, this is the first paper solving a (non-degenerate) constrained Markov decision process that naturally arises from a dynamic information design problem.

I show that the optimal design features a sequence of time-specific thresholds, which generally vary in a U-shaped pattern over the product's life. At any time, the platform should recommend the product if its current belief of the product's quality being high is above the current threshold. Intuitively, this suggests that the platform should set a time-varying recommendation standard, which first goes down and then goes up as the product ages. Underlying this time-pattern is a tension between the platform's desire to create informational value for future consumers and its need to satisfy the current consumer's IC constraint. When the product is very young, the informational value is high due to a long remaining product lifespan, but consumers are "skeptical" about following recommendations because they know that even the platform has not acquired much information about the product yet. This implies a binding IC constraint and necessitates a picky censorship regarding when to recommend the product. As time passes, consumers become easier to convince as they expect that the platform may have gotten better informed by previous signals. The recommendation standard can hence be lowered.[6] This continues until the standard has become sufficiently low such that trials with beliefs further below it are no longer worthwhile. Thereafter, the optimal threshold will gradually go up because the informational value of consumption dwindles as the product approaches its end of life.

The result above implies an interesting prediction about the optimal recommendation dynamic: it can feature *temporary* recommendation suspensions following negative consumer feedback for young products. Specifically, following a negative feedback, the platform's belief of high quality can drop below the recommendation threshold, which suspends recommendation and learning. However, if we are in the early phase of the product's life where the threshold is declining, the threshold can fall below the belief again a few periods later, which restarts recommendation. In practice, it is well-documented that temporary recommendation suspensions can be used for punishing misconduct of a

[5]See Altman (1999) for a textbook treatment to constrained Markov decision processes.
[6]See discussion about Figure 3 in Section 4.1 for a more concrete intuition behind this.

product supplier.[7] My finding here suggests another motivation for taking such actions, which is to induce more efficient social learning about new products given the inadequacy of individual consumers' incentives.

My characterization of the optimal design also enables several comparative static analyses. The first one considers how the design should be adjusted when consumption becomes more likely to yield non-neutral signals about the product's quality (e.g., due to developments in feedback acquiring technologies). Using a coupling approach for comparing stochastic processes, I show that the recommendation standards should be uniformly lowered over the product's life when this happens. The intuition is twofold. First, as trials become more informative, the platform ideally wants to carry out more of them, which motivates lower standards. Second, knowing information has been accrued at a faster speed, the consumer in every period should be more willing to follow recommendation if the threshold is unchanged, which makes it feasible to indeed lower the standards.

The second analysis incorporates random consumer arrivals. I find that if the platform enjoys a thicker market where consumers arrive more frequently, the optimal policy should feature lower recommendation standards over the product's life. The intuition is again twofold. First, with more consumers expected to come, the value of information at any time is higher and the platform thus wants to induce more trials. Second, with a higher consumer arrival frequency, the consumer in every period will expect the platform to have acquired more signals ceteris paribus, which relaxes the IC constraints and leaves room for implementing lower standards.

In the third analysis, I extend the model to accommodate biased platforms. Specifically, I assume that the platform can enjoy extra commission from consumption and wants to maximize the sum of consumer surplus and its own commission. In this scenario, my characterization of the optimal design easily extends, and comparative statics with respect to the platform's bias can be provided. I show that when the platform's commission increases, the optimal recommendation standard will generally remain unchanged in the early periods of the product's life, but will go deeper down after the original bottom before bouncing back. Intuitively, when the platform becomes more biased, it wants to lower the standards more, but this can be done if and only if the consumers' IC constraints have turned non-binding without the change.

I will relax the binary quality assumption in an appendix (Appendix A). Although a full characterization for the optimal design is not available there, my duality approach still helps to reveal certain properties of it. In particular, I show that the optimal policy generally features a partial-order monotone structure, which can be considered as a generalization of the threshold structure. I will discuss implications of this result for algorithmic recommendation design.

[7] For a list of real-world examples, see Table 1 in Liang et al. (2020).

Related literature – My study is closely related to Kremer et al. (2014) and Che & Hörner (2018), who also study the optimal recommendation design when early consumption yields informational externality to later consumers.[8] These papers have focused on special classes of consumption-generated signals. Specifically, the main model in Kremer et al. (2014) considers fully revealing signals, i.e., the underlying quality will be fully revealed after a single trial. They thus focus on the decision problem about when to induce the first trial based on the platform's initial information.[9] Che & Hörner (2018) considers a Poisson learning environment with binary quality levels in continuous time. They assume that at any time the platform has either received no news, or has received conclusive news that fully reveals the product's quality. The design problem then boils down into a deterministic control problem about the recommendation intensity following the history without any news arrival. Unlike these papers, my study accommodates general non-conclusive signals. My characterization of the optimal design is thus about whether to recommend the product in each period based on any current belief of the platform, which goes beyond timing of the first trial or recommendation intensity without previous news. This allows me to examine the dynamic pattern of time-varying recommendation standards and necessitates the more general mathematical formulation. Moreover, the general setting also enables my comparative static analyses, which have no counterparts in the previous papers.

An extension in Kremer et al. (2014) and a strand of subsequent algorithm-oriented research have studied environments more general than the main model of Kremer et al. (2014), which do allow for non-conclusive consumption-generated signals (Papanastasiou et al., 2018; Mansour et al., 2020).[10] The goal of this literature is to propose algorithms that can achieve better asymptotic performance as the number of consumers coming in sequence goes to infinity, which is often measured by the decay rate of per-consumer welfare loss compared to the full-information first-best benchmark. While such measurement reflects an important aspect of the design's performance, it ignores the early consumers' potential loss from trying a low-quality product within any finite time horizon, and can be insensitive to multiplicative changes in the total welfare loss.[11] Hence, for the algorithms proposed in this literature, little is known about their finite-horizon efficiency, and little has been done to improve their non-asymptotic performance. My paper complements

[8] Lorecchio & Monte (2021) also considers a setting where the designer records previous agents' feedbacks to guide later agents' decisions. However, their designer has state-independent payoff, rely on restricted communication rules, and only focuses on the long-run stationary equilibrium, which makes their paper distinct from mine.

[9] More precisely, the initial information in their paper is about an alternative consumption option, which is always tried by the first consumer with its quality fully revealed since then.

[10] Also see, for examples, Bahar et al. (2015), Mansour et al. (2016), Chen et al. (2018), Immorlica et al. (2019) and Bahar et al. (2021) for a variety of extensions.

[11] To see this, notice that an average loss function $L(t)$ is considered to have the same decay rate in t as $\alpha L(t)$ for any $\alpha > 0$.

the literature by solving the optimal design in a stylized setting with fixed time horizon, which may serve as a performance benchmark for evaluating any algorithm and help to inspire new algorithms with a non-asymptotic focus.[12]

Another growing literature also considers the optimal information provision by a platform to a sequence of short-lived agents (Glazer et al., 2021; Komiyama & Noda, 2021; Küçükgül et al., 2022). In these papers, the agents are either endowed with or are able to acquire private signals, and a central task for the platform is to infer these private signals from the agents' decisions. These papers thus consider very different information sources of the platform and explore design concerns distinct from mine.

More generally, my paper belongs to the broad literature on information design (Kamenica & Gentzkow, 2011; Rayo & Segal, 2010), and especially to studies on dynamic designs (e.g., Ely, 2017; Renault et al., 2017; Smolin, 2021; Ely & Szydlowski, 2020; Ball, 2019; Orlov et al., 2020; Lorecchio, 2021). One difference between many of the studies in this literature and mine is that I consider a designer whose private information flow is controlled by the receivers' decisions, rather than being exogenous. My analysis illustrates how such a setting naturally leads to a constrained Markov decision process after simplification by the revelation principle of Myerson (1986),[13] and how the Lagrangian duality approach can be useful for solving it.[14]

The paper is organized as follows. Section 2 presents the main model. Section 3 derives the optimal design. Section 4 explores dynamic properties of the optimal design. Section 5 considers comparative statics. Section 6 provides additional discussions. Section 7 concludes with some methodological remarks. Appendix A considers the extension with general quality support. All proofs are provided in Appendix B.

2 The Main Model

I first describe the model, and then discuss several underlying assumptions in Section 2.2.

[12] Papanastasiou et al. (2018) does investigate finite-horizon design in a particular setting, but in that setting the initial information is such that either no exploration can ever happen or consumer IC constraints are never binding, which makes the optimal design obvious. They also propose to formulate the designer's problem as a constrained Markov decision process in a more general setting, but concluded it to be computationally infeasible and did not derive analytical results from it except for giving a bound on how many belief states need to involve randomization in an optimal design.

[13] I provide a more detailed remark on this in Section 7.

[14] Beutler & Ross (1985) and Beutler & Ross (1986) were the first to use a Lagrangian approach to study constrained Markov decision processes. The method is subsequently developed and applied in many mathematical and engineering papers (see, e.g., Section 1.2 of Altman (1999) for a brief survey). These studies typically only involve a few aggregate constraints corresponding to different design criteria. In contrast, my problem features one constraint for each period, which leads to a novel dynamic aspect of the problem.

2.1 The Setting and the Design Problem

The model features a platform, a sequence of short-lived Bayes-rational consumers and a product. The product is launched in period 1 and will remain available for consumption over $T < \infty$ periods. In each period $t = 1, ..., T$, a consumer arrives at the platform and decides whether to consume the product. I denote the consumer's decision as $a_t \in \{0, 1\}$ with $a_t = 1$ meaning consumption occurs. Without consuming the product, the consumer will receive her outside option, whose value is normalized to zero. If she consumes the product, the consumer's utility will be equal to $\tilde{\theta}$, which is a random variable taking values in $\{\theta_L, \theta_H\}$, with $\theta_L < 0 < \theta_H$. This $\tilde{\theta}$ measures the underlying quality of the product, which is fixed over time but initially unknown. I assume that the platform and the consumers share a common prior p_0 for $\theta = \theta_H$.

At the beginning of period 1, the platform receives a signal s_0 about $\tilde{\theta}$, which captures its internal research or private data on past performance of similar products. Subsequently, an additional signal will be generated to the platform whenever a consumer consumes the product. Let s_i denote the signal from the i'th consumption of the product. Conditional on $\tilde{\theta}$, I assume that $s_1, s_2, ...$ are i.i.d. and are independent from s_0.

In every period, the platform can compute its posterior belief about the product's quality based on previous signals. I use p_t to denote the platform's belief about $\tilde{\theta} = \theta_H$ at the beginning of period t. Let μ_1 denote the distribution of $p_1 = \mathbb{P}(\tilde{\theta} = \theta_H | s_0)$; let $G(\cdot|\cdot)$ denote the transition kernel of belief reflecting Bayesian updating when consumption occurs; let $D(\cdot|p)$ denote the Dirac measure at p. The process of $(p_t)_{t=1}^T$ then follows the following transition rule:

$$p_1 \sim \mu_1 \tag{1}$$

$$p_{t+1} | p_t, a_t \sim a_t G(\cdot|p_t) + (1 - a_t) D(\cdot|p_t) \tag{2}$$

Before the realization of s_0, the platform can commit to an information transmission policy that decides what message to convey to the coming consumer in each period based on the information available at that time. I assume that the consumer can neither observe previous messages, nor observe decisions of earlier consumers.[15] She only observes her arrival time and her own message, and then decides whether to consume the product.

The timeline of the environment is summarized as follows:

1. Before period 1, the platform (publicly) commits to an information transmission policy, and the product's quality $\tilde{\theta}$ is secretly realized. Then, the platform privately receives its initial signal s_0 about $\tilde{\theta}$.
2. In every period $t = 1, ..., T$, a consumer arrives and receives a message from the

[15]The model therefore mainly fits applications where the platform is using private recommendations sent to each individual consumer instead of using public recommendations.

platform, which is generated according to the information transmission policy. She then decides whether to consume the product. If she is the n'th consumer who consumes the product, signal s_n will be generated to the platform. The economy then enters into the next period.

The Designer's Problem: I look for the information transmission policy that maximizes the total consumer surplus to be generated over the product's life, i.e., $\sum_{t=1}^{T} \mathbb{E}[a_t\tilde{\theta}]$.[16] By the revelation principle (Myerson, 1986), it suffices to consider *incentive-compatible recommendation policies*, which just decide whether to *recommend* the product for consumption in every period, subject to the requirement that a Bayes-rational consumer will follow the recommendation. Since the belief p_t summarizes all the payoff-relevant information of the platform in period t, standard argument implies that we can further focus on randomized Markov policies with respect to the process of $(p_t)_{t=1}^{T}$.Formally, a *randomized Markov recommendation policy* is a sequence of measurable mappings $\phi := (\phi_t : t = 1, ..., T)$, where each $\phi_t : [0,1] \to [0,1]$ decides the probability of recommending the product at time t given any $p_t \in [0,1]$. For the rest of the paper, by "policy" I will be referring to a policy of this type.

I impose the following assumption on consumption-generated signals throughout:

Assumption 1. Assume the following:

(i) $\mathbb{P}(\mathbb{E}[\tilde{\theta}|s_0] > 0) > 0$;

(ii) For any $i \geq 0$, we have $\mathbb{P}(s_i \in A|\tilde{\theta} = \theta_L) < \mathbb{P}(s_i \in A|\tilde{\theta} = \theta_H)$ for some (measurable) set A in the realization space of s_i;

(iii) For any $i \geq 0$, we have $\mathbb{P}(s_i \in A|\tilde{\theta} = \theta_L) > 0 \Leftrightarrow \mathbb{P}(s_i \in A|\tilde{\theta} = \theta_H) > 0$ for any (measurable) set A in the realization space of s_i.

Condition (i) implies that it is optimal for the first consumer to consume given some realizations of the platform's initial information.[17] Without such a condition, the first consumer will never want to consume the product, knowing which the second consumer will never consume either. Induction then implies that no consumption can ever happen under any design. Condition (i) rules out such a trivial scenario. Condition (ii) guarantees that the signals are indeed informative about $\tilde{\theta}$. Condition (iii) implies that no signal realization can conclusively reveal the quality level. It helps to simplify the exposition of certain proofs, but is not essential for results in the paper.

[16] In practice, the platform may want to maximize total consumer surplus for different reasons. For example, it may be able to extract that surplus by charging a subscription fee. My analyses can also accommodate biased platforms who wants to maximize the sum of consumer surplus and some extra commission from consumption, which are considered in Section 5.3.

[17] To guarantee this, one may instead impose the slightly weaker condition $\mathbb{P}(\mathbb{E}[\tilde{\theta}|s_0] \geq 0) > 0$. I impose the stronger condition for a technical reason when deriving the duality result.

2.2 Discussion on Model Assumptions

1. **Consumer information on product launch time.** An important assumption of the model is that each consumer can observe when the product was launched. This is reasonable for many products like books, TV-series, restaurants or hotels, which typically have a public release or opening time. For some other products, however, the consumer may only have a rough idea about their ages. In such a case, my design setting can be considered as being robust to uncertainty about the consumer's exact information. Under the assumption that consumers can perfectly observe the product's launch time, the optimal design being derived will be incentive-compatible no matter what information consumers actually have about the product's age. It thus provides the best guaranteed performance.
2. **Information v.s. monetary incentive.** The model assumes that the platform cannot directly pay early consumers for trying out the new product. While it can work well in some applications, the use of monetary incentive may be problematic in others. In particular, if consumers are only attracted by the monetary incentive instead of the product itself, they may *pretend* to consume the product and leave some artificial feedback just to earn the money, especially when the product's pecuniary price is zero (e.g., digital contents on a subscribed platform).[18] Moreover, monetary subsidy can sometimes attract a biased pool of consumers, whose feedback can be unhelpful or even misleading.[19] I hence focus on the design of information instead of monetary incentives in this paper.
3. **Timing and consumer flow.** A feature of the model is that consumer arrivals are distributed evenly over time. Instead of treating this as a simplifying assumption, one can understand it in terms of how the model times $t = 1, 2, ...$ are defined. That is, the model time flow is *arranged* so that the consumers will arrive evenly over time. This implies that, for example, if consumers arrive more frequently in the first calendar month after the product's launch than in the second month, we should have denser model periods in the first month than in the second month. In this way, the model can capture product life-cycle with time-varying consumer flow flexibly.

[18]For example, one may play a movie at background without watching it, and then fabricate some feedback to be paid. This kind of moral hazard problem makes it generally hard to buy information directly with monetary payment.

[19]For example, if we subsidize for trying a fine-dining restaurant, we may attract someone who will never eat there without the subsidy. Feedback from such consumers can be misleading as the product is actually targeting someone else.

3 Characterization for the Optimal Design

3.1 The Constrained Markov Decision Process

Let Φ denote the set of all (measurable) policies. Given any $\phi \in \Phi$, I use $\mathbb{P}_\phi$ to denote the probability measure over events about $\big((a_t)_{t=1}^T, (p_t)_{t=1}^T\big)$ provided that consumers follow the recommendations, and use $\mathbb{E}_\phi$ to denote the corresponding expectation operator. Then, the incentive-compatibility (IC) constraint for a time-t consumer can be written as:

$$\mathbb{P}_\phi(a_t = 1) > 0 \Rightarrow \mathbb{E}_\phi[u(p_t)|a_t = 1] \geq 0 \tag{3}$$

$$\mathbb{P}_\phi(a_t = 0) > 0 \Rightarrow \mathbb{E}_\phi[u(p_t)|a_t = 0] \leq 0 \tag{4}$$

where $u(p) := p\theta_H + (1-p)\theta_L$, i.e., the expected consumption utility given belief p. These respectively guarantee that the consumer will follow the recommendation when the product is recommended ($a_t = 1$) and when it is not ($a_t = 0$). Since one's consumption of the product generally benefits later consumers by yielding information, the designer will never want to recommend $a_t = 0$ when consumption is optimal for the current consumer. This implies that the second constraint above is non-restrictive for the designer and can thus be omitted. For the first constraint, we can more compactly write it as $\mathbb{E}_\phi[u(p_t)|a_t = 1]\mathbb{P}_\phi(a_t = 1) \geq 0$, which is equivalent to $\mathbb{E}_\phi[a_t u(p_t)] \geq 0$. Notice $\mathbb{E}_\phi[a_t u(p_t)]$ is the expected surplus of the time-t consumer when she follows the recommendation under ϕ. The designer's problem can hence be formulated as:[20]

$$\max_{\phi\in\Phi} \Big\{ \sum_{t=1}^{T} \mathbb{E}_\phi[a_t u(p_t)] \Big\} \tag{5}$$

$$\text{s.t. } \mathbb{E}_\phi[a_t u(p_t)] \geq 0 \quad \forall t = 1, ..., T \tag{6}$$

Due to the presence of the expectation operator, each constraint in (6) is not just restricting the recommendation decision following a particular realization of p_t, but involves integration over the entire distribution of p_t. Such an aggregated IC constraint arises because the payoff-relevant process $(p_t)_{t=1}^T$ is only observed by the platform but not by the consumer, who therefore must integrate over the equilibrium distribution of p_t when computing her posterior belief given any recommendation message. The presence of such constraints makes the problem a *constrained* Markov decision process, which cannot be directly handled by dynamic programming with p_t being the state variable. To overcome this difficulty, I will provide a dual characterization in Section 3.2, which allows

[20] Papanastasiou et al. (2018) first proposed the constrained Markov decision process formulation for this kind of design problem. However, they do not pursue much further analysis with it. See footnote 12 for details.

me to partially reduce the problem into an unconstrained one.

The following lemma reveals the key properties of the belief process needed for later analyses. In the statement of property (P5) below, $\bar{p}$ is defined as the indifferent belief for a consumer, i.e., $u(\bar{p}) = 0$.

Lemma 1. *The belief process (1) – (2) has the following properties:*
(P1) *$G(\cdot|p)$ as a measure-valued function of p is weakly continuous.*
(P2) *$\int_{p'} u(p')G(dp'|p) = u(p)$.*
(P3) *$G(\cdot|p)$ increases in p in terms of first-order stochastic dominance.*
(P4) *$G\big([0,p)|p\big)$ and $G\big((p,1]|p\big)$ are strictly positive for any $p \in (0,1)$.*
(P5) *$\mu_1\big((\bar{p},1]\big) > 0$.*

Property (P1) means that small changes in the prior can only lead to small changes in the posterior, which is a technical result that guarantees the existence of the optimal design. Property (P2) is implied by the standard law of iterated expectation. Property (P3) is an inertia property of the belief process, which roughly says that a product looking more promising today is also more likely to look promising tomorrow. It will be important for showing the threshold structure of the optimal policy. Property (P4) is directly implied by the assumption that signals are informative, which guarantees that the belief process will not stay constant for sure following consumption. Property (P5) is directly implied by Assumption 1(i), which allows consumption and learning to occur with some probability in period 1. I note that these five properties are all one needs to know about the belief process for later analyses, which abstract away from other details of the learning process.

The following result guarantees the existence of optimal design.

Proposition 1. *There exists an optimal solution to the designer's problem (5) – (6).*

3.2 The Dual Characterization

Given any vector of Lagrangian multipliers $\lambda \in \mathbb{R}_+^T$ associated to the IC constraints, I define the Lagrangian function of the designer's problem as:

$$\mathcal{L}(\phi;\lambda) = \sum_{t=1}^{T} \mathbb{E}_\phi[(1+\lambda_t)a_t u(p_t)] \tag{7}$$

Then, we have the following strong-duality result.

Lemma 2. *Let w^* denote the optimal value of the designer's problem. Then,*

$$w^* = \min_{\lambda\in\mathbb{R}_+^T} \sup_{\phi\in\Phi} \mathcal{L}(\phi;\lambda) \tag{8}$$

where the minimum is achieved by some λ^**. Given any such* λ^**, a policy* ϕ^* *is optimal for the designer's problem if and only if:*

(i) $\phi^* \in \arg\max_{\phi \in \Phi} \mathcal{L}(\phi; \lambda^*)$

(ii) $\lambda_t^* \mathbb{E}_{\phi^*}[a_t u(p_t)] = 0$, $\forall t = 1, ..., T$

(iii) $\mathbb{E}_{\phi^*}[a_t u(p_t)] \geq 0$, $\forall t = 1, ..., T$

To see how this result is helpful, notice that once we know a solution λ^* to the dual problem (8), which intuitively measures the "shadow values" of the IC constraints, the lemma implies that any optimal policy must solve $\max_{\phi \in \Phi} \mathcal{L}(\phi; \lambda^*)$, which is an unconstrained problem. The optimal design can then be characterized by studying this unconstrained problem with the standard dynamic programming approach.

The problem here, however, is that the value of λ^* is not available. Generally, deriving it requires one to either solve the min-max problem in (8) or solve the fixed-point problem defined by conditions (i) – (iii) jointly for (λ^*, ϕ^*), both of which are difficult. Fortunately, as I show below, a property of λ^* can be directly derived from the dual problem, which turns out to enable a sharp characterization for the optimal policy.

3.3 Main Structures of the Optimal Design

The following lemma is a key result derived from the dual problem (8).

Lemma 3. *There exists* $\lambda^* \in \arg\min_{\lambda \in \mathbb{R}_+^T} \sup_{\phi \in \Phi} \mathcal{L}(\phi; \lambda)$ *such that* $\lambda_t^* \geq \lambda_{t+1}^*$ *for all* $t = 1, ..., T-1$*.*

To see an intuition behind this result, assume that the dual problem has a unique solution λ^*. As usual, we can interpret λ_t^* as the shadow value of marginally relaxing the time-t IC constraint for the designer's problem. As time passes, two changes happen in the designer's problem. First, as information accumulates over time, we will be able to make recommendation selections more wisely. This makes it possible to obey the consumer's IC constraint with less sacrifice of socially desirable consumption. Second, as the remaining lifetime of the product gets shorter, the dynamic value from having additional myopically suboptimal consumption drops. These both suggest that relaxing later IC constraints is less helpful than relaxing the earlier ones. Hence, the associated shadow values should decrease over time. While this argument is intuitive, formalizing it can be difficult. I hence instead develop an "inter-change" argument for the lemma's proof. In particular, given any λ^* solving the dual problem, I show that if two adjacent components of it violate the time pattern, then interchanging them will lead to a new solution to the dual problem. Starting with any solution to the dual problem, one can thus construct a solution satisfying the time pattern by making such interchanges repeatedly.

An important implication of Lemma 3 is that the optimal design will generally feature a *two-phase structure*. In the first phase, $\lambda_t^* > 0$ and the IC constraints are thus binding;

in the second phase, $\lambda_t^* = 0$ and the IC constraints are thus essentially non-restrictive.[21] As we will see later, this two-phase structure is critical in constructing the optimal design, and has important implications on the dynamic pattern of optimal recommendations.

The time-pattern of λ^* found in Lemma 3 also turns out to induce a simple solution structure for the Lagrangian function optimization $\max_{\phi \in \Phi} \mathcal{L}(\phi; \lambda^*)$. To state this result, I define *threshold policies* as follows:

Definition 1. A time-t policy $\phi_t : [0,1] \to [0,1]$ is called a *threshold time-t policy* if there exists a threshold $\eta_t \in [0,1]$ such that $p > \eta_t \Rightarrow \phi_t(p) = 1$ and $p < \eta_t \Rightarrow \phi_t(p) = 0$. A policy ϕ is called a *threshold policy* if ϕ_t is a threshold time-t policy for every t.

Namely, a threshold policy will recommend the product when the current belief of $\tilde{\theta} = \theta_H$ is above a threshold, and will not recommend when the belief is below the threshold. It can also involve randomization at the threshold. By applying backward induction on the dynamic programming of $\max_{\phi \in \Phi} \mathcal{L}(\phi; \lambda)$, I show the following result:

Lemma 4. *Given any non-increasing sequence of multipliers $(\lambda_t)_{t=1}^T$, any solution to $\max_{\phi \in \Phi} \mathcal{L}(\phi; \lambda)$ is almost surely equivalent to a threshold policy. Moreover, $p_t > \bar{p} \Rightarrow a_t = 1$ a.s. under such a policy.*[22]

Together with Lemma 3 and the dual characterization for the optimal design, this directly implies the threshold structure of the optimal design:

Proposition 2. *Any optimal policy is almost surely equivalent to a threshold policy. Moreover, $p_t > \bar{p} \Rightarrow a_t = 1$ a.s. under it for any t.*

I note that although threshold policies are practically appealing, their optimality is not obvious in my setting. While the myopic value of consumption always increases in p_t, the dynamic informational value of it does not. Given the presence of IC constraints, even measuring such dynamic value is not straightforward, as one not only needs to consider the direct benefit to later consumers, but also needs to consider how better information may help to relax the IC constraints of later consumers and thereby facilitate more information generation from them. The duality approach I take partly resolves such difficulty by characterizing the shadow values of those IC constraints. Given the monotonicity property of $(\lambda_t^*)_{t=1}^T$, I show that when p_t is increased, the positive change in the myopic value of consumption always dominates the potentially indeterminate change in the dynamic informational value measured in the continuation problem of $\max_{\phi \in \Phi} \mathcal{L}(\phi; \lambda^*)$. The

[21] It is easy to see that the second phase includes at least the last period, since the optimal policy there will be myopically optimal. The first phase is non-empty as long as the prior on $\tilde{\theta}$ is not sufficiently favorable to support first-best learning.

[22] By saying $A \Rightarrow B$ almost surely (a.s.), I mean that the event in which A happens but B does not happen is of zero probability.

total value of consumption is thus always increasing in p_t, which implies the threshold structure of the optimal design.

Lemma 4 above also helps to characterize the *dictator's optimal policy*, where by "dictator" I mean a social planner who can dictate consumers' decisions without obeying their IC constraints. Notice that if $\lambda_t = 0$ for all t, the Lagrangian optimization $\max_{\phi\in\Phi} \mathcal{L}(\phi;\lambda)$ is reduced to the dictator's problem. Lemma 4 then implies that the dictator's problem also features a threshold solution. This solution will be used in the construction of the optimal design below.

3.4 The Optimal Policy

Notation: For any vector indexed by time, I will use subscription "$\geq t$" to indicate the sub-vector corresponding to time no earlier than t. For example, $\phi_{\geq t}$ will denote the continuation policy since time t. Notations like $\phi_{>t}$ and $\phi_{<t}$ are similarly defined.

Based on the two-phase structure of the optimal design implied by Lemma 3 and the threshold structure stated in Proposition 2, one can explicitly construct the optimal design using a forward induction algorithm. To do so, I define ϕ^d to be the "most conservative" optimal policy for the dictator's problem (i.e., the designer's problem without IC constraints), whose details are provided in Appendix B.5. When the dictator's problem admits multiple solutions, ϕ^d is the most conservative one in the sense that it always breaks ties in favor of non-recommendation.

A candidate optimal threshold policy ϕ^o, together with a cutoff time point $\widehat{t}$, can be inductively defined as follows.

Definition 2. A policy ϕ^o, a sequence of distributions $(\mu_t^o)_{t=1}^T$ over $[0,1]$ and a time point $\widehat{t} \in \{1, ..., T\}$ are defined with the following algorithm:

Start with $t = 1$ and define $\mu_1^o = \mu_1$.

- Step 1: If $\int_p [\phi_t^d(p)u(p)]\mu_t^o(dp) \geq 0$, then end the algorithm while defining:
  - $\widehat{t} = t$;
  - $\phi_{\geq t}^o = \phi_{\geq t}^d$;
  - $(\mu_s^o)_{s>t}$ to be the marginal distributions of $(p_s)_{s>t}$ under $\phi_{\geq t}^d$ given $p_t \sim \mu_t^o$.

  Otherwise, go to the next step.
- Step 2: Define ϕ_t^o to be a threshold time-t policy such that:[23]
  (i) $\phi_t^o(p) = 1$ for all $p > \bar{p}$;
  (ii) $\int_p [\phi_t^o(p)u(p)]\mu_t^o(dp) = 0$.

  Also define μ_{t+1}^o to be the distribution of p_{t+1} under ϕ_t^o given $p_t \sim \mu_t^o$. Then go back to step 1 with t replaced by $t + 1$.

[23]See Appendix B.5.2 for details of the construction of ϕ_t^o.

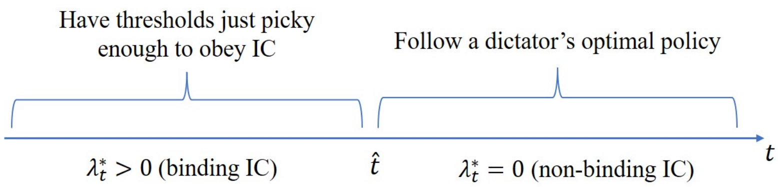

Figure 1: Two-Phase Structure of the Optimal Policy ϕ^o.

Intuitively, beginning with the initial period, the algorithm will progressively check in every period t whether it is incentive-compatible to follow the dictator's policy given $\phi^o_{<t}$ previously determined, i.e., whether $\int_p [\phi^d_t(p)u(p)]\mu^o_t(dp) \geq 0$. If so, the current period will be marked as period $\widehat{t}$, and ϕ^o will be set to follow ϕ^d for all later periods; otherwise, ϕ^o_t will be defined to have its threshold just high enough to satisfy the current consumer's IC constraint, which is what conditions (i) and (ii) in step 2 guarantee. The policy ϕ^o such defined will be in accordance with the two-phase structure of the optimal design found in Section 3.3, which is illustrated in Figure 1. In its early phase, the consumers' IC constraints will be binding and ϕ^o will be just "picky" enough to satisfy those constraints as equalities. In the later phase, the IC constraints will become non-restrictive and ϕ^o will follow the dictator's unconstrained optimal policy. The cutoff time $\widehat{t}$ is when ϕ^o shifts from the first phase to the second, after which it will follow ϕ^d.

The following proposition shows that ϕ^o is indeed an optimal policy and fully characterizes any optimal design.

Proposition 3. *Any policy ϕ^* is optimal for the designer's problem (5) – (6) if and only if: (i) $\phi^*_{<\widehat{t}}$ agrees with $\phi^o_{<\widehat{t}}$ almost surely; (ii) given $p_{\widehat{t}} \sim \mu^o_{\widehat{t}}$, $\phi^*_{\geq \widehat{t}}$ is incentive-compatible and is optimal for the dictator's continuation problem starting from time $\widehat{t}$. In particular, ϕ^o is optimal.*

The characterization in Proposition 3 makes it convenient to explore dynamic features of the optimal design and study how it should be tailored to market details. I pursue these in the following sections.

4 Dynamic Properties of the Optimal Design

4.1 Time Pattern of the Recommendation Standards

Threshold policies can be naturally interpreted as policies setting the minimum age-specific standards for a product to qualify for recommendations. Given the dynamic nature of the problem, it is conceivable that such a minimum standard should evolve over the product's life. I explore this time pattern below.

For the ease of exposition, I impose the following full-support and atomless assumption on the belief process.

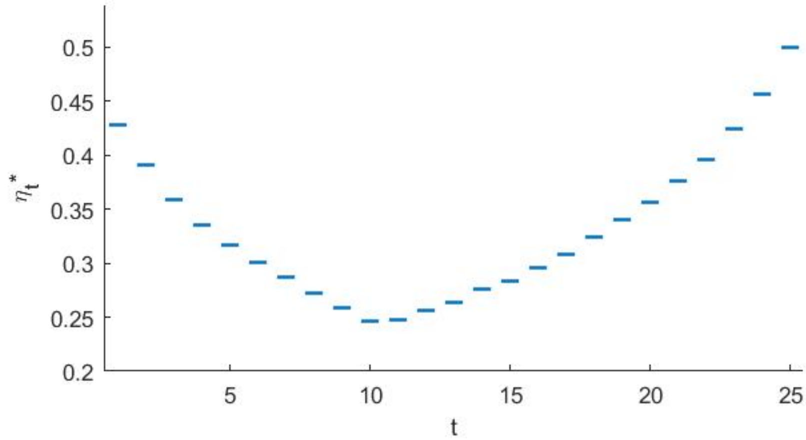

Figure 2: The optimal threshold policy for a numerical example: $T = 25$, $\theta_L = -1$, $\theta_H = 1$, $p_0 = 0.2$, $s_i|\theta \sim Normal(\theta, \sigma^2)$ with $\sigma = 3$ for all $i \geq 0$.

Assumption 2. The signals (s_0 and $\{s_i\}_{i\geq 1}$) are such that the marginal distributions of $p_1, ..., p_T$ are atomless and have full support over $(0, 1)$ under any policy.

The atomless assumption renders randomization at the threshold irrelevant, so we can solely focus on the thresholds themselves. The full-support assumption guarantees that any deviation in the recommendation threshold matters, which avoids the need to discuss off-path indeterminacy of the optimal policy. Assumption 2 holds, in particular, if the log-likelihood ratios of the signals $s_0, s_1, ...$ are continuous random variables with full support over $\mathbb{R}$.[24]

Recall that $\widehat{t}$ is the first time when it is incentive-compatible for ϕ^o to follow the dictator's optimal policy. The following result characterizes the time pattern of the optimal recommendation standards.

Proposition 4. *Under Assumption 2, the thresholds $(\eta^*_t)^T_{t=1}$ of any optimal threshold policy satisfy: (a) $\eta^*_{t-1} > \eta^*_t$ for all $t \leq \widehat{t} - 1$; (b) $\eta^*_t < \eta^*_{t+1}$ for all $t \geq \widehat{t}$. Moreover, $\eta^*_t \leq \bar{p}$ for any t.*

The proposition suggests that the optimal recommendation standard should first decrease and then increase over the product's life, which corresponds to the two phases with binding and non-binding IC constraints respectively. Figure 2 presents a numerical example. Underlying this result is a tension between the platform's desire to create dynamic informational value for future consumers and the need to fulfill the current consumer's IC constraint. In the early phase, the dynamic value is generally high since the product has a long future to go, but consumers are more "skeptical" about following

[24]That is, for both $i = 0$ and $i \geq 1$, s_i admits density functions f^L_i and f^H_i conditional on $\tilde{\theta} = \theta_L$ and $\tilde{\theta} = \theta_H$ respectively such that $\log\left(\frac{f^H_i(s_i)}{f^L_i(s_i)}\right)$ is a continuous random variable with full support over $\mathbb{R}$.

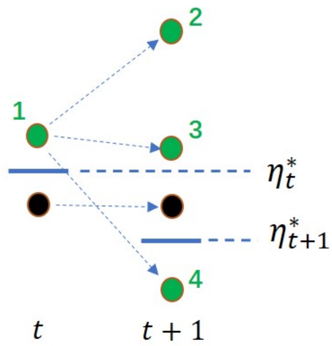

Figure 3: Explanation for decreasing recommendation thresholds before time $\widehat{t}$. The dots represent possible realizations of p_t or p_{t+1}. The arrows indicate the evolution of these beliefs. The blue bars represent the optimal recommendation thresholds.

recommendations because they know that even the platform has not acquired much information yet. This implies a binding IC constraint and necessitates a picky censorship over the beliefs eligible for recommendations. As time proceeds, consumers become easier to convince and the recommendation criterion can thus be relaxed. This continues until the standard is already sufficiently low such that trials with beliefs further below it are no longer worthwhile given the remaining time of the product. The IC constraint then turns non-restrictive. Thereafter, the optimal design just follows the dictator's optimal policy, where the recommendation standard gradually goes up since the dynamic value from myopically suboptimal consumption shrinks as the product approaches its end of life.

Figure 3 explains why the optimal threshold drops before time $\widehat{t}$ more precisely. The two dots on the left represent two possible realizations of p_t, and the blue bar between them represents the optimal threshold in period t. Since the green dot is above the threshold, it is associated with a recommendation and will hence split in a mean-preserving spread manner, which leads to realizations of p_{t+1} represented by the three new green dots in period $t + 1$. This reflects the new information generated by consumption. The black-dot belief in period t does not qualify for a recommendation and is thus carried over into period $t + 1$. Now, suppose the designer keeps the threshold unchanged over the two periods. Then in period $t + 1$, the lower green dot (dot 4) will be excluded from the recommendation region. Since belief evolves in a mean-preserving manner, the joint belief of those green dots remaining in the recommendation region (i.e., dots 2 and 3) will then be more favorable than their predecessor (dot 1). This implies that the consumer's IC constraint in period $t+1$ must turn slack. Intuitively, the better information in period $t + 1$ would have induced a more favorable selection for the consumer if the threshold were kept the same as before. This leaves room for the designer to also include the black dot into the recommendation region. When $t + 1 < \widehat{t}$, the designer indeed wants to do so since my previous characterization has shown that the consumer's IC constraint should

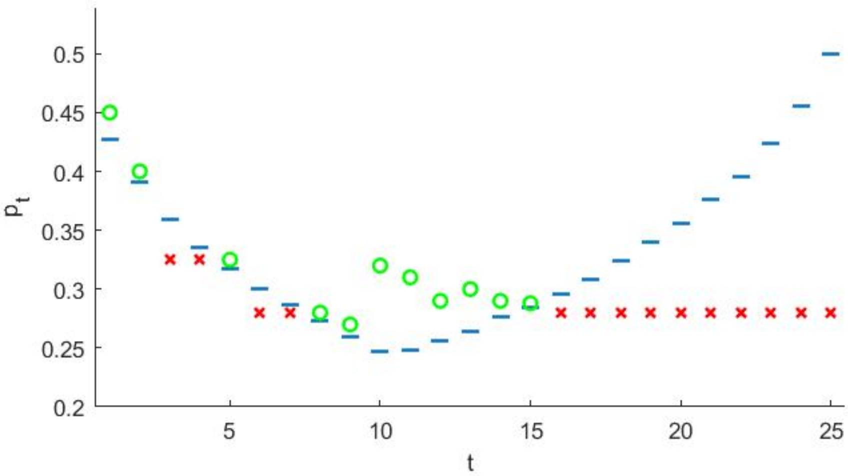

Figure 4: A realized path of recommendations under the optimal design in the numerical example of Figure 2. Blue bars represent the optimal thresholds. The crosses and circles track the platform's realized beliefs $(p_t)_{t=1}^T$, where green circles mean that recommendation is made and red crosses mean the opposite.

keep binding before time $\widehat{t}$. This implies that the threshold in period $t+1$ should be lowered.

The U-shaped pattern of recommendation thresholds has interesting implications for the optimal recommendation and learning dynamics, which I examine in the next subsection.

4.2 Optimal Recommendation Dynamic

Figure 4 demonstrates an example path of realized recommendations under the optimal policy in the numerical example of Figure 2.[25] The sequence of blue bars still represents the optimal age-specific thresholds. The series of crosses and circles tracks a realized path of $(p_t)_{t=1}^T$, where green circles mean that the current belief is above the current threshold and the product is hence recommended, while red crosses mean the opposite.

The figure highlights an interesting property of the optimal recommendation and learning dynamic – recommendation and learning can be *temporarily* suspended following negative feedback from the last consumption (e.g., the suspension in periods 3 - 4). Such kind of suspension is beneficial because it allows us to support myopically suboptimal exploration for products looking more promising at the same age (i.e., with a p_t being higher but still less than $\bar{p}$) without violating the IC constraint. However, the suspension may not need to last forever when further exploration is still socially desirable. As soon as the threshold for recommendation drops below the current belief, trials for the

[25]The figure presents a case where recommendation is eventually abandoned, which is more likely to happen when the product's true quality is low. However, the property discussed below does not rely on this.

product should be resumed. Notice that such a restart of recommendations can only happen in the early phase of the product's life, where the IC constraints are binding and the recommendation threshold decreases over time. If it were in the later phase, any suspension of recommendation would be permanent.

The learning dynamic described above and the associated inefficiency interestingly contrasts with those in the classic social learning literature.[26] In those early models, any stop of learning is permanent, and inefficiency arises because the stop is too early. Under the optimal design in the current setting, in comparison, stops can be temporary, and inefficiency can arise because these unwanted pauses slow down learning.[27]

In practice, *temporary* recommendation suspensions (or deprioritization) following negative consumer feedback are often used for punishing misconduct of a product supplier (e.g., a seller or content provider).[28] My finding here suggests another motivation for taking such actions, which is to induce social exploration on new products in a more efficient way given the inadequacy of individual consumer's incentive. Compared to those used for punishing supplier misconduct, the temporary recommendation suspensions in my model have two distinct features. First, they can happen following negative feedback regarding the product's innate features instead of unsatisfactory behaviors of the supplier. Second, they only happen for young products. Presuming that platforms in practice are indeed trying to enhance social learning about new products by tailoring their recommender systems, these may serve as concrete predictions of the model that can be tested with real data on platform recommendations.

5 Comparative Statics and Extensions

In this section, I provide comparative statics about how the optimal recommendation policy should be tailored to market details. When needed, the basic model will also be extended in certain ways.

5.1 Information Generation Rate

In many applications, having someone trying out a product is not guaranteed to generate meaningful information about the product's quality. The consumer may not leave feedback, or may give a piece of feedback that is too cursory to be authenticated.[29] These

[26]For classic papers on social experimentation, see, e.g., Bolton & Harris (1999) and Keller et al. (2005). For papers on observational learning, see, e.g., Banerjee (1992) and Bikhchandani et al. (1992).

[27]Under the optimal design in my setting, permanent stop of learning needs not be inefficient. In particular, any permanent stop happening in the later phase of the product's life is actually efficient because the thresholds there are the same as the dictator's optimal ones.

[28]See, e.g., Table 1 in Liang et al. (2020) for a list of examples.

[29]In particular, platforms like Amazon or Yelp often rely on textual analysis to filter out fake reviews. If a review is not material enough to pass such a test, it will be disregarded or attached with little weight

will lead to little post-consumption information generation.

How should the optimal design be adjusted when consumption becomes more likely to yield information due to, for example, better feedback elicitation design or improved satisfaction measurement technologies? To study this, I introduce an *information generation rate* of consumption into the model. Specifically, I assume that following one's consumption, the signal s_i being generated will be a compounded signal. It has probability $\alpha \in (0,1)$ to be an informative signal and has probability $1-\alpha$ to be uninformative, and the platform can tell which type the signal is.[30] Let $G^I(\cdot|\cdot)$ denote the transition kernel for the platform's belief following an informative signal. Given any α, the transition kernel for p_t following one's consumption then becomes:

$$G(\cdot|p) = \alpha G^I(\cdot|p) + (1-\alpha)D(\cdot|p) \tag{9}$$

The following proposition provides the comparative statics result with respect to the information generation rate α. For simplicity, I still impose the full-support and atomless assumption as in Section 4.1.

Proposition 5. *Assume Assumption 2 holds.*[31] *Given any α, let $(\eta_t^*(\alpha))_{t=1}^T$ denote the thresholds of the optimal threshold policy. Then $\alpha_a < \alpha_b \Rightarrow \eta_t^*(\alpha_a) \geq \eta_t^*(\alpha_b)\, \forall t$, where the inequality is strict for all $t \in (1,T)$.*

The proposition suggests that if the information generation rate is improved, the optimal recommendation standard should be lowered throughout the product's life. Roughly speaking, there are two forces behind this change. First, a higher α implies greater informational value from one's consumption, which motivates more exploration. Second, with a higher α, information from past consumption is accumulated at a faster rate even if the thresholds remain the same as before. This enables better-informed recommendations at any time, which makes consumers more willing to follow the recommendations ceteris paribus. We thus have room to lower the recommendation standards without violating the IC constraints. Together, these lead to looser recommendation criteria in the optimal design.

The formal proof of the proposition is technically involved because it requires comparing two controlled Markov processes corresponding to different α. Central to the proof is a coupling argument, where I explicitly construct the belief processes under the optimal designs corresponding to different α in the same probability space. This allows a direct comparison between them. I refer interested readers to the proof of Observation 3 in Appendix B.7.

in any recommendation algorithm. This is important to deter fake reviews.

[30] This compounded signal satisfies Assumption 1 as long as the informative component does.

[31] This holds, in particular, if the log-likelihood ratios of s_0 and subsequent *informative* signals are continuous random variables, and the log-likelihood ratio of s_0 has full support over $\mathbb{R}$.

5.2 Consumer Arrival Frequency

In practice, different platforms and different products have different frequencies of potential customer arrivals. A platform with a larger user base is expected to have more visits per unit of time than one with a smaller user base; a mass-market product is expected to have greater potential customer flow than a niche product. How should the optimal design be tailored to these market heterogeneity? To answer this question, I extend my framework to accommodate random consumer arrivals below, and then examine how the optimal design should depend on the arrival rate of consumers.

Formally, I modify the model in Section 2 by assuming that in every period a consumer will arrive with probability $\rho \in (0,1)$. The arrivals are independent over time and independent from other random objects in the model. Accordingly, I re-interpret a_t as the consumption decision (or the platform's recommendation) of an arriving consumer. Then, a signal of quality will be generated following period t if and only if $a_t = 1$ and a consumer does arrive in that period. The transition rule of the belief process $(p_t)_{t=1}^T$ then becomes:

$$p_{t+1}|p_t, a_t \sim a_t\big[\rho G(\cdot|p_t) + (1-\rho)D(\cdot|p_t)\big] + (1-a_t)D(\cdot|p_t) \tag{10}$$

Compared to the transition rule in (2), the change is that $G(\cdot|\cdot)$ is now replaced by $\rho G(\cdot|\cdot) + (1-\rho)D(\cdot|\cdot)$. This is the only change we have in the design environment.[32]

Because Lemma 1 still holds with $G(\cdot|\cdot)$ replaced by $\rho G(\cdot|\cdot) + (1-\rho)D(\cdot|\cdot)$, all of my previous characterizations for the optimal design will remain valid. The following proposition reveals how the consumer arrival frequency matters for the optimal design.

Proposition 6. *Assume Assumption 2 holds.*[33] *Given any arrival rate ρ, let $(\eta_t^*(\rho))_{t=1}^T$ denote the thresholds of the optimal threshold policy. Then $\rho_a < \rho_b \Rightarrow \eta_t^*(\rho_a) \geq \eta_t^*(\rho_b)\, \forall t$, where the inequality is strict for all $t \in (1,T)$.*

The proposition suggests that when the platform faces a thicker market where consumers arrive more frequently, the recommendation standards over the product's life should be lower. The intuition is again twofold. First, a higher arrival rate implies that more consumers are likely to come in the future. This increases the informational value of early consumption. Second, with a higher arrival rate, the platform will in expectation have more consumption and hence more signals accumulated before any given period if the standards are kept unchanged. This allows the platform to indeed implement lower recommendation standards while obeying the IC constraints.

[32] Another change to the designer's problem is that we should now replace $u(\cdot)$ in the designer's objective function (5) with $\rho u(\cdot)$, which reflects the fact that a consumer arrives only with probability ρ. This does not matter for the optimization, however, since it only rescales the objective function.

[33] This still holds, in particular, under the conditions in footnote 24, but now there is no need to require $\log\big(\frac{f_i^H(s_i)}{f_i^L(s_i)}\big)$ $(i \geq 1)$ to have full support over $\mathbb{R}$.

5.3 Platform Bias

In some applications, the platform may be inherently biased towards recommending the product it carries instead of willing to maximize consumer surplus, especially when it earns extra commission from consumption. My analysis can be easily extended to incorporate this interest conflict between the platform and consumers.

Formally, we can modify the main model by assuming that the platform will earn commission $\beta \geq 0$ whenever a consumer consumes the product, and it wants to maximize the sum of its commission and consumer surplus. The platform's problem then becomes:

$$\max_{\phi \in \Phi} \Big\{ \sum_{t=1}^{T} \mathbb{E}_\phi \big[a_t \big(u(p_t) + \beta \big) \big] \Big\} \tag{11}$$

$$\text{s.t. } \mathbb{P}_\phi(a_t = 1) > 0 \Rightarrow \mathbb{E}_\phi[u(p_t)|a_t = 1] \geq 0 \tag{12}$$

$$\mathbb{P}_\phi(a_t = 0) > 0 \Rightarrow \mathbb{E}_\phi[u(p_t)|a_t = 0] \leq 0 \tag{13}$$

As in the basic model, IC constraint (13) can be ignored because the platform prefers consumption more than the consumers. Through the same argument as in Section 3.1, the optimization can then be written as:

$$\max_{\phi \in \Phi} \Big\{ \sum_{t=1}^{T} \mathbb{E}_\phi \big[a_t \big(u(p_t) + \beta \big) \big] \Big\} \tag{14}$$

$$\text{s.t. } \mathbb{E}_\phi[a_t u(p_t)] \geq 0 \quad \forall t = 1, ..., T \tag{15}$$

The only difference from the optimization in Section 3 is that we now have the β term in the objective function measuring the platform's bias. It is easy to check that all proofs in Section 3 can be extended straightforwardly. Especially, the optimal policy will still be characterized by Proposition 3, with ϕ^o being defined by the algorithm in Definition 2. The analysis in Section 4 also remains qualitatively the same, which implies that the optimal recommendation policy generally features U-shaped standards over time.

To see how β shapes the optimal design, notice that it affects the dictator's optimal solution ϕ^d. Intuitively, the larger is β, the more willing the dictator will be to recommend the product and thus the lower thresholds ϕ^d will have. Since ϕ^d is a key ingredient in constructing the optimal policy in Definition 2, β then influences the optimal design. This is actually the only channel through which β matters. The following proposition provides comparative statics with respect to changes in β.

Proposition 7. *Assume Assumption 2 holds. For any $\beta \geq 0$, let $(\eta_t^*(\beta))_{t=1}^T$ denote the thresholds of the optimal threshold policy, and let $\widehat{t}(\beta)$ denote the cutoff time defined in Definition 2. Then, given any $\beta_b > \beta_a \geq 0$, we have:*

(a) $\widehat{t}(\beta_b) \geq \widehat{t}(\beta_a)$

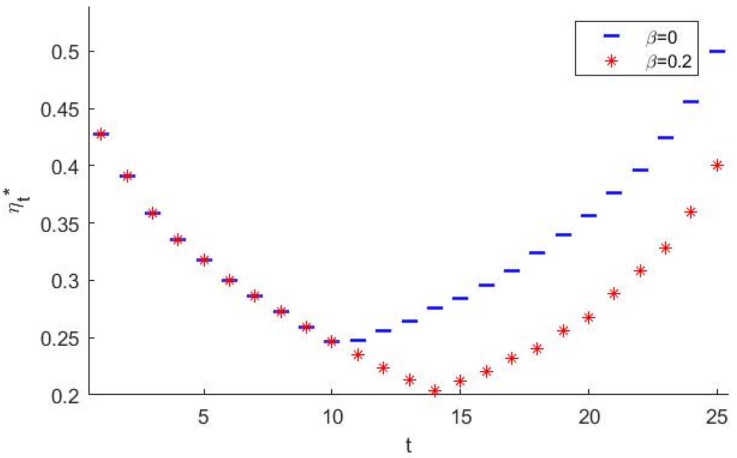

Figure 5: Optimal recommendation thresholds with $\beta = 0$ and $\beta = 0.2$ respectively, for the numerical example in Figure 2.

(b) $\eta^*_t(\beta_b) = \eta^*_t(\beta_a)$ *for all* $t < \widehat{t}(\beta_a)$

(c) $\eta^*_t(\beta_b) \leq \eta^*_t(\beta_a)$ *for all* $t \geq \widehat{t}(\beta_a)$*, where the inequality holds strictly for all* $t > \widehat{t}(\beta_a)$*.*

Figure 5 illustrates the proposition's results. Following an increase in β, the optimal thresholds will remain the same as before during the phase where the thresholds were originally decreasing, but will then go deeper down after the original bottom and thereafter remain lower than the original thresholds. Intuitively, after β increases, the platform would like to set lower standards and thereby earn more commission. However, during the initial periods (before period 10 in the figure), the consumer's IC constraints have already been binding. The platform can thus do nothing but keeping the original standards. During the later periods, on the contrary, the consumers' IC constraints were originally slack. This enables the platform to implement lower standards.

6 Additional Discussion

6.1 Non-binary Quality Levels

One simplifying assumption in my main model is that the product's quality can only take binary values. As in many other studies on information design or social learning, this allows one to represent the evolving belief with a single-dimensional variable, which significantly eases the analysis.[34]

In Appendix A, I extend the model to allow general quality support. Although a full characterization of the optimal design is not available, the duality approach does help to extend certain structures of the optimal design to that general setting. In particular, I show that the optimal design still features a two-phase structure implied by Lemma 3.

[34] See, e.g., section 2 in Hörner & Skrzypacz (2017) for papers on social experimentation.

Moreover, the optimal policy should be more inclined to recommend the product when the platform's current belief about quality is higher in the likelihood-ratio order. This extends the threshold structure to the case with multi-dimensional beliefs. I will discuss how this result can be helpful for algorithmic recommendation design in the appendix.

6.2 Comparison to Previous Studies

As has been mentioned in the introduction, my study is closely related to Kremer et al. (2014) and Che & Hörner (2018), who also study how platform recommendations can improve social learning efficiency when early consumption produces information for later consumers through the platform. Similar to my paper, they also reveal how past information accrued to the platform enables it to persuade more consumers into socially desirable explorations in the future. However, the three papers provide different characterizations for the optimal design.

Because Kremer et al. (2014) assumes fully revealing consumption-generated signals, their design is mainly about when to induce the first trial of the product.[35] The main result is that a product that looks better based on the platform's initial information (in its quality relative to an alternative option) should receive the first trial earlier. Che & Hörner (2018) assumes that the platform learns from conclusive news that fully reveals the product's quality upon its arrival. The design in their paper is hence about "how much" to recommend the product following the history without news arrival. They show that myopically suboptimal recommendation, given no news arrival, should gain increasing intensity as the product ages until being ceased at some point. In contrast, my study accommodates general non-conclusive consumption-generated signals. My prediction of the optimal design is therefore about whether to recommend the product in each period based on any current belief of the platform. This allows me to interpret my results as being about the time-varying recommendation standards. In particular, I have found that the optimal design features threshold policies with respect to the evolving belief, and the recommendation standard should vary in a U-shaped pattern as the product ages.[36]

Allowing non-conclusive signals also enables richer predictions about the optimal recommendation dynamic. In particular, the phenomenon of temporary recommendation suspensions following negative consumer feedback in Section 4.2 cannot exist with conclusive signals, since conclusive negative feedback should stop recommendation forever. Moreover, both of the previous papers suggest that exploration (i.e., myopically suboptimal trials) should stop after some middle age of the product. With general consumption-generated signals, however, Proposition 4 implies that exploration can happen until the

[35] After the first trial, quality is fully revealed and recommendations should be myopically optimal.

[36] With binary product quality, the result in Kremer et al. (2014) can be considered as a special case of mine.

last period of the product's life, although the belief region for exploration gradually shrinks after some point.

Finally, my study has also provided several comparative statics regarding the optimal recommendation standards in Section 5, which have no counterparts in the previous papers.

7 Conclusion and Methodological Remarks

I have studied the optimal design of platform recommendations when early consumption of a product yields informational externality to later consumers. The optimal design is shown to feature threshold policies with the recommendation standard varying in a U-shaped pattern over the product's life. An interesting implication is that recommendations can be temporarily suspended following negative feedback for young products. My characterization also enables a couple of comparative statics. In particular, I have shown that the recommendation standard should be lower throughout the product's life when consumption is more likely to generate informative feedback or when consumers are arriving more frequently over time.

My model accommodates non-conclusive consumption-generated signals. Consequently, compared to the existing literature, it requires a more general formulation of the designer's problem as a *constrained* Markov decision process. I argue that such a mathematical formulation naturally arises in dynamic information design problems where the designer's private information flow is controlled by the receivers' decisions. Specifically, if one focuses on direct mechanisms, then those design problems can be treated as Markov decision processes where the designer decides the receivers' actions and thereby controls his own information flow subject to the receivers' IC constraints. Since the receivers do not observe the designer's information, their IC constraints need to involve taking expectations over it. This then leads to the aggregated constraints that cannot be directly handled in dynamic programming (like constraint (6)). I expect that such a formulation and the Lagrangian duality approach I take can be generally useful for this kind of dynamic information design problems.

A Non-binary Quality Levels

In this appendix, I extend the model to allow for non-binary product quality and generalize certain characterizations of the optimal design.

A.1 The General Setting

Consider the same setting as in Section 2.1 except that the support of $\tilde{\theta}$ can be an arbitrary set $\Theta \subset \mathbb{R}$ now. I assume that the joint distribution of $\tilde{\theta}$ and the signals $(s_0, s_1, ...)$ is such that the platform's posterior belief is always within a family of distributions $\{Q_z\}_{z\in Z}$, where $Z \subset \mathbb{R}^n$ is a countable parameter set.[37] I assume that $\{Q_z\}_{z\in Z}$ admits density functions $\{q_z\}_{z\in Z}$ with respect to some common dominating measure over $\mathbb{R}$, and $q_z(\cdot) > 0$ on Θ for all $z \in Z$. Moreover, I impose the following assumption:

Assumption A.1. Assume the following:
(i) $\mathbb{P}(\mathbb{E}[\tilde{\theta}|s_0] > 0) > 0$.
(ii) For any $i \geq 1$, s_i takes values in some set $S \subset \mathbb{R}$. Its distribution conditional on $\tilde{\theta}$ admits a conditional density function $\ell(\cdot|\cdot)$ (w.r.t. some dominating measure over $\mathbb{R}$) such that: $\ell(s|\theta) > 0$ for all $s \in S$ and $\theta \in \Theta$; $\ell(\cdot|\theta)$ increases in θ in the likelihood-ratio order.

Condition (i) plays the same role as its counterpart in Assumption 1. Condition (ii) implies that higher realizations of s_i suggest that the product is more likely to be of higher quality. This framework is general enough to incorporate many parametric learning models with congruent prior and signals (e.g., the Beta-Binomial model). Moreover, it accommodates any learning model with finite support of $\tilde{\theta}$ that satisfies Assumption A.1.

Let $z_t \in Z$ denote the platform's belief parameter at the beginning of period t. Since z_t (or Q_{z_t}) summarizes all information available to the platform at time t, we can focus on (randomized) Markov recommendation policies w.r.t. $(z_t)_{t=1}^T$. Formally, any policy of this type is a sequence of measurable mappings $\phi := (\phi_t : t = 1, ..., T)$, where each $\phi_t : Z \to [0, 1]$ decides the probability of recommending the product at time t given any belief parameter $z_t \in Z$.

As in Section 5.2, I also allow i.i.d. random consumer arrivals and use ρ to denote the arrival probability.

[37] I assume Z to be countable for simplicity, which avoids the need to discuss certain measurability issues about the optimal policy.

A.2 Characterizations

Following similar arguments as those in Section 3.1 and Section 5.2, the designer's problem can be formulated as follows:

$$\max_{\phi\in\Phi}\Big\{\sum_{t=1}^{T}\mathbb{E}_{\phi}[a_t u(z_t)]\Big\}$$

$$\text{s.t. } \mathbb{E}_{\phi}[a_t u(z_t)] \geq 0 \quad \forall t = 1, ..., T$$

$$z_{t+1}|z_t, a_t \sim a_t\big[\rho G(\cdot; z_t) + (1-\rho)D(\cdot; z_t)\big] + (1-a_t)D(\cdot; z_t)$$

$$z_1 \sim \mu_1$$

where $u(z_t) := \int_{\theta\in\Theta}\theta Q_{z_t}(d\theta)$ (i.e., the expected consumption surplus given belief Q_{z_t}), $G(\cdot;\cdot)$ is the transition kernel for z_t following one's consumption, and μ_1 is the distribution of z_1. Compared to the model with binary qualities, the process of $(z_t)_{t=1}^T$ now replaces the role of $(p_t)_{t=1}^T$. I define the Lagrangian function $\mathcal{L}(\phi;\lambda)$ and the dual problem similarly as those in Section 3. The following lemma follows easily from my assumptions and the definition of $(z_t)_{t=1}^T$.

Lemma A.1. *The belief (parameter) process has the following properties:*
(P1') $\big[\rho G(\cdot; z) + (1-\rho)D(\cdot; z)\big]$ *as a measure-valued function of* z *is weakly continuous.*[38]
(P2') $\int_{z'} u(z')\big[\rho G(dz'; z) + (1-\rho)D(dz'; z)\big] = u(z)$.
(P5') $\mu_1\big(\{z : u(z) > 0\}\big) > 0$.

These properties are the counterparts to properties (P1), (P2) and (P5) in Lemma 1. Because Lemmas 2 and 3 in the main text only rely on these properties in Lemma 1, they still hold in the current setting.[39] In particular, we still have the following time pattern of Lagrangian multipliers derived from the dual problem:

Lemma A.2. *There exists* $\lambda^* \in \arg\min_{\lambda\in\mathbb{R}_+^T}\sup_{\phi\in\Phi}\mathcal{L}(\phi;\lambda)$ *such that* $\lambda_t^* \geq \lambda_{t+1}^*$ *for all* $t = 1, ..., T-1$.

One particular implication of this lemma is that the multiplier will stay at zero once dropping to it. As in the main model, this implies that any optimal design generally features a two-phase structure. In the first phase, IC constraints are binding and hence the recommendation policy is just picky enough for the consumers to follow; in the second phase, the IC constraints become non-restrictive and the optimal design follows the optimal continuation policy of the dictator.

While I cannot fully pin down the optimal design, the proposition below reveals an important structure of it. Let $\geq_{LR}$ denote dominance in the likelihood-ratio order.

[38] This is trivially true since Z is countable.

[39] The proofs for them remain the same as before except that the role of $(p_t)_{t=1}^T$ is now replaced by $(z_t)_{t=1}^T$.

Proposition A.1. *Assume the consumer arrival rate* $\rho \in (0,1)$.[40] *Any optimal policy is almost surely equivalent to some policy* ϕ^* *such that for all* t*: if* $Q_{z'} \geq_{LR} Q_z$ *and* $\int_\theta \theta dQ_{z'}(\theta) > \int_\theta \theta dQ_z(\theta)$*, then* $\phi_t^*(z) > 0 \Rightarrow \phi_t^*(z') = 1$.

Intuitively, the proposition roughly suggests that any optimal policy should be more inclined to recommend the product when the current belief of quality is higher in the likelihood-ratio order. This naturally extends the threshold structure of the optimal design in my main model to the current setting, where the platform's belief is in a multi-dimensional space only endowed with a partial order.

Although the aforementioned property is intuitively appealing, it is actually violated by many recommendation algorithms proposed in studies that only focus on the design's asymptotic performance. For example, the algorithm in Mansour et al. (2020) introduces randomized exploration to fulfill the consumers' IC constraints, which necessarily violates the property.[41] My result suggests that modifying their algorithm to be more consistent with this property may help to improve the algorithm's finite-horizon performance. This may be an interesting topic for future algorithm-oriented research.

B Proofs

B.1 Proof for Lemma 1

Proof. Property (P2) is implied by the law of iterated expectation; property (P4) is obvious given Assumption 1(ii) (i.e., the signals are not completely uninformative); property (P5) is directly implied by Assumption 1(i). I show (P1) and (P3) below.

Fix any $i \geq 1$. Let S denote the signal realization space of s_i. Let $f_L(\cdot)$ and $f_H(\cdot)$ denote s_i's conditional density functions conditional on $\tilde{\theta} = \theta_L$ and $\tilde{\theta} = \theta_H$ respectively, with respect to some dominating measure m over S. Without loss of generality, we can choose m s.t. $f_L(s)$ and $f_H(s)$ are not both equal to zero m-a.s. Since I assume no signal realization fully reveals the value of $\tilde{\theta}$ (i.e., Assumption 1(iii)), we also have $f_L(s) \neq 0 \Leftrightarrow f_H(s) \neq 0$ m-a.s. Thus $f_L(s)$ and $f_H(s)$ are non-zero m-a.s. Define the log-likelihood ratio $\ell_i = \log\big(f_H(s_i)/f_L(s_i)\big)$, and let J_L and J_H denote its distribution given $\tilde{\theta} = \theta_L$ and $\tilde{\theta} = \theta_H$ respectively.[42]

We have the following observation:

Claim *(a).* For any a, $J_H(a) = \int_{\ell \leq a} e^\ell dJ_L(\ell)$.

[40] Although I conjecture that the result should also hold with $\rho = 1$, my current proof requires $\rho < 1$ to avoid some technical subtly.

[41] The algorithm in Mansour et al. (2020) is not Markovian. Hence, more precisely, it is the randomized Markov policy outcome-equivalent to their algorithm that does not satisfy the property.

[42] Such log-likelihood ratio representation of a signal has been previously used in Smith & Tian (2018).

Proof for Claim (a). The following equalities hold:

$$\begin{aligned}
&\int_{\ell\le a} e^{\ell} dJ_L(\ell) \overset{①}{=} \mathbb{E}[\mathbb{1}_{\{\ell_i\le a\}} e^{\ell_i}|\tilde{\theta}=\theta_L] \overset{②}{=} \mathbb{E}\Big[\mathbb{1}_{\left\{\log\frac{f_H(s_i)}{f_L(s_i)}\le a\right\}} \frac{f_H(s_i)}{f_L(s_i)}\Big|\tilde{\theta}=\theta_L\Big] \\
&\overset{③}{=} \int \mathbb{1}_{\left\{\log\frac{f_H(s)}{f_L(s)}\le a\right\}} \frac{f_H(s)}{f_L(s)} f_L(s)m(ds) \overset{④}{=} \int \mathbb{1}_{\left\{\log\frac{f_H(s)}{f_L(s)}\le a\right\}} f_H(s)m(ds) \\
&\overset{⑤}{=} \mathbb{E}[\mathbb{1}_{\{\ell_i\le a\}}|\tilde{\theta}=\theta_H] \overset{⑥}{=} J_H(a)
\end{aligned}$$

where the first and the last equalities hold by the definitions of J_L and J_H respectively; the second equality holds by the definition of ℓ_i; the third and the fifth equalities hold by the definitions of f_L and f_H respectively; the fourth equality is a trivial identity. □

Now, given any prior belief p about $\tilde{\theta}=\theta_H$, let $\tilde{p}$ denote the posterior belief given s_i. Then by the Bayes rule we have: $\log\frac{\tilde{p}}{1-\tilde{p}} = \log\frac{p}{1-p} + \ell_i$. Let $\mathbb{P}_p$ denote the probability measure given prior p. This then implies that

$$\mathbb{P}_p(\tilde{p}\le x) = \mathbb{P}_p\big(\ell_i \le \log\frac{x}{1-x} - \log\frac{p}{1-p}\big) \tag{B.1}$$

$$= pJ_H\big(\log\frac{x}{1-x} - \log\frac{p}{1-p}\big) + (1-p)J_L\big(\log\frac{x}{1-x} - \log\frac{p}{1-p}\big) \tag{B.2}$$

$$= \int_{\ell\le\log\frac{x}{1-x}-\log\frac{p}{1-p}} [pe^{\ell} + (1-p)]dJ_L(\ell) \tag{B.3}$$

where the last equality holds by Claim (a) above. Now, pick any $p^*\in\mathbb{R}$ and a sequence of $(p_n)_n \to p^*$. When expression (B.3) is continuous in x at $x=x_0$ given $p=p^*$, obviously we must have $J_L(\ell)$ being continuous at $\ell = \log\frac{x_0}{1-x_0} - \log\frac{p^*}{1-p^*}$, which further implies that the expression (B.3) is continuous in p at $p=p^*$ given $x=x_0$. Thus $\mathbb{P}_{p^*}(\tilde{p}\le x)$ being continuous in x at $x=x_0$ implies $\mathbb{P}_{p_n}(\tilde{p}\le x_0)\to\mathbb{P}_{p^*}(\tilde{p}\le x_0)$ as $n\to\infty$. Therefore, the distribution of $\tilde{p}$ given prior p is weakly continuous in p. This proves the weak continuity condition for $G(\cdot|p)$ in (P1).

To check property (P3), we need the following observation:

Claim *(b).* For any a, $\int_{\ell\le a} e^{\ell} dJ_L(\ell) \le \int_{\ell\le a} dJ_L(\ell)$.

Proof for Claim (b). Notice by Claim (a) above, $e^{\ell}dJ_L(\ell)$ just equals to $dJ_H(\ell)$. We can thus treat both $e^{\ell}dJ_L(\ell)$ and $dJ_L(\ell)$ as probability measures over $\mathbb{R}$, with densities e^{ℓ} and 1 respectively w.r.t. the dominating measure $dJ_L(\ell)$. Since e^{ℓ} is increasing in ℓ, we then have $e^{\ell}dJ_L(\ell)$ dominating $dJ_L(\ell)$ in the likelihood-ratio order.[43] This further implies dominance in first-order stochastic dominance and thus $\int_{\ell\le a} e^{\ell}dJ_L(\ell) \le \int_{\ell\le a} dJ_L(\ell)$. □

[43]See, e.g., section 1.4 in Müller & Stoyan (2002) for an introduction to such order.

Now, pick any p_a and p_b s.t. $p_a < p_b$, we have

$$\int_{\ell \le \log \frac{x}{1-x} - \log \frac{p_b}{1-p_b}} [p_b e^{\ell} + (1-p_b)] dJ_L(\ell) \le \int_{\ell \le \log \frac{x}{1-x} - \log \frac{p_a}{1-p_a}} [p_b e^{\ell} + (1-p_b)] dJ_L(\ell)$$
$$\le \int_{\ell \le \log \frac{x}{1-x} - \log \frac{p_a}{1-p_a}} [p_a e^{\ell} + (1-p_a)] dJ_L(\ell)$$

where the second inequality holds due to Claim (b). Together with equations (B.1)–(B.3), this implies that $\mathbb{P}_p(\tilde{p} \le x)$ is weakly decreasing in p for any x. Thus we have the property of (P3).

Q.E.D.

B.2 Proof for Proposition 1

Proof. The proof basically applies Lemma 1(iv) in Feinberg & Piunovskiy (2000) to my setting. Specifically, define

$$\mathcal{V} = \{v \in \mathbb{R}^{T+1} : \exists \phi \in \Phi \text{ s.t. } v_t = \mathbb{E}_\phi[a_t u(p_t)] \, \forall t = 1, ...T \text{ and } v_{T+1} = \sum_{t=1}^{T} \mathbb{E}_\phi[a_t u(p_t)]\}$$

Notice for each admissible policy ϕ, the first T arguments of the corresponding vector v are the values of the IC constraints and the $(T+1)$'th argument of v is just the total surplus in the designer's objective. Lemma 1(iv) in Feinberg & Piunovskiy (2000) implies that $\mathcal{V}$ is a compact set. This further implies that the set $\mathcal{V} \cap \{v \in \mathbb{R}^{T+1} : v_t \ge 0 \, \forall t = 1, ...T\}$ is compact, and thus when we maximize over its $(T+1)$'th dimension, the supremum is achievable. By the definition of $\mathcal{V}$, this is equivalent to that the designer's problem has its supremum achieved by some ϕ.

Now, it suffices to check that the four conditions of Lemma 1 in Feinberg & Piunovskiy (2000) are indeed satisfied in my setting. Condition 1 holds because my state space $[0, 1]$ is closed, and the set of feasible actions $A = \{0, 1\}$ is finite and does not vary in time and state. Conditions 2 and 4 hold because the flow payoffs in my setting are bounded and continuous in the pair of action and state, and is non-zero for only finitely many periods. For Condition 3, we just need to show the transition probability $aG(\cdot|p) + (1-a)D(\cdot|p)$ is weakly-continuous in $(a, p) \in \{0, 1\} \times [0, 1]$. With $\{0, 1\}$ endowed with the discrete topology, it suffices to check weak continuity in p when $a = 1$ and $a = 0$ separately. These are respectively implied by the weak continuity of $G(\cdot|p)$ (Property (P1) in Lemma 1) and $D(\cdot|p)$ in p.[44]

Q.E.D.

[44] To see $D(\cdot|p)$ is weakly-continuous in p, notice its cdf is just $\mathbb{1}_{\{x \ge p\}}$. Given any sequence $(p_n)_n \to p^*$, we have $\mathbb{1}_{\{x \ge p_n\}} \to \mathbb{1}_{\{x \ge p^*\}}$ for any $x \ne p^*$. The weak-continuity is thus implied (see, e.g., Section 3.2 in Durrett (2019) for conditions of weak-continuity).

B.3 Proof for Lemma 2

Proof. To use the Lagrangian duality theorem, I first transform the designer's problem into a linear program. Throughout, I fix the initial belief state distribution μ_1. Given any policy $\phi \in \Phi$, let m_t^ϕ denote the distribution of (a_t, p_t) under it.[45]

Let $\mathcal{M}$ denote the set of all sequences of such distributions under some ϕ, i.e., $\mathcal{M} = \{(m_t^\phi)_{t=1}^T : \phi \in \Phi\}$. A (standard) characterization for this set is that $(m_t)_{t=1}^T \in \mathcal{M}$ if and only if:

$$m_1(\{0,1\} \times B) = \mu_1(B) \tag{B.4}$$

$$m_{t+1}(\{0,1\} \times B) = \int_{p\in[0,1]} \sum_{a\in\{0,1\}} [aG(B|p) + (1-a)D(B|p)] m_t(a, dp)\ \forall t = 1, ..., T-1 \tag{B.5}$$

for any $B \in \mathcal{B}_{[0,1]}$ (Borel σ-field of $[0,1]$). I use $\widehat{\mathcal{M}}$ to denote the set of $(m_t)_{t=1}^T$ satisfying these conditions. The fact that $\mathcal{M} \subset \widehat{\mathcal{M}}$ is obvious since any $(m_t^\phi)_{t=1}^T$ must be consistent with μ_1 and the transition probabilities, and thus satisfies (B.4) and (B.5).

To see $\widehat{\mathcal{M}} \subset \mathcal{M}$, pick any $(m_t^*)_{t=1}^T \in \widehat{\mathcal{M}}$. Let ϕ^* be a (randomized) Markov policy such that ϕ_t^* is just the conditional probability mass function of a_t given p_t under m_t^*. Formally, for any m_t, treat $m_t(a, dp)$ $(a = 1, 2)$ as a measure over $[0,1]$ s.t. $m_t(a, B) = m_t(\{a\} \times B)$, $\forall B \in \mathcal{B}_{[0,1]}$. Then ϕ^* is defined as (an arbitrary version of) the Radon-Nikodym derivative of $m_t^*(1, dp)$ w.r.t. $m_t^*(0, dp) + m_t^*(1, dp)$. (Notice $m_t^*(0, dp) + m_t^*(1, dp)$ is just the marginal distribution of m_t^* over $[0,1]$ and the Radon-Nikodym derivative is by definition measurable.) Then, we can show ϕ^* implements $(m_t^*)_{t=1}^T$ by induction in t. Let $m_t^{\phi^*}$ denote the joint distribution of (a_t, p_t) under ϕ^* for any t. For $t = 1$, we have for all $B \in \mathcal{B}_{[0,1]}$:

$$\begin{aligned} m_1^{\phi^*}(\{1\} \times B) = \int_{p\in B} \phi_1^*(p)\mu_1(dp) &= \int_{p\in B} \phi_1^*(p)[m_1^*(0, dp) + m_1^*(1, dp)] \\ &= \int_{p\in B} m_1^*(1, dp) = m_1^*(\{1\} \times B) \end{aligned}$$

where the second equality holds by condition (B.4) and the third equality holds by the definition of ϕ^*. Since $m_1^{\phi^*}(\{1\} \times B) + m_1^{\phi^*}(\{0\} \times B) = \mu_1(B) = m_1^*(\{1\} \times B) + m_1^*(\{0\} \times B)$, we also have $m_1^{\phi^*}(\{0\} \times B) = m_1^*(\{0\} \times B)$. Thus $m_1^{\phi^*} = m_1^*$.

Now, assume $m_t^{\phi^*} = m_t^*$ and consider the result for $t + 1$. Because condition (B.5) holds for m_{t+1}^*, we know that the marginal distribution over $[0,1]$ under m_{t+1}^* given m_t^* is

[45] As is standard, we can construct the underlying measurable space for the process as $(\{0,1\} \times [0,1])^T$ with the Borel σ-field, and treat the corresponding random variables as identity mappings $(\{0,1\} \times [0,1])^T \to (\{0,1\} \times [0,1])^T$. Thus we can treat any distribution for those random variables equivalently as a measure over the underlying measurable space, which is typically how I interpret those distributions.

determined by the same rule as that determines the marginal distribution over $[0,1]$ under $m^{\phi^*}_{t+1}$ given $m^{\phi^*}_t$. Thus $m^*_t = m^{\phi^*}_t$ implies $m^*_{t+1}(\{0,1\}\times B) = m^{\phi^*}_{t+1}(\{0,1\}\times B)$, $\forall B \in \mathcal{B}_{[0,1]}$. This further implies:

$$m^{\phi^*}_{t+1}(\{1\} \times B) = \int_{p\in B} \phi^*_{t+1}(p)[m^{\phi^*}_{t+1}(0,dp) + m^{\phi^*}_{t+1}(1,dp)]$$
$$= \int_{p\in B} \phi^*_{t+1}(p)[m^*_{t+1}(0,dp) + m^*_{t+1}(1,dp)] = m^*_{t+1}(\{1\} \times B)$$

where the second equality holds because the two measures are equal as mentioned right above and the third equality holds by the definition of ϕ^*. Together with $m^*_{t+1}(\{0,1\} \times B) = m^{\phi^*}_{t+1}(\{0,1\} \times B)$, this also implies $m^{\phi^*}_{t+1}(\{0\} \times B) = m^*_{t+1}(\{0\} \times B)$. Therefore, $m^*_{t+1} = m^{\phi^*}_{t+1}$. This completes the induction proof for showing that m^* is implemented with ϕ^*.

The above discussion has shown $\mathcal{M} = \widehat{\mathcal{M}}$. We can thus rewrite the designer's problem as

$$\max_{(m_t)_{t=1}^T \in \widehat{\mathcal{M}}} \Big\{ \sum_{t=1}^T \Big[\int_{p\in[0,1]} \sum_{a\in\{0,1\}} au(p)m_t(a,dp) \Big] \Big\}$$
$$\text{s.t.} \int_{p\in[0,1]} \sum_{a\in\{0,1\}} au(p)m_t(a,dp) \geq 0 \quad \forall t = 1,...,T$$

Since conditions (B.4) and (B.5) are affine in $(m_t)_{t=1}^T$, the set $\widehat{\mathcal{M}}$ is a convex subset of $\big\{\text{signed Borel measures on } \{0,1\} \times [0,1]\big\}^T$. The optimization above is thus a linear program over this convex set $\widehat{\mathcal{M}}$.

Let $\widehat{\mathcal{L}}((m_t)_{t=1}^T; \lambda) := \sum_{t=1}^T (1+\lambda_t)\big[\int_{p\in[0,1]} \sum_{a\in\{0,1\}} au(p)m_t(a,dp)\big]$, i.e., the Lagrangian function associated to the linear program. Since $u(\cdot)$ is bounded by Lemma 1 and $T < \infty$, the optimal value w^* is finite. Standard Lagrangian duality (e.g., Theorem 1 in Section 8.6 of Luenberger (1997)) then implies:[46]

$$w^* = \min_{\lambda\in\mathbb{R}^T_+} \sup_{(m_t)_{t=1}^T \in \widehat{\mathcal{M}}} \widehat{\mathcal{L}}((m_t)_{t=1}^T; \lambda)$$

where the minimum is achieved by some non-negative λ^*. Given any such λ^*, $(m^*_t)_{t=1}^T \in \widehat{\mathcal{M}}$ solves the linear program if and only if:

(i) $(m^*_t)_{t=1}^T \in \arg\max_{(m_t)_{t=1}^T \in \widehat{\mathcal{M}}} \widehat{\mathcal{L}}((m_t)_{t=1}^T; \lambda^*)$

(ii) $\lambda^*_t \int_{p\in[0,1]} \sum_{a\in\{0,1\}} au(p)m^*_t(a,dp) = 0$, $\forall t = 1,...,T$

(iii) $\int_{p\in[0,1]} \sum_{a\in\{0,1\}} au(p)m^*_t(a,dp) \geq 0$, $\forall t = 1,...,T$

[46] Theorem 1 in Section 8.6 of Luenberger (1997) does not directly state the sufficiency of conditions (i) – (iii) for optimality. However, this is obvious as those conditions together imply $(m^*_t)_{t=1}^T$ is feasible and $\sum_{t=1}^T \big[\int_{p\in[0,1]} \sum_{a\in\{0,1\}} au(p)m^*_t(a,dp)\big] = \sum_{t=1}^T (1+\lambda_t)\big[\int_{p\in[0,1]} \sum_{a\in\{0,1\}} au(p)m^*_t(a,dp)\big] = w^*$.

To check the corresponding Slater's condition, notice by properties (P2) and (P5) in Lemma 1, the consumer's surplus will be strictly positive at all t under the myopically optimal policy, and thus all IC constraints can hold strictly.

Finally, notice $\mathcal{M} = \widehat{\mathcal{M}}$ just means that $(m_t)_{t=1}^T \in \widehat{\mathcal{M}}$ if and only if it is induced by some $\phi \in \Phi$. Thus the above results are equivalent to the statements in the lemma.

Q.E.D.

B.4 Analyses and Proofs for Section 3.3

B.4.1 Preliminaries

Towards using the duality result, I start with examining the Lagrangian function optimization $\max_{\phi\in\Phi} \mathcal{L}(\phi; \lambda)$ given any generic multiplier $\lambda \in \mathbb{R}_+^T$. For this *unconstrained* Markov decision problem, let $V_\lambda(\cdot, t)$ be the value function at time t, which is inductively defined with the Bellman equation:

$$V_\lambda(\cdot, T+1) \equiv 0 \tag{B.6}$$

$$V_\lambda(p, t) = \max\left\{(1+\lambda_t)u(p) + \int_{p'} V_\lambda(p', t+1)G(dp'|p),\, V_\lambda(p, t+1)\right\} \forall t = 1, ..., T \tag{B.7}$$

where the two arguments in the maximization correspond to the values with and without time-t consumption of the product respectively. I define $H_\lambda(p, t)$ to be the difference between these two values, i.e.,

$$H_\lambda(p, t) := (1+\lambda_t)u(p) + \int_{p'} V_\lambda(p', t+1)G(dp'|p) - V_\lambda(p, t+1) \,\forall t = 1, ..., T \tag{B.8}$$

Intuitively, $V_\lambda(p, t)$ is the continuation value for the Lagrangian optimization at time t given $p_t = p$; $H_\lambda(p, t)$ measures the net benefit from inducing the time-t consumption given $p_t = p$. A preliminary result needed later is that $H_\lambda(\cdot, t)$ is continuous.

Lemma B.1. *$H_\lambda(\cdot, t)$ is continuous for any t.*

Proof. Since $u(\cdot)$ is continuous by definition, we just need to show $V_\lambda(p, t+1)$ and $\int_{p'} V_\lambda(p', t+1)G(dp'|p)$ are continuous in p. When $t = T$, these hold by the definition of $V_\lambda(\cdot, T+1)$. Given that they hold for time t, by the Bellman equation we know $V_\lambda(p, t)$ is also continuous in p. Furthermore, because $G(dp'|p)$ is weakly continuous in p and $V_\lambda(p, t)$ is bounded due to the boundedness of $u(\cdot)$, this also implies the continuity of $\int_{p'} V_\lambda(p', t)G(dp'|p)$ in p.[47] The proof is thus completed by (backward) induction in t.

Q.E.D.

[47] Pick any $(p_n)_n \to p^*$, the weak continuity implies $G(dp'|p_n) \xrightarrow{w} G(dp'|p^*)$, which further implies $\int f(p')G(dp'|p_n) \to \int f(p')G(dp'|p^*)$ for any bounded and continuous function f.

We have the standard dynamic programming result:

Lemma B.2. *Given any multiplier λ and initial belief state distribution μ_1, we have:*

(a) $\max_{\phi\in\Phi}\mathcal{L}(\phi;\lambda)=\int_p V_\lambda(p,1)\mu_1(dp)$*;*

(b) A policy $\phi^\lambda\in\arg\max_{\phi\in\Phi}\mathcal{L}(\phi;\lambda)$ if and only if $H_\lambda(p_t,t)>0\Rightarrow a_t=1$ and $H_\lambda(p_t,t)<0\Rightarrow a_t=0$ almost surely under it.

Proof. Pick any $\phi\in\Phi$. Notice that by the definition of V_λ, we have

$$\begin{aligned}V_\lambda(p_t,t)&\geq\phi_t(p_t)\Big[(1+\lambda_t)u(p_t)+\int_{p'}V_\lambda(p',t+1)G(dp'|p_t)\Big]+\big(1-\phi_t(p_t)\big)V_\lambda(p_t,t+1)\\&\overset{\text{a.s.}}{=}\mathbb{E}_\phi[(1+\lambda_t)a_tu(p_t)|p_t]+\mathbb{E}_\phi[V_\lambda(p_{t+1},t+1)|p_t]\end{aligned}\tag{B.9}$$

where the inequality holds as equality if and only if $H_\lambda(p_t,t)>0\Rightarrow\phi_t(p_t)=1$ and $H_\lambda(p_t,t)<0\Rightarrow\phi_t(p_t)=0$.

Using this repeatedly, we have:

$$\begin{aligned}\int_p V_\lambda(p,1)\mu_1(dp)=\mathbb{E}_\phi[V_\lambda(p_1,1)]&\geq\mathbb{E}_\phi[(1+\lambda_1)a_1u(p_1)]+\mathbb{E}_\phi[V_\lambda(p_2,2)]\\&\geq\mathbb{E}_\phi[(1+\lambda_1)a_1u(p_1)]+\mathbb{E}_\phi[(1+\lambda_2)a_2u(p_2)]+\mathbb{E}_\phi[V_\lambda(p_3,3)]\\...&\geq\mathbb{E}_\phi[(1+\lambda_1)a_1u(p_1)]+...+\mathbb{E}_\phi[(1+\lambda_T)a_Tu(p_T)]+\underbrace{\mathbb{E}_\phi[V_\lambda(p_{T+1},T+1)]}_{=0}\end{aligned}$$

Notice the last line is just $\mathcal{L}(\phi;\lambda)$. This shows $\int_p V_\lambda(p,1)\mu_1(dp)\geq\sup_{\phi\in\Phi}\mathcal{L}(\phi;\lambda)$. Moreover, notice that all these inequalities hold as equalities if and only if the inequality in (B.9) holds as equality for all t almost surely under ϕ. As is mentioned earlier, this is equivalent to $H_\lambda(p_t,t)>0\Rightarrow\phi_t(p_t)=1$ and $H_\lambda(p_t,t)<0\Rightarrow\phi_t(p_t)=0$ almost surely under ϕ. If there is indeed a measurable ϕ satisfying these properties, then we have $\int_p V_\lambda(p,1)\mu_1(dp)=\sup_{\phi\in\Phi}\mathcal{L}(\phi;\lambda)$, the supremum is achieved by such policy, and any other admissible policy is optimal if and only if it also satisfies these properties. Therefore, to prove the lemma, it now suffices to show that there is indeed a measurable policy satisfying $H_\lambda(p_t,t)>0\Rightarrow\phi_t(p_t)=1$ and $H_\lambda(p_t,t)<0\Rightarrow\phi_t(p_t)=0$ almost surely. We can construct such policy by defining $\phi_t(p)=\mathbb{1}_{\{H_\lambda(p,t)\geq0\}}$, where $\mathbb{1}$ is the indicator function. It is indeed measurable since $H_\lambda(\cdot,t)$ is continuous by Lemma B.1. *Q.E.D.*

I now provide some basic properties for the value function V_λ.

Lemma B.3. *Given any $\lambda\in\mathbb{R}^T_+$, we have: (a) $V_\lambda(p,t)$ is (weakly) increasing in p; (b) $\int_{p'}V_\lambda(p',t)G(dp'|p)\geq V_\lambda(p,t)$ for any pair of (p,t).*

Proof. The results can be shown by backward induction in t using the Bellman equation. Since $V_\lambda(\cdot,T+1)\equiv0$, both properties hold trivially for $t=T+1$.

Now, assume property (a) holds for $V_\lambda(\cdot, t+1)$. Then property (P3) in Lemma 1 implies that $\int_{p'} V_\lambda(p', t+1)G(dp'|p)$ increases in p. Together with the monotonicity of $u(\cdot)$ and the fact that $\lambda_t \geq 0$, we know the RHS of the Bellman equation (B.7) is increasing in p. This shows property (a) also holds for $V_\lambda(\cdot, t)$ and concludes the proof for part (a).

For property (b), assuming it holds for all periods later than t, we need to show $\int_{p'} V_\lambda(p', t)G(dp'|p) \geq V_\lambda(p, t)$. Substituting the Bellman equation in, we know this is equivalent to:

$$\int_{p'} \max\Big\{(1+\lambda_t)u(p') + \int_{p''} V_\lambda(p'', t+1)G(dp''|p'),\, V_\lambda(p', t+1)\Big\}G(dp'|p)$$
$$\geq \max\Big\{(1+\lambda_t)u(p) + \int_{p'} V_\lambda(p', t+1)G(dp'|p),\, V_\lambda(p, t+1)\Big\}$$

It then suffices to check:

$$\int_{p'} \Big((1+\lambda_t)u(p') + \int_{p''} V_\lambda(p'', t+1)G(dp''|p')\Big)G(dp'|p)$$
$$\geq (1+\lambda_t)u(p) + \int_{p'} V_\lambda(p', t+1)G(dp'|p)$$
$$\text{and} \quad \int_{p'} V_\lambda(p', t+1)G(dp'|p) \geq V_\lambda(p, t+1)$$

The second of these inequalities is directly implied by the induction hypothesis. To check the first one, notice by property (P2) in Lemma 1, we have $\int_{p'}(1+\lambda_t)u(p')G(dp'|p) = (1+\lambda_t)u(p)$. Moreover, by the induction hypothesis we have $\int_{p''} V_\lambda(p'', t+1)G(dp''|p') \geq V_\lambda(p', t+1)$ for any p', and thus $\int_{p'}\int_{p''} V_\lambda(p'', t+1)G(dp''|p')G(dp'|p) \geq \int_{p'} V_\lambda(p', t+1)G(dp'|p)$. These together imply the first inequality above. Thus property (b) holds for period t. This completes the proof by induction.

Q.E.D.

B.4.2 Proof for Lemma 3

Proof. Pick any $\lambda' \in \mathbb{R}^T_+$ such that $\lambda'_\tau < \lambda'_{\tau+1}$ for some τ. Define λ'' to be equal to λ' except for the terms of time τ and $\tau+1$, which are defined as: $\lambda''_\tau = \lambda'_{\tau+1}$ and $\lambda''_{\tau+1} = \lambda'_\tau$. The key to the proof is the following observation:

Claim. Let λ' and λ'' be defined as above. Then $\sup_{\phi\in\Phi} \mathcal{L}(\phi; \lambda') \geq \sup_{\phi\in\Phi} \mathcal{L}(\phi; \lambda'')$.

Proof for the Claim. Since λ' and λ'' agree for $t < \tau$, any policy will lead to the same flow payoffs for periods before τ for both $\mathcal{L}(\phi; \lambda')$ and $\mathcal{L}(\phi; \lambda'')$. It thus suffices to show $V_{\lambda'}(p, \tau) \geq V_{\lambda''}(p, \tau)$ given any p.

Since λ' and λ'' also agree for $t \geq \tau+2$, we have $V_{\lambda'}(\cdot, \tau+2) = V_{\lambda''}(\cdot, \tau+2)$. I can thus let $V_*(\cdot, \tau+2)$ denote both of them, i.e., $V_*(\cdot, \tau+2) := V_{\lambda'}(\cdot, \tau+2)$ $(= V_{\lambda''}(\cdot, \tau+2))$.

Also let $(y_1, y_2) := (\lambda'_\tau, \lambda'_{\tau+1})$. Then $y_1 < y_2$ and $(\lambda''_\tau, \lambda''_{\tau+1}) = (y_2, y_1)$. By the Bellman equation (B.7), we have:

$$V_{\lambda''}(p,\tau) = \max\Big\{(1+y_2)u(p) + \int_{p'} V_{\lambda''}(p',\tau+1)G(dp'|p),\ V_{\lambda''}(p,\tau+1)\Big\} \tag{B.10}$$

Consider the following two cases:

- **Case 1:** the maximum in equation (B.10) is achieved with $a_\tau = 0$ (i.e., no consumption).
  In this case, we have:

$$\begin{aligned} V_{\lambda''}(p,\tau) =& V_{\lambda''}(p,\tau+1) \\ =& \max\Big\{(1+y_1)u(p) + \int_{p'} V_*(p',\tau+2)G(dp'|p),\ V_*(p,\tau+2)\Big\} \\ \leq& \max\Big\{(1+y_1)u(p) + \int_{p'} V_{\lambda'}(p',\tau+1)G(dp'|p),\ V_{\lambda'}(p,\tau+1)\Big\} \\ =& V_{\lambda'}(p,\tau) \end{aligned}$$

  where the second equality holds by the Bellman equation for $V_{\lambda''}(p,\tau+1)$; the inequality holds because the Bellman equation for $V_{\lambda'}(p,\tau+1)$ implies that $V_*(p,\tau+2) \leq V_{\lambda'}(p,\tau+1)$ for any p; the last equality is just the Bellman equation for $V_{\lambda'}(p,\tau)$.
- **Case 2:** the maximum in equation (B.10) is achieved with $a_\tau = 1$.
  In this case, we have:

$$\begin{aligned} V_{\lambda''}(p,\tau) =& (1+y_2)u(p) + \int_{p'} V_{\lambda''}(p',\tau+1)G(dp'|p) \qquad \text{(B.11)} \\ =& (1+y_1)u(p) + \int_{p'} \big[\underbrace{(y_2-y_1)u(p') + V_{\lambda''}(p',\tau+1)}_{=:M(p')}\big]G(dp'|p) \qquad \text{(B.12)} \end{aligned}$$

  where the second equality holds because $u(p) = \int_{p'} u(p')G(dp'|p)$ by property (P2) in Lemma 1. Using the Bellman equation for $V_{\lambda''}(p',\tau+1)$, we know the term $M(p')$ satisfies:

$$\begin{aligned} M(p') =& (y_2-y_1)u(p') + \max\Big\{(1+y_1)u(p') + \int_{p''} V_*(p'',\tau+2)G(dp''|p'), V_*(p',\tau+2)\Big\} \\ =& \max\Big\{(1+y_2)u(p') + \int_{p''} V_*(p'',\tau+2)G(dp''|p'), V_*(p',\tau+2) + (y_2-y_1)u(p')\Big\} \end{aligned} \tag{B.13}$$

  Now, consider properties of $M(p')$ in two different scenarios about p':
  - Scenario 1: $u(p') > 0$.
    Notice Lemma B.3(b) implies $\int_{p''} V_*(p'',\tau+2)G(dp''|p') \geq V_*(p',\tau+2)$. When

$u(p') > 0$, we thus have $(1+y_2)u(p') + \int_{p''} V_*(p'', \tau+2)G(dp''|p') \geq V_*(p', \tau+2) + (y_2 - y_1)u(p')$. Therefore, (B.13) implies:

$$\begin{aligned} M(p') &= (1+y_2)u(p') + \int_{p''} V_*(p'', \tau+2)G(dp''|p') \\ &= \max\left\{(1+y_2)u(p') + \int_{p''} V_*(p'', \tau+2)G(dp''|p'), V_*(p', \tau+2)\right\} \\ &= V_{\lambda'}(p', \tau+1) \end{aligned}$$

where the second equality holds since $\int_{p''} V_*(p'', \tau+2)G(dp''|p') \geq V_*(p', \tau+2)$ and $u(p') > 0$ imply $(1+y_2)u(p') + \int_{p''} V_*(p'', \tau+2)G(dp''|p') \geq V_*(p', \tau+2)$.

– Scenario 2: $u(p') \leq 0$.
In this case, we have $(y_2 - y_1)u(p') \leq 0$. (B.13) thus implies:

$$\begin{aligned} M(p') &\leq \max\left\{(1+y_2)u(p') + \int_{p''} V_*(p'', \tau+2)G(dp''|p'), V_*(p', \tau+2)\right\} \\ &= V_{\lambda'}(p', \tau+1) \end{aligned}$$

where the equality holds by the Bellman equation.

In both scenarios, we always have $M(p') \leq V_{\lambda'}(p', \tau+1)$. Together with inequality (B.12), this implies that $V_{\lambda''}(p, \tau) \leq (1+y_1)u(p) + \int_{p'} V_{\lambda'}(p', \tau+1)G(dp'|p) \leq V_{\lambda'}(p, \tau)$, where the second inequality is due to the Bellman equation for $V_{\lambda'}(p, \tau)$.

In sum, $V_{\lambda''}(p, \tau) \leq V_{\lambda'}(p, \tau)$ in both cases, which completes the proof for the claim.

□

I now go back to the main proof for Lemma 3. Pick any $\lambda^0 \in \arg\min_{\lambda \in \mathbb{R}_+^T} \sup_{\phi \in \Phi} \mathcal{L}(\phi; \lambda)$ (whose existence is guaranteed by Lemma 2). If λ_t^0 is already non-increasing in t, we are done; if $\lambda_t^0 < \lambda_{t+1}^0$ for some t, then the claim above implies that by interchanging terms λ_t^0 and λ_{t+1}^0, we will get a new multiplier still in $\arg\min_{\lambda \in \mathbb{R}_+^T} \sup_{\phi \in \Phi} \mathcal{L}(\phi; \lambda)$. By repeatedly making such interchanges, we can then derive a multiplier $\lambda^* \in \arg\min_{\lambda \in \mathbb{R}_+^T} \sup_{\phi \in \Phi} \mathcal{L}(\phi; \lambda)$ such that $\lambda_t^* \geq \lambda_{t+1}^*$, $\forall t$.[48]

[48]More specifically, starting with $n = 0$, we can run the following algorithm:

1. Let $\tau_n = \inf\{t : \lambda_t^n < \lambda_{t+1}^n\}$. If $\tau_n = +\infty$, end the algorithm and out-put λ^n; otherwise, go to the next step.
2. Let $s = \max\left\{0, \sup\{t < \tau_n : \lambda_t^n \geq \lambda_{\tau_n+1}^n\}\right\}$ and define

$$\lambda_t^{n+1} = \begin{cases} \lambda_{\tau_n+1}^n & \text{if } t = s+1; \\ \lambda_{t-1}^n & \text{if } t = s+2, ..., \tau_n+1; \\ \lambda_t^n & \text{elsewhere} \end{cases}$$

Then, repeat the procedures with n replaced by $n+1$.

Intuitively, in step 2 of the algorithm we advance the first term in λ^n greater than its predecessor to an earlier position such that the first $\tau_n + 1$ terms will be in descending order. It is then easy to see that this algorithm will end in finite time and the vector it delivers will be non-increasing over t. Moreover, *Q.E.D.*

B.4.3 Properties of H_λ

Next, I provide some properties of H_λ, which are key to the proof for Lemma 4.

Lemma B.4. *H_λ satisfies the following properties:*

(a) $p > \bar{p} \Rightarrow H_\lambda(p,t) > 0$, $\forall t$;

(b) If λ_t is non-increasing in t, then $H_\lambda(p,t)$ is (weakly) increasing in p for any t;

(c) If λ_t is non-increasing in t, then for any x, y s.t. $x < y$ and $H_\lambda(y,t) \le 0$, we have $H_\lambda(x,t) < H_\lambda(y,t)$ (thus $H_\lambda(\cdot,t)$ has at most one root);

(d) If $\lambda_t = \lambda_{t+1}$, then $H_\lambda(p,t) \le 0 \Rightarrow H_\lambda(p,t+1) < 0$.

Proof. **Part (a):** Because $\int_{p'} V_\lambda(p',t+1)G(dp'|p) - V_\lambda(p,t+1) \ge 0$ according to Lemma B.3, part (a) is directly implied by the definition of H_λ.

Part (b): I prove (b) by backward induction in t. By definition, $H_\lambda(p,T) = (1+\lambda_T)u(p)$ is strictly increasing in p and thus the monotonicity property holds for $H_\lambda(p,T)$. Now, assuming it holds for $H_\lambda(p,t+1)$, I show it also holds for $H_\lambda(p,t)$. Notice the following equations hold:

$$\begin{aligned}
H_\lambda(p,t) =& (1+\lambda_t)u(p) + \int_{p'} V_\lambda(p',t+1)G(dp'|p) - V_\lambda(p,t+1) \\
=& (1+\lambda_t)u(p) + \int_{p'} \Big(\max\{H_\lambda(p',t+1),0\} + V_\lambda(p',t+2) \Big) G(dp'|p) \\
& - \Big(\max\{H_\lambda(p,t+1),0\} + V_\lambda(p,t+2) \Big) \\
=& (1+\lambda_t)u(p) + \int_{p'} V_\lambda(p',t+2)G(dp'|p) - V_\lambda(p,t+2) \\
& + \int_{p'} \max\{H_\lambda(p',t+1),0\}G(dp'|p) - \max\{H_\lambda(p,t+1),0\} \\
=& (1+\lambda_t)u(p) - (1+\lambda_{t+1})u(p) + H_\lambda(p,t+1) \\
& + \int_{p'} \max\{H_\lambda(p',t+1),0\}G(dp'|p) - \max\{H_\lambda(p,t+1),0\} \\
=& (\lambda_t - \lambda_{t+1})u(p) + \min\{H_\lambda(p,t+1),0\} + \int_{p'} \max\{H_\lambda(p',t+1),0\}G(dp'|p)
\end{aligned} \tag{B.14}$$

to derive λ^{n+1} from λ^n in step 2, one can just interchange the (τ_n+1)'th term with the τ_n'th term, then interchange the (new) τ_n'th term with the (τ_n-1)'th term,..., and finally interchange the (new) $(s+2)$'th term with the $(s+1)$'th term. In each of these steps, we interchange two adjacent terms with the latter greater than the former. By the claim proved above, this keeps each of the (intermediate) vector within $\arg\min_{\lambda\in\mathbb{R}^T_+} \sup_{\phi\in\Phi} \mathcal{L}(\phi;\lambda)$. Thus the multiplier we derive in the end remains in $\arg\min_{\lambda\in\mathbb{R}^T_+} \sup_{\phi\in\Phi} \mathcal{L}(\phi;\lambda)$. This completes the proof.

where the second equality holds because $V_\lambda(p, t+1) = \max\{H_\lambda(p, t+1), 0\} + V_\lambda(p, t+2)$ according to the Bellman equation (B.7); the first and the fourth equalities are directly implied by the definition of H_λ; the other two are trivial identities. Recall that: $\lambda_t \geq \lambda_{t+1}$ by assumption; $H_\lambda(p, t+1)$ weakly increases in p by the induction hypothesis; and $G(\cdot|p)$ increases in first-order stochastic dominance in p by property (P3) of Lemma 1. These imply that all of the three terms in the last expression are (weakly) increasing in p. Thus $H_\lambda(p, t)$ is (weakly) increasing in p. This completes the proof for (b).

Part (c): I still prove by induction. The result holds obviously for $H_\lambda(p, T) = (1 + \lambda_T)u(p)$. Now, assuming it holds for period $t+1$, I show it also holds for period t. In particular, with any $x < y$ in $[0, 1]$, we want to show $H_\lambda(y, t) \leq 0 \Rightarrow H_\lambda(x, t) < H_\lambda(y, t)$. Given result (b) and equation (B.14) derived above, $H_\lambda(x, t) < H_\lambda(y, t)$ obviously holds when $\lambda_t > \lambda_{t+1}$, since $u(p) = \theta_H p + \theta_L(1-p)$ is strictly increasing in p. It thus suffices to assume $\lambda_t = \lambda_{t+1}$. In this case, I have the following observation:

Claim. If $\lambda_t = \lambda_{t+1}$ and $H_\lambda(y, t) \leq 0$, then $H_\lambda(y, t+1) \leq 0$.
Proof for the claim. Given $\lambda_t = \lambda_{t+1}$ and $H_\lambda(y, t) \leq 0$, suppose $H_\lambda(y, t+1) > 0$. Then equation (B.14) implies that $\int_{p'} \max\{H_\lambda(p', t+1), 0\} G(dp'|y) = H_\lambda(y, t) \leq 0$. However, the belief process must have $G\big([y, 1]|y\big) > 0$ (as is implied by property (P2) in Lemma 1). Moreover, by the monotonicity of $H_\lambda(\cdot, t+1)$ proved in part (b), we know $H_\lambda(p', t+1) \geq H_\lambda(y, t+1)$ for any $p' \geq y$. Together with the hypothesis $H_\lambda(y, t+1) > 0$, these then imply $\int_{p'} \max\{H_\lambda(p', t+1), 0\} G(dp'|y) > 0$, which contradicts with the previous conclusion. Thus we must have $H_\lambda(y, t+1) \leq 0$. □

Now, go back to the main proof for part (c). Notice when $\lambda_t = \lambda_{t+1}$ and $H_\lambda(y, t) \leq 0$, the following holds:

$$\begin{aligned} H_\lambda(y, t) &= \min\{H_\lambda(y, t+1), 0\} + \int_{p'} \max\{H_\lambda(p', t+1), 0\} G(dp'|y) \\ &= H_\lambda(y, t+1) + \int_{p'} \max\{H_\lambda(p', t+1), 0\} G(dp'|y) \\ &> H_\lambda(x, t+1) + \int_{p'} \max\{H_\lambda(p', t+1), 0\} G(dp'|x) = H_\lambda(x, t) \end{aligned}$$

The first equality is just by equation (B.14) with $\lambda_t = \lambda_{t+1}$. The second equality holds because the claim proved above implies $H_\lambda(y, t+1) \leq 0$. The strict inequality holds because: (i) $H_\lambda(y, t+1) \leq 0$ further implies $H_\lambda(x, t+1) < H_\lambda(y, t+1)$ by the induction hypothesis; (ii) $G(\cdot|y)$ first order stochastic dominates $G(\cdot|x)$; and (iii) $H_\lambda(\cdot, t+1)$ is increasing by part (b). The last equality holds also because of equation (B.14) and the fact that $H_\lambda(x, t+1) < 0$. This completes the proof by induction.

Part (d): Suppose $\lambda_t = \lambda_{t+1}$, $H_\lambda(p,t) \leq 0$, but $H_\lambda(p,t+1) \geq 0$. Equation B.14 would imply

$$H_\lambda(p,t) = \int_{p'} \max\{H_\lambda(p',t+1),0\}G(dp'|p)$$

I now argue that the RHS above must be strictly positive. Notice $H_\lambda(p,t) \leq 0$ obviously imply $p < 1$. Property (P4) in Lemma 1 then implies that $G((p,1]|p) > 0$. Moreover, since $H_\lambda(p,t+1) \geq 0$, parts (b) and (c) proved earlier imply $H_\lambda(p',t+1) > 0$ for any $p' > p$. These together imply $\int_{p'} \max\{H_\lambda(p',t+1),0\}G(dp'|p) > 0$. This contradicts with $H_\lambda(p,t) \leq 0$ given the equation above. Thus the result in part (d) holds.

Q.E.D.

B.4.4 Proof for Lemma 4

The proof for Lemma 4 easily follows from Lemma B.4.

Proof. For each t, I construct the threshold η_t as follows:

- Case 1: $\{p : H_\lambda(p,t) = 0\} = \emptyset$.
  In this case, define $\eta_t = \inf\{p : H_\lambda(p,t) > 0\}$.
- Case 2: $\{p : H_\lambda(p,t) = 0\} \neq \emptyset$.
  In this case, Lemma B.4(c) implies that $\{p : H_\lambda(p,t) = 0\}$ contains a single element. Define η_t to be this element.

By the monotonicity property of $H_\lambda(\cdot,t)$ in Lemma B.4(b), in both cases we have $p_t > \eta_t \Rightarrow H_\lambda(p_t,t) > 0$ and $p_t < \eta_t \Rightarrow H_\lambda(p_t,t) < 0$. Together with Lemma B.2(b), this implies that under any solution to $\max_{\phi\in\Phi} \mathcal{L}(\phi;\lambda)$, we have $p_t > \eta_t \Rightarrow a_t = 1$ a.s. and $p_t < \eta_t \Rightarrow a_t = 0$ a.s. Hence any solution is almost surely equivalent to a threshold policy with thresholds being $(\eta_t)_{t=1}^T$.

Moreover, notice Lemma B.4(a) implies $u(p_t) > 0 \Rightarrow H_\lambda(p_t,t) > 0$. Together with Lemma B.2(b), this then implies $u(p_t) > 0 \Rightarrow a_t = 1$ a.s. under any optimal solution to $\max_{\phi\in\Phi} \mathcal{L}(\phi;\lambda)$.

Q.E.D.

B.5 Proofs for Section 3.4

B.5.1 Definition of ϕ^d

As is mentioned in the main text, I define ϕ^d as the "most conservative" optimal policy for the dictator's problem. Formally, for any $t = 1, ..., T$:

$$\phi_t^d(p) := \begin{cases} 1 & \text{if } H_{\mathbf{0}}(p,t) > 0; \\ 0 & \text{otherwise} \end{cases} \tag{B.15}$$

where $H_{\mathbf{0}}$ is as defined in (B.8) in Appendix B.4 with $\lambda = \mathbf{0}$. By Lemma B.2, it is easy to see that we not only have ϕ^d being optimal for the dictator's problem, but also have $\phi^d_{\geq t}$ to be optimal for the dictator's continuation problem starting from time t regardless of the belief distribution at t. Moreover, ϕ^d (or $\phi^d_{\geq t}$) is conservative in the sense that it breaks any tie in favor of non-recommendation (i.e., $H_{\mathbf{0}}(p, t) = 0 \Rightarrow \phi^d_t(p) = 0$). This makes it most favorable to the current consumer among all dictator's optimal (continuation) policies.

B.5.2 Details in the Construction of ϕ^o and a Uniqueness Property

I first construct a threshold time-t policy ϕ^o_t satisfying the requirements in step 2 of the algorithm in Definition 2. Given μ^o_t, define the policy's threshold as

$$\eta^o_t = \inf\{x \in [0,1] : \int_{p>x} u(p)\mu^o_t(dp) > 0\}$$

and define the recommendation probability at the threshold as

$$\phi^o_t(\eta^o_t) = -\frac{\int_{p>\eta^o_t} u(p)\mu^o_t(dp)}{u(\eta^o_t)\mu^o_t(\{\eta^o_t\})}$$

When the denominator above is zero, I define $\phi^o_t(\eta^o_t) = 0$ for simplicity. Now, I check the ϕ^o_t such defined indeed satisfies the desired properties.

Claim. ϕ^o_t above is well-defined and satisfies the properties in step 2 of the algorithm.

Proof. First notice properties (P2) and (P5) in Lemma 1 together imply $\mu^o_t((\bar{p}, 1]) > 0$ and thus $\int_{p>\bar{p}} u(p)\mu^o_t(dp) > 0$. This implies $\eta^o_t \leq \bar{p}$ and thus the first property is satisfied.

Now, I argue that $\int_{p>\eta^o_t} u(p)\mu^o_t(dp) \geq 0$. To see this, notice by the definition of η^o_t, there exists a sequence $\{x_n\} \downarrow \eta^o_t$ such that $\int_{p>x_n} u(p)\mu^o_t(dp) > 0$ for all n. Since $\mathbb{1}_{\{p>x_n\}} \to \mathbb{1}_{\{p>\eta^o_t\}}$ and $u(\cdot)$ is bounded, we must have $\int_{p>\eta^o_t} u(p)\mu^o_t(dp) \geq 0$ by the dominated convergence theorem.

Next, I argue that $\int_{p\geq\eta^o_t} u(p)\mu^o_t(dp) \leq 0$. As the algorithm has not been ended in step 1, we must have $\int_{p\in[0,1]} u(p)\mu^o_t(dp) < 0$.[49] Thus the argument is true when $\eta^o_t = 0$. When $\eta^o_t > 0$, notice by the definition of η^o_t, there exists a sequence $\{x_n\} \uparrow \eta^o_t$ such that $\int_{p>x_n} u(p)\mu^o_t(dp) \leq 0$ for all n. Since $\mathbb{1}_{\{p>x_n\}} \to \mathbb{1}_{\{p\geq\eta^o_t\}}$, by the dominated convergence theorem we then must have $\int_{p\geq\eta^o_t} u(p)\mu^o_t(dp) \leq 0$.

Now, consider two cases:

- Case 1: $\mu^o_t(\{\eta^o_t\}) = 0$. In this case, the above arguments imply $\int_{p>\eta^o_t} u(p)\mu^o_t(dp) = 0$. Therefore the second property is satisfied.

[49] Notice $\phi^d_t(p) = 1$ for any $p > \bar{p}$ by the definition of ϕ^d and Lemma B.4(a).

- Case 2: $\mu_t^o(\{\eta_t^o\}) > 0$. In this case, the above arguments imply $\int_{p>\eta_t^o} u(p)\mu_t^o(dp) \geq 0$ and $\int_{p>\eta_t^o} u(p)\mu_t^o(dp) + u(\eta_t^o)\mu_t^o(\{\eta_t^o\}) \leq 0$. These imply $0 \geq \frac{\int_{p>\eta_t^o} u(p)\mu_t^o(dp)}{u(\eta_t^o)\mu_t^o(\{\eta_t^o\})} \geq -1$ whenever $u(\eta_t^o)\mu_t^o(\{\eta_t^o\}) \neq 0$. Thus $\phi_t^o(\eta_t^o)$ defined above is a valid probability. By the definition of $\phi_t^o(\eta_t^o)$, we have $\int_p \phi_t^o(p)u(p)\mu_t^o(dp) = \int_{p>\eta_t^o} u(p)\mu_t^o(dp) + \phi_t^o(\eta_t^o)u(\eta_t^o)\mu_t^o(\{\eta_t^o\}) = 0$. Thus ϕ_t^o satisfies the second desired property.

Q.E.D.

Due to the possible existence of off-path belief states, there can also be other forms of ϕ_t satisfying the desired properties in step 2 of the algorithm. However, the following lemma implies that any such policy must μ_t^o-a.e. agree with ϕ_t^o.

Lemma B.5. *Given any probability measure μ over $[0,1]$ such that $\mu((\bar{p},1]) > 0$,[50] any threshold time-t policies satisfying $p > \bar{p} \Rightarrow \phi_t(p) = 1$ (μ-a.e.) and $\int_p \phi_t(p)u(p)\mu(dp) = 0$ must agree μ-a.e.*

Proof. For $i = 1,2$, let ϕ_t^i be a threshold time-t policy with threshold η_t^i, which satisfies $\int_p \phi_t^i(p)u(p)\mu(dp) = 0$ and $p > \bar{p} \Rightarrow \phi_t^i(p) = 1$ (μ-a.e.). Without loss of generality, assume $\eta^1 \leq \eta^2$. Notice under the assumption $\mu((\bar{p},1]) > 0$, $\int_p \phi_t^i(p)u(p)\mu(dp) = 0$ implies that we must have $\phi_t^i(p) > 0$ for some p with $u(p) < 0$. Thus the threshold structure implies $\phi_t^i(p) = 1$ for all p such that $u(p) \geq 0$, which holds for both $i = 1,2$. Then, notice we have:

$$0 = \int_p \phi_t^1(p)u(p)\mu(dp) - \int_p \phi_t^2(p)u(p)\mu(dp) = \int_{p:u(p)<0} \big(\phi_t^1(p) - \phi_t^2(p)\big)u(p)\mu(dp)$$

If $\eta^1 < \eta^2$, then $\phi_t^1(p) \geq \phi_t^2(p)$ for all p because of the threshold structure. Supposing the policies do not agree μ-a.e. , which can only happen when $u(p) < 0$, then $\phi_t^1(p) > \phi_t^2(p)$ for a positive μ-measure set of p with $u(p) < 0$. This implies that the last expression above is strictly negative, which is a contradiction.

If $\eta^1 = \eta^2 =: \eta$, then the two policies can only differ at $p = \eta$ with $u(\eta) < 0$. Supposing they do not agree μ-a.e, we must have $\mu(\{\eta\}) > 0$ and $\phi_t^1(\eta) \neq \phi_t^2(\eta)$. These imply that the last expression above equals to $[\phi_t^1(\eta) - \phi_t^2(\eta)]u(\eta)\mu(\{\eta\}) \neq 0$, which is a contradiction. Thus the policies must agree μ-a.e. *Q.E.D.*

B.5.3 Proof for Proposition 3 and Related Results

We need to first prove some properties of ϕ^d.

Lemma B.6. *Set $\{p : \phi_t^d(p) = 1\}$ shrinks in set inclusion order as t increases.*

[50] Notice properties (P2) and (P5) in Lemma 1 together imply that this holds for time-t belief distribution μ_t under any policy.

Proof. Since by construction $\phi_t^d(p) = 1$ if and only if $H_{\mathbf{0}}(p,t) > 0$, the result is directly implied by Lemma B.4(d) in Section B.4. *Q.E.D.*

This observation leads to the following result:

Lemma B.7. *Given any time-t belief distribution μ_t, we have $\int_p \phi_t^d(p)u(p)\mu_t(dp) \le \int_p \phi_{t+1}^d(p)u(p)\mu_{t+1}(dp)$, where μ_{t+1} is the period $t+1$ belief distribution under ϕ_t^d given μ_t. (That is, the consumer's expected payoff is weakly higher in period $t+1$ than in period t under ϕ^d.)*

Proof. The result is proved by the following arguments:

$$\begin{aligned}
&\int_{p'} \phi_t^d(p')u(p')\mu_t(dp') \\
=&\int_p \int_{p'} u(p')\phi_{t+1}^d(p')[\phi_t^d(p)G(dp'|p) + (1-\phi_t^d(p))D(dp'|p)]\mu_t(dp) \\
=&\int_p \int_{p'} u(p')\phi_{t+1}^d(p')G(dp'|p)\phi_t^d(p)\mu_t(dp) \ge \int_p u(p)\phi_t^d(p)\mu_t(dp)
\end{aligned}$$

The first equality holds by the transition rule for p. The second equality holds because $\phi_t^d(p) = 0 \Rightarrow \phi_{t+1}^d(p) = 0$ by Lemma B.6, and thus $\int_{p'} u(p')\phi_{t+1}^d(p')(1-\phi_t^d(p))D(dp'|p) = u(p)\phi_{t+1}^d(p)(1-\phi_t^d(p)) = 0$. The last inequality holds because $\int_{p'} u(p')\phi_{t+1}^d(p')G(dp'|p) \ge \int_{p'} u(p')G(dp'|p) = u(p)$, where the "$\ge$" is due to $u(p') > 0 \Rightarrow \phi_{t+1}^d(p') = 1$ and the "$=$" is implied by property (P2) in Lemma 1.

Q.E.D.

To ease notation, let IC_t denote the IC constraint for time-t consumer. Repeated use of Lemma B.7 implies that given any μ_t, if $\phi_{\ge t}^d$ satisfies IC_t, then it satisfies all later IC's. We are now ready to prove the proposition.

Proof for Proposition 3. First consider the "only if" part. Let ϕ^{opt} be any optimal policy for the designer and let μ_t^{opt} be the distribution of p_t under it for any t. Due to Proposition 2, I can assume ϕ^{opt} is a threshold policy without loss of generality. Let λ^* be a Lagrangian multiplier solving the dual problem that is non-increasing over t (which exists by Lemma 3). I first check condition (i):

***Claim** (a).* $\phi_{<\hat{t}}^{opt}$ agrees with $\phi_{<\hat{t}}^o$ a.s.

Proof for Claim (a). I check by forward induction in t for $t < \hat{t}$. For $t = 1$, $t < \hat{t}$ implies that ϕ_1^d violates IC_1 given the initial state distribution μ_1. In this case, we must have $\lambda_1^* > 0$ and thus IC_1 is binding under ϕ^{opt}. (Suppose not. Then $\lambda_t^* = 0$ for all $t \ge 1$ since it is non-increasing in t, and thus the Lagrangian problem $\max \mathcal{L}(\phi; \lambda^*)$ would coincide with the dictator's problem. By Lemma 2, this implies that ϕ^{opt} must also solve the dictator's problem. However, by construction ϕ^d provides the highest time-1 expected consumer surplus given μ_1 among all optimal policies for the dictator's problem. Thus when ϕ_1^d violates IC_1, so does ϕ_1^{opt}, which is a contradiction.) Also notice the optimality of ϕ^{opt} implies $p > \bar{p} \Rightarrow \phi_1^{opt}(p) = 1$ (μ_1-a.e.). Thus ϕ_1^{opt} satisfies the properties of ϕ_1^o in step 2 of the algorithm. By Lemma B.5, we then must have ϕ_1^{opt} and ϕ_1^o agree μ_1-a.e.

Now, assume ϕ^{opt} and ϕ^o agree a.s. for all periods before t and $t < \widehat{t}$. Then, they induce the same distribution for p_t, which is just μ_t^o constructed in the algorithm. I want to show ϕ^{opt} must also satisfy $p > \bar{p} \Rightarrow \phi_t^{opt}(p) = 1$ μ_t^o-a.s. and $\int_p \phi_t^{opt}(p)u(p)\mu_t^o(dp) = 0$. The former is directly implied by the optimality of ϕ^{opt}. For the latter, notice $t < \widehat{t}$ implies that ϕ^d violates IC_t given $p_t \sim \mu_t^o$. We then must have $\lambda_t^* > 0$ and thus IC_t is binding under ϕ^{opt}. (Suppose not. Then $\lambda_{t'}^* = 0$ for all $t' \geq t$, and thus the continuation Lagrangian problem starting with time t coincides with the corresponding dictator's continuation problem. The optimality of ϕ^{opt} then implies that $\phi_{\geq t}^{opt}$ must be optimal for this dictator's continuation problem given $p_t \sim \mu_t^o$. However, among all such policies, $\phi_{\geq t}^d$ delivers the highest time-t expected consumer surplus. Thus when ϕ_t^d violates IC_t, so does ϕ_t^{opt}, which is a contradiction.) Thus $\int_p \phi_t^{opt}(p)u(p)\mu_t^o(dp) = 0$. Again by Lemma B.5, we must have ϕ_t^{opt} and ϕ_t^o agree μ_t^o-a.e. The proof is then completed by induction. □

Now, since $\phi_{<\widehat{t}}^{opt}$ and $\phi_{<\widehat{t}}^o$ agree almost surely, they lead to the same distribution for $p_{\widehat{t}}$, which is just $\mu_{\widehat{t}}^o$. Hence $\phi_{\geq\widehat{t}}^{opt}$ must satisfy all IC constraints after time $\widehat{t}$ given $p_{\widehat{t}} \sim \mu_{\widehat{t}}^o$. The following claim checks the rest of condition (ii) in the proposition.

Claim *(b).* $\phi_{\geq\widehat{t}}^{opt}$ is optimal for the dictator's continuation problem since time $\widehat{t}$ given $p_{\widehat{t}} \sim \mu_{\widehat{t}}^o$.

Proof for Claim (b). By the definition of $\widehat{t}$, $\phi_{\widehat{t}}^d$ satisfies $IC_{\widehat{t}}$ given $p_{\widehat{t}} \sim \mu_{\widehat{t}}^o$ and thus $\phi_{\geq\widehat{t}}^d$ satisfies all later IC constraints by Lemma B.7 given $p_{\widehat{t}} \sim \mu_{\widehat{t}}^o$. This implies that if we deviate from $\phi_{\geq\widehat{t}}^{opt}$ to $\phi_{\geq\widehat{t}}^d$ since period $\widehat{t}$, no IC constraint will be violated. Also notice such deviation can only improve the total surplus since $\phi_{\geq\widehat{t}}^d$ is optimal for the dictator's continuation problem, which is more relaxed than the original continuation problem. For such a deviation to be unprofitable, we then need $\phi_{\geq\widehat{t}}^{opt}$ to achieve the same value as $\phi_{\geq\widehat{t}}^d$ for that continuation problem and thus $\phi_{\geq\widehat{t}}^{opt}$ is also optimal for it. □

Now, I turn to the "if" part. Given the existence of optimal policy (guaranteed by Proposition 1), it suffices to see that all policies satisfying the two conditions (i) and (ii) are feasible and yield the same total payoff for the designer. They generate the same total payoff since: for periods before $\widehat{t}$, all such policies agree a.s. and lead to $p_{\widehat{t}} \sim \mu_{\widehat{t}}^o$; for periods $t \geq \widehat{t}$, all such policies achieve the same total payoff as that under $\phi_{\geq\widehat{t}}^d$ given $p_{\widehat{t}} \sim \mu_{\widehat{t}}^o$. To show they are feasible, it suffices to check $\phi_{<\widehat{t}}^o$ is feasible. This is true since $\phi_{<\widehat{t}}^o$ satisfies all IC constraints for $t < \widehat{t}$ as equalities by construction. This completes the proof for the "if" part.

Finally, notice by construction ϕ^o does satisfy conditions (i) and (ii). In particular, Lemma B.7 implies that $\phi^d_{\geq \widehat{t}}$ satisfies all IC constraints after time $\widehat{t}$ given $p_{\widehat{t}} \sim \mu^o_{\widehat{t}}$, which implies the feasibility of ϕ^o following time $\widehat{t}$. Thus ϕ^o is optimal.

Q.E.D.

B.6 Proof for Proposition 4

Proof. Let ϕ^* denote an optimal policy, let $(\eta^*_t)_{t=1}^T$ denote its thresholds, and let μ^*_t denote the distribution of p_t under it. Notice that under the full support condition in Assumption 2, Proposition 2 implies that we must have $\eta^*_t \leq \bar{p}$ for all t. (Recall that $\bar{p}$ is the myopic threshold, i.e., $u(\bar{p}) = 0$.) Moreover, under the atomless condition in Assumption 2, randomization at the thresholds does not matter.

For part (a), I first show the following observation:

Claim. For any $\eta \in (0, \bar{p}]$, we have $\int_{p\geq\eta} \left[\int_{p'\geq\eta} u(p')G(dp'|p)\right] \mu^*_t(dp) > \int_{p\geq\eta} u(p)\mu^*_t(dp)$.

Proof for the claim. Given any $\eta \in (0, \bar{p}]$, recall that property (P4) in Lemma 1 implies $G([0, \eta)|\eta) > 0$. By the weak continuity of $G(\cdot|p)$ on p (i.e., property (P1) in Lemma 1), this further implies that there exists $\delta > 0$ s.t. $G([0, \eta)|p) > 0\, \forall p \in [\eta, \eta + \delta]$.[51] Together with the full support assumption on μ^*_t, we then have $\int_{p\geq\eta} G\big([0, \eta)|p\big)\mu^*_t(dp) > 0$. Since $\eta \leq \bar{p}$, $u(p) < 0$ for all $p < \eta$. Thus $\int_{p\geq\eta} \int_{p'<\eta} u(p')G(dp'|p)\mu^*_t(dp) < 0$. This then implies $\int_{p\geq\eta} \int_{p'\geq\eta} u(p')G(dp'|p)\mu^*_t(dp) > \int_{p\geq\eta} \int_{p'} u(p')G(dp'|p)\mu^*_t(dp) = \int_{p\geq\eta} u(p)\mu^*_t(dp)$, where the equality holds by property (P2) in Lemma 1. □

Now, I argue that the following holds given any $t \leq \widehat{t} - 2$:

$$\begin{aligned}
&\int_{p\geq\eta^*_t} u(p)\mu^*_{t+1}(dp) \\
&= \int_p \int_{p'\geq\eta^*_t} u(p')\left[\mathbb{1}_{\{p\geq\eta^*_t\}} G(dp'|p) + \mathbb{1}_{\{p<\eta^*_t\}} D(dp'|p)\right]\mu^*_t(dp) \\
&= \int_{p\geq\eta^*_t} \left[\int_{p'\geq\eta^*_t} u(p')G(dp'|p)\right]\mu^*_t(dp) + \int_{p\geq\eta^*_t} \mathbb{1}_{\{p<\eta^*_t\}} u(p)\mu^*_t(dp) \\
&= \int_{p\geq\eta^*_t} \left[\int_{p'\geq\eta^*_t} u(p')G(dp'|p)\right]\mu^*_t(dp) \\
&> \int_{p\geq\eta^*_t} u(p)\mu^*_t(dp) = 0
\end{aligned}$$

The first equality holds by the transition rule of p_t; the second equality is trivial identity; the third equality holds because the second term in line 3 is obviously zero; the last expression equals to zero because the IC constraint is binding for any $t < \widehat{t}$ by Proposition 3. To see the inequality holds, notice that $t \leq \widehat{t} - 2$ necessarily implies $\eta^*_t > 0$, since

[51] See Theorem 3.2.11 in Durrett (2019) (equivalence between conditions (i) and (ii)).

otherwise ϕ_t^d would be feasible at time t and the algorithm in Definition 2 would have stopped in step 1 at time t. The desired inequality is then directly implied by the claim proved above.

For part (b), notice under the atomless assumption in Assumption 2, Lemma B.4(c) (see Appendix B.4) implies that $H_{\mathbf{0}}(p_t, t)$ is non-zero almost surely under any policy. Lemma B.2(b) then implies that any optimal policy for the dictator must almost surely agree with ϕ^d. By Proposition 3, this further implies that $\phi^*_{\geq \hat{t}}$ must almost surely agree with $\phi^d_{\geq \hat{t}}$ (given $p_t \sim \mu^*_{\hat{t}}$). Under the full support assumption in Assumption 2, this then requires that $\phi^*_{\geq \hat{t}}$ and $\phi^d_{\geq \hat{t}}$ share the same thresholds. It thus suffices to prove the desired property for ϕ^d.

Let $(\eta_t^d)_{t=1}^T$ denote the sequence of thresholds of ϕ^d. Recall that $\phi_t^d(p) = \mathbb{1}_{\{H_{\mathbf{0}}(p,t)>0\}}$ (as is defined in Appendix B.5.1). By the continuity of $H_{\mathbf{0}}(\cdot, t)$ (Lemma B.1), we then have $H_{\mathbf{0}}(\eta_t^d, t) = 0$ for any t.[52] By Lemma B.4(d), this further implies $H_{\mathbf{0}}(\eta_t^d, t+1) < 0$. Since $H_{\mathbf{0}}(\cdot, t+1)$ is increasing (Lemma B.4(b)), we thus must have $\eta_{t+1}^d > \eta_t^d$ for any t. This completes the proof for part (b).

Finally, $\eta_t^* \leq \bar{p}$ for all t is directly implied by Proposition 2 under the full support assumption in Assumption 2.

Q.E.D.

B.7 Proof for Proposition 5

The designer's problem is formally written as:

$$\max_{\phi \in \Phi} \Big\{ \sum_{t=1}^{T} \mathbb{E}_\phi[a_t u(p_t)] \Big\}$$

$$\text{s.t. } \mathbb{E}_\phi[a_t u(p_t)] \geq 0 \quad \forall t = 1, ..., T$$

$$p_{t+1} | p_t, a_t \sim a_t \big[\alpha G^I(\cdot|p_t) + (1-\alpha) D(\cdot|p_t) \big] + (1-a_t) D(\cdot|p_t)$$

$$p_1 \sim \mu_1$$

Pick α_a and α_b with $\alpha_a < \alpha_b$. Corresponding to these two information generation rates respectively, let G^a and G^b be the transition kernels of p_t following one's consumption, as is defined in equation (9); let $V_{\mathbf{0}}^a$ and $V_{\mathbf{0}}^b$ be the value functions for the dictator's problem (i.e., with $\lambda = \mathbf{0}$), as is defined in Section B.4; let $H_{\mathbf{0}}^a$ and $H_{\mathbf{0}}^b$ be the associated H-functions (with $\lambda = \mathbf{0}$) as in equation (B.8); let $\hat{t}^a$ and $\hat{t}^b$ denote the critical time points defined in Definition 2; let ϕ^a and ϕ^b be the optimal threshold policies and denote their sequences of thresholds as $(\eta_t^a)_{t=1}^T$ and $(\eta_t^b)_{t=1}^T$. I prove the proposition by showing a sequence of observations below.

[52] Notice it is easy to see that $H_{\mathbf{0}}(0, t) < 0$ and $H_{\mathbf{0}}(1, t) > 0$.

Observation 1: $\eta_T^a = \eta_T^b$

This is obvious since at the last period the optimal threshold just equals to the myopically optimal threshold.

Observation 2: $\eta_t^a > \eta_t^b$ **for all** $t \in [\hat{t}^b, T)$**.**

To show this observation, first notice that since higher α is beneficial, we have the following non-surprising result for $V_{\mathbf{0}}^a$ and $V_{\mathbf{0}}^b$:

Claim *(a).* $V_{\mathbf{0}}^b(p,t) \geq V_{\mathbf{0}}^a(p,t)$ for any pair of (p,t).

Proof for Claim (a). I show by backward induction on t. For $t = T+1$, the result holds trivially. Assuming it holds for all $t' > t$, I now consider time t.

By the Bellman equation, we have:

$$\begin{aligned} V_{\mathbf{0}}^b(p,t) - V_{\mathbf{0}}^a(p,t) = &\max\left\{u(p) + \int_{p'} V_{\mathbf{0}}^b(p',t+1)G^b(dp'|p),\, V_{\mathbf{0}}^b(p,t+1)\right\} \\ &- \max\left\{u(p) + \int_{p'} V_{\mathbf{0}}^a(p',t+1)G^a(dp'|p),\, V_{\mathbf{0}}^a(p,t+1)\right\} \end{aligned}$$

By the induction hypothesis, we know $V_{\mathbf{0}}^b(p,t+1) \geq V_{\mathbf{0}}^a(p,t+1)$. It thus suffices to check $\int_{p'} V_{\mathbf{0}}^b(p',t+1)G^b(dp'|p) \geq \int_{p'} V_{\mathbf{0}}^a(p',t+1)G^a(dp'|p)$. Notice the following relations hold:

$$\begin{aligned} &\int_{p'} V_{\mathbf{0}}^b(p',t+1)G^b(dp'|p) - \int_{p'} V_{\mathbf{0}}^a(p',t+1)G^a(dp'|p) \\ =&\alpha_b \int_{p'} V_{\mathbf{0}}^b(p',t+1)G^I(dp'|p) + (1-\alpha_b)V_{\mathbf{0}}^b(p,t+1) \\ &- \alpha_a \int_{p'} V_{\mathbf{0}}^a(p',t+1)G^I(dp'|p) - (1-\alpha_a)V_{\mathbf{0}}^a(p,t+1) \\ \geq&\alpha_a \int_{p'} V_{\mathbf{0}}^b(p',t+1)G^I(dp'|p) + (1-\alpha_a)V_{\mathbf{0}}^b(p,t+1) \\ &- \alpha_a \int_{p'} V_{\mathbf{0}}^a(p',t+1)G^I(dp'|p) - (1-\alpha_a)V_{\mathbf{0}}^a(p,t+1) \end{aligned}$$

where the equality is by the definition of G^a and G^b. To see the inequality holds, recall that Lemma B.3(b) implies $\int_{p'} V_{\mathbf{0}}^b(p',t+1)G^b(dp'|p) \geq V_{\mathbf{0}}^b(p,t+1)$, which further implies $\int_{p'} V_{\mathbf{0}}^b(p',t+1)G^I(dp'|p) \geq V_{\mathbf{0}}^b(p,t+1)$ since $G^b(\cdot|p)$ is a weighted average of $G^I(\cdot|p)$ and $D(\cdot|p)$. The above inequality thus holds given $\alpha_b > \alpha_a$. Now, notice the induction hypothesis implies that the last expression above is indeed non-negative. We thus have the desired result. □

Next, we can show the following claim:

Claim *(b).* $H_{\mathbf{0}}^b(p,t) \leq 0 \Rightarrow H_{\mathbf{0}}^a(p,t) < 0$ for all pairs of (p,t) with $t < T$.

Proof for Claim (b). We can first prove $H_{\mathbf{0}}^b(p,t) \leq 0 \Rightarrow H_{\mathbf{0}}^a(p,t) \leq 0$. To show this, notice by Lemma B.4(d) we know that $H_{\mathbf{0}}^b(p,t) \leq 0$ implies $H_{\mathbf{0}}^b(p,t') < 0$ for all $t' > t$. With $\alpha = \alpha_b$, it is thus optimal for the dictator to stop recommendation from time t on given $p_t = p$, which leads to the optimal continuation value being zero. Hence $V_{\mathbf{0}}^b(p,t) = 0$. By Claim (a), this implies $V_{\mathbf{0}}^a(p,t) \leq 0$ and it is thus also optimal for the dictator to stop recommendation at (p,t) given $\alpha = \alpha_a$. This then implies $H_{\mathbf{0}}^a(p,t) \leq 0$.

Now, fix $t < T$. It suffices to rule out the possibility that $H_{\mathbf{0}}^b(p,t) \leq 0$ but $H_{\mathbf{0}}^a(p,t) = 0$. Supposing we do have $H_{\mathbf{0}}^b(p,t) \leq 0$ and $H_{\mathbf{0}}^a(p,t) = 0$, I will derive a contradiction below.

First, notice the following holds:

$$\begin{aligned} H_{\mathbf{0}}^b(p,t) = u(p) + \int_{p'} V_{\mathbf{0}}^b(p',t+1)[\alpha_b G^I(dp'|p) + (1-\alpha_b)D(dp'|p)] - V_{\mathbf{0}}^b(p,t+1) \\ = u(p) + \alpha_b \int_{p'} V_{\mathbf{0}}^b(p',t+1) G^I(dp'|p) \end{aligned}$$

The first equality holds by the definition of $H_{\mathbf{0}}$; the second equality holds because $H_{\mathbf{0}}^b(p,t) \leq 0$ implies it's optimal for the dictator to stop recommendation from time t on given $p_t = p$ (as is mentioned earlier) and thus $V_{\mathbf{0}}^b(p,t+1) = 0$.

The same argument also applies to $H_{\mathbf{0}}^a$. Thus $H_{\mathbf{0}}^a(p,t) = 0$ implies $H_{\mathbf{0}}^a(p,t) = u(p) + \alpha_a \int_{p'} V_{\mathbf{0}}^a(p',t+1) G^I(dp'|p)$.

Combining the above results, we then must have:

$$u(p) + \alpha_b \int_{p'} V_{\mathbf{0}}^b(p',t+1) G^I(dp'|p) \leq 0 \tag{B.16}$$

$$u(p) + \alpha_a \int_{p'} V_{\mathbf{0}}^a(p',t+1) G^I(dp'|p) = 0 \tag{B.17}$$

Notice $H_{\mathbf{0}}^b(p,t) \leq 0 \Rightarrow H_{\mathbf{0}}^b(p,t+1) < 0$ (by Lemma B.4(d)), which further implies that $u(p) < 0$.[53] Equation (B.17) then implies $\alpha_a \int_{p'} V_{\mathbf{0}}^a(p',t+1) G^I(dp'|p) > 0$. Since $\alpha_a < \alpha_b$ and $0 \leq V_{\mathbf{0}}^a(\cdot,\cdot) \leq V_{\mathbf{0}}^b(\cdot,\cdot)$ (by Claim (a)), we then must have $u(p) + \alpha_b \int_{p'} V_{\mathbf{0}}^b(p',t+1) G^I(dp'|p) > u(p) + \alpha_a \int_{p'} V_{\mathbf{0}}^a(p',t+1) G^I(dp'|p) = 0$, which violates equation (B.16). This contradiction implies that we cannot have $H_{\mathbf{0}}^b(p,t) \leq 0$ but $H_{\mathbf{0}}^a(p,t) = 0$.

□

With Claim (b), we can now prove the desired observation that $\eta_t^a > \eta_t^b$ for all $t \in [\hat{t}^b, T)$. Fix $t \in [\hat{t}^b, T)$. By Proposition 3 we must have threshold η_t^b to be optimal for the dictator given $\alpha = \alpha_b$. Under the full support condition in Assumption 2, this implies that $H_{\mathbf{0}}^b(\eta_t^b, t) = 0$.[54] Claim (b) then implies $H_{\mathbf{0}}^a(\eta_t^b, t) < 0$ and thus even the dictator's optimal threshold at time t given $\alpha = \alpha_a$ would be strictly greater than η_t^b. Thus $\eta_t^a > \eta_t^b$.

[53] If $u(p) = 0$, then recommending the product cannot harm and thus $H_{\mathbf{0}}^b(p,t+1)$ is at least zero.

[54] If not, then by the continuity of $H_{\mathbf{0}}^b$ we must have $H_{\mathbf{0}}^b(\cdot,t)$ to be strictly positive or strictly negative over some neighborhood of η_t^b, which implies η_t^b to be a suboptimal threshold for the dictator.

This completes the proof for Observation 2.

If $\widehat{t}^b = 1$, we are already done with proving the proposition. When $\widehat{t}^b > 1$, it remains to show Observation 3 below, which is the most central part of the proof.

Observation 3: When $\widehat{t}^b > 1$, we have: $\eta_1^a \geq \eta_1^b$ and $\eta_t^a > \eta_t^b$ for all $t \in (1, \widehat{t}^b)$.

When $\widehat{t}^b > 1$, we know the consumer's IC constraint in period 1 is binding given the period-1 threshold being η_1^b. It is thus not incentive compatible to choose a period-1 threshold less than η_1^b. We thus have $\eta_1^a \geq \eta_1^b$.

I now turn to prove $\eta_t^a > \eta_t^b$ for $t \in (1, \widehat{t}^b)$. This is done by constructing two belief processes $(p_t^a)_{t=1}^T$ and $(p_t^b)_{t=1}^T$ on the same probability space, where $(p_t^a)_{t=1}^T$ follows the transition rule under ϕ^a given $\alpha = \alpha_a$ and $(p_t^b)_{t=1}^T$ follows the transition rule under ϕ^b given $\alpha = \alpha_b$.

Specifically, fix a probability space on which a Markov process $(x_n)_{n=0}^\infty$ and a sequence of i.i.d. random variables $(\xi_t)_{t=1}^\infty$ with $\xi_t \sim \text{Uniform}[0,1]$ are defined. $(x_n)_{n=0}^\infty$ is independent from $(\xi_t)_{t=1}^\infty$ and satisfies:

$$x_0 \sim \mu_1$$
$$x_{n+1} | x_n \sim G^I(\cdot | x_n), \ \forall n$$

Intuitively, one can interpret x_n as the value that the platform's belief will take after receiving the n'th informative signal from consumers; ξ_t will serve as a randomization device deciding whether an informative signal will be generated after consumption at time t. For any $k \in \{a, b\}$, I now define process $(p_t^k)_{t=1}^T$ together with an auxiliary process $(n_t^k)_{t=1}^T$ by the following rule:

$$n_1^k = 0; \quad p_1^k = x_0 \tag{B.18}$$
$$n_{t+1}^k = n_t^k + \mathbb{1}_{\{p_t^k > \eta_t^k\}} \mathbb{1}_{\{\xi_t < \alpha_k\}}; \quad p_{t+1}^k = x_{n_{t+1}^k} \quad \forall t \tag{B.19}$$

Intuitively, given (α_k, ϕ^k), n_t^k tracks how many informative signals will have been recorded at the beginning of time t. It is added by 1 after each period if and only if consumption has been made in that period (i.e., $p_t^k > \eta_t^k$) and an informative signal has been generated (which is assumed to happen when $\xi_t < \alpha_k$).[55] Given that n_t^k informative signals have been received, p_t^k just equals to $x_{n_t^k}$, which reflects the posterior belief given those signals. It is easy to check that $(p_t^k)_{t=1}^{\tau}$ indeed satisfy the initial distribution and the transition rule of the belief process under policy ϕ^k given response rate α_k.

We can prove a couple of claims about the process of $\left((p_t^a, p_t^b)\right)_{t=1}^T$ constructed above.

[55] Under Assumption 2, what happens when $p_t^k = \eta_t^k$ does not matter since it has zero probability to occur.

Claim *(d).* Pick any $\tau \in (1, \widehat{t}^b)$ and assume $\eta_t^a \geq \eta_t^b$ for all $t < \tau$. Then $p_\tau^a < \eta_\tau^b \Rightarrow p_\tau^a = p_\tau^b$.
Proof for Claim (d). First notice that by Proposition 4, we know $\eta_\tau^b < \min_{t<\tau}\{\eta_t^b\}$. This implies that the process of $\big((p_t^b, n_t^b)\big)_{t=1}^\tau$ will stop once p_t^b hits into $[0, \eta_\tau^b]$ according to transition rule (B.19).

Now, suppose $p_\tau^a < \eta_\tau^b$. Then we have $x_{n_\tau^a} < \eta_\tau^b$. Since by assumption $\eta_t^a \geq \eta_t^b$ for all $t < \tau$ and $\alpha_b > \alpha_a$, it is easy to see that $n_\tau^b \geq n_\tau^a$ for sure.[56] This implies that there must exist a time $t' \leq \tau$ such that $n_{t'}^b = n_\tau^a$. At this time t', we then have $p_{t'}^b = x_{n_\tau^a} = p_\tau^a$. Since $p_\tau^a < \eta_\tau^b$, the process of $(p_t^b)_{t=1}^\tau$ will stop there and thus $p_\tau^b = p_{t'}^b = p_\tau^a$. □

Claim *(e).* Pick any $\tau \in (1, \widehat{t}^b)$. We have $\mathbb{P}(p_\tau^a > \eta_\tau^b > p_\tau^b) > 0$.
Proof for Claim (e). First, I show that $\mathbb{P}(p_1^a > \eta_\tau^b > p_\tau^b) > 0$. Supposing not, since $p_1^a = x_0 = p_1^b$, we then must have $\mathbb{P}(p_1^b > \eta_\tau^b > p_\tau^b) = 0$. This implies that $p_1^b > \eta_\tau^b \Rightarrow p_\tau^b > \eta_\tau^b$ a.s. Moreover, since $(\eta_t^b)_{t=1}^\tau$ is decreasing over time, $p_1^b \leq \eta_\tau^b$ would imply $p_1^b \leq \min_{t\leq\tau}\{\eta_t^b\}$ and thus $p_\tau^b = p_1^b \leq \eta_\tau^b$. We therefore have $p_1^b > \eta_\tau^b \Leftrightarrow p_\tau^b > \eta_\tau^b$ a.s. Given this, the following must hold:

$$\begin{aligned} 0 &= \mathbb{E}[u(p_\tau^b)\mathbb{1}_{\{p_\tau^b > \eta_\tau^b\}}] = \mathbb{E}[u(p_\tau^b)\mathbb{1}_{\{p_1^b > \eta_\tau^b\}}] = \mathbb{E}\big[u(p_1^b)\mathbb{1}_{\{p_1^b > \eta_\tau^b\}}\big] \\ &= \mathbb{E}\big[u(p_1^b)\mathbb{1}_{\{p_1^b > \eta_1^b\}}\big] + \mathbb{E}\big[u(p_1^b)\mathbb{1}_{\{\eta_\tau^b < p_1^b < \eta_1^b\}}\big] = \mathbb{E}\big[u(p_1^b)\mathbb{1}_{\{\eta_\tau^b < p_1^b < \eta_1^b\}}\big] < 0 \end{aligned}$$

which leads to a contradiction. The first equality above holds because the consumer's IC is binding at time τ under the scenario of (ϕ^b, α^b) (since $\tau < \widehat{t}^b$); the second equality holds because $p_1^b > \eta_\tau^b \Leftrightarrow p_\tau^b > \eta_\tau^b$ a.s as is shown earlier; the third equality holds because $u(\cdot)$ is affine and $\mathbb{E}[p_\tau^b | p_1^b] = p_1^b$; the fourth equality is trivial given $\eta_\tau^b < \eta_1^b$; the fifth equality holds because the consumer's IC is binding in period 1 under the scenario of (ϕ^b, α^b) (since $1 < \widehat{t}^b$) and thus $\mathbb{E}\big[u(p_1^b)\mathbb{1}_{\{p_1^b > \eta_1^b\}}\big] = 0$; the last inequality holds because η_1^b is no greater than the myopic optimal threshold and $\mathbb{P}(\eta_\tau^b < p_1^b < \eta_1^b) > 0$ under the full support assumption (i.e., Assumption 2). The contradiction implies that we must have $\mathbb{P}(p_1^a > \eta_\tau^b > p_\tau^b) > 0$.

Now, look into the event $\{p_1^a > \eta_\tau^b > p_\tau^b\}$. Notice that we can decompose it into:

$$\{p_1^a > \eta_\tau^b > p_\tau^b\} = \cup_{v\in\{0,1\}^\tau}\{p_1^a > \eta_\tau^b > p_\tau^b, (\mathbb{1}_{\{\xi_t < \alpha_b\}})_{t=1}^\tau = v\}$$

Since the union is over finitely many sets, the fact that the LHS has positive probability implies that for some $v \in \{0,1\}^\tau$, we have $\mathbb{P}(p_1^a > \eta_\tau^b > p_\tau^b, (\mathbb{1}_{\{\xi_t<\alpha_b\}})_{t=1}^\tau = v) > 0$. Notice that given $(\mathbb{1}_{\{\xi_t<\alpha_b\}})_{t=1}^\tau = v$, whether or not $p_1^a > \eta_\tau^b > p_\tau^b$ solely depends on the realizations of $(x_n)_{n=1}^\tau$. Define

$$X := \{(x_n)_{n=1}^\tau : p_1^a > \eta_\tau^b > p_\tau^b \text{ given } (\mathbb{1}_{\{\xi_t<\alpha_b\}})_{t=1}^\tau = v\}$$

[56] By construction, whenever $n_t^b = n_t^a$, we have $p_t^b = p_t^a$ and hence $n_{t+1}^a = n_t^a + 1 \Rightarrow n_{t+1}^b = n_t^b + 1$. Thus the sequence of $(n_t^a)_t$ can never surpass $(n_t^b)_t$.

Then, the independence between $(x_n)_{n=1}^{\tau}$ and $(\xi_t)_{t=1}^{\tau}$ implies $\mathbb{P}(p_1^a > \eta_\tau^b > p_\tau^b, (\mathbb{1}_{\{\xi_t<\alpha_b\}})_{t=1}^{\tau} = v) = \mathbb{P}\big((x_n)_{n=1}^{\tau} \in X\big)\mathbb{P}((\mathbb{1}_{\{\xi_t<\alpha_b\}})_{t=1}^{\tau} = v)$. Thus we must have $\mathbb{P}\big((x_n)_{n=1}^{\tau} \in X\big) > 0$.

Now, define $Y = \{(\xi_t)_{t=1}^{\tau} : (\mathbb{1}_{\{\xi_t<\alpha_b\}})_{t=1}^{\tau} = v, \xi_t > \alpha_a \forall t = 1, ..., \tau\}$. Since $\alpha_a < \alpha_b$, we obviously have $\mathbb{P}((\xi_t)_{t=1}^{\tau} \in Y) > 0$. Together with $\mathbb{P}\big((x_n)_{n=1}^{\tau} \in X\big) > 0$, we then have:

$$\mathbb{P}\big((x_n)_{n=1}^{\tau} \in X, (\xi_t)_{t=1}^{\tau} \in Y\big) = \mathbb{P}\big((x_n)_{n=1}^{\tau} \in X\big)\mathbb{P}((\xi_t)_{t=1}^{\tau} \in Y) > 0$$

Notice that by the construction of sets X and Y, we know $(x_n)_{n=1}^{\tau} \in X$ and $(\xi_t)_{t=1}^{\tau} \in Y$ together imply $p_\tau^a = p_1^a$ (since all those ξ_t are greater than α_a) and $p_1^a > \eta_\tau^b > p_\tau^b$. Therefore we have $\mathbb{P}(p_\tau^a > \eta_\tau^b > p_\tau^b) > 0$.

□

We are now ready to show $\eta_t^a > \eta_t^b$ for all $t \in (1, \hat{t}^b)$. Since I have shown $\eta_1^a \geq \eta_1^b$, by an easy induction argument, it suffices to show that $\eta_t^a \geq \eta_t^b \forall t < \tau \Rightarrow \eta_\tau^a > \eta_\tau^b$ for any $\tau \in (1, \hat{t}^b)$.

Fix $\tau \in (1, \hat{t}^b)$ and assume $\eta_t^a \geq \eta_t^b \forall t < \tau$. By using the results in Claims (d) and (e), we can show:

$$\begin{aligned}\mathbb{E}[u(p_\tau^b)\mathbb{1}_{\{p_\tau^b<\eta_\tau^b\}}] &= \mathbb{E}[u(p_\tau^b)\mathbb{1}_{\{p_\tau^b<\eta_\tau^b,\, p_\tau^a<\eta_\tau^b\}}] + \mathbb{E}[u(p_\tau^b)\mathbb{1}_{\{p_\tau^b<\eta_\tau^b\leq p_\tau^a\}}] \\ &< \mathbb{E}[u(p_\tau^b)\mathbb{1}_{\{p_\tau^b<\eta_\tau^b,\, p_\tau^a<\eta_\tau^b\}}] = \mathbb{E}[u(p_\tau^a)\mathbb{1}_{\{p_\tau^a<\eta_\tau^b\}}]\end{aligned}$$

The first equality is a trivial identity. To see the why the inequality holds, notice $u(p_\tau^b) < 0$ for any $p_\tau^b < \eta_\tau^b$ since η_τ^b is no greater than the myopic optimal threshold. Together with the fact that $\mathbb{P}(p_\tau^b < \eta_\tau^b \leq p_\tau^a) > 0$ (by Claim (e)), this implies $\mathbb{E}[u(p_\tau^b)\mathbb{1}_{\{p_\tau^b<\eta_\tau^b\leq p_\tau^a\}}] < 0$. The last equality above holds because $p_\tau^a < \eta_\tau^b \Rightarrow p_\tau^a = p_\tau^b$ (by Claim (d)).

Notice $\mathbb{E}[u(p_\tau^a)] = \mathbb{E}[u(p_0)] = \mathbb{E}[u(p_\tau^b)]$ by the law of iterated expectation. Thus the above result $\mathbb{E}[u(p_\tau^b)\mathbb{1}_{\{p_\tau^b<\eta_\tau^b\}}] < \mathbb{E}[u(p_\tau^a)\mathbb{1}_{\{p_\tau^a<\eta_\tau^b\}}]$ implies $\mathbb{E}[u(p_\tau^b)\mathbb{1}_{\{p_\tau^b>\eta_\tau^b\}}] > \mathbb{E}[u(p_\tau^a)\mathbb{1}_{\{p_\tau^a>\eta_\tau^b\}}]$. Because $\mathbb{E}[u(p_\tau^b)\mathbb{1}_{\{p_\tau^b>\eta_\tau^b\}}] = 0$ (since $\tau < \hat{t}^b$ and thus consumer's IC must be binding at time τ under the scenario of (ϕ^b, α_b)), we then must have $\mathbb{E}[u(p_\tau^a)\mathbb{1}_{\{p_\tau^a>\eta_\tau^b\}}] < 0$. This implies that η_τ^a must be strictly greater than η_τ^b to obey consumer's IC at time τ under policy ϕ^a given $\alpha = \alpha_a$. This completes the proof for $\eta_t^a > \eta_t^b$ for all $t \in (1, \hat{t}^b)$.

B.8 Proof for Proposition 6

Proof. Based on my discussion in the main text, the designer's problem is equivalent to:

$$\begin{aligned}&\max_{\phi\in\Phi}\Big\{\sum_{t=1}^{T}\mathbb{E}_\phi[a_t u(p_t)]\Big\} \\ &\text{s.t. } \mathbb{E}_\phi[a_t u(p_t)] \geq 0 \quad \forall t = 1, ..., T \\ &\qquad p_{t+1}|p_t, a_t \sim a_t\big[\rho G(\cdot|p_t) + (1-\rho)D(\cdot|p_t)\big] + (1-a_t)D(\cdot|p_t),\ p_1 \sim \mu_1\end{aligned}$$

With ρ replaced by α and G replaced by G^I, this is equivalent to the designer's problem studied in Appendix B.7. Thus the effect of an increment in ρ here is equivalent to the effect of an increment in α there. The result is hence directly implied by Proposition 5.

Q.E.D.

B.9 Proof for Proposition 7

Proof. Given platform biases $\beta_a < \beta_b$, let $V^a_{\mathbf{0}}$ and $V^b_{\mathbf{0}}$ denote the corresponding value functions of the dictator's problem (i.e., with $\lambda = \mathbf{0}$), as is defined in Section B.4; let $H^a_{\mathbf{0}}$ and $H^b_{\mathbf{0}}$ denote the associated H-functions (with $\lambda = \mathbf{0}$), i.e.,

$$H^k_{\mathbf{0}}(p,t) := u(p) + \beta_k + \int_{p'} V^k_{\mathbf{0}}(p', t+1)G(dp'|p) - V^k_{\mathbf{0}}(p, t+1),\ \forall t = 1, ..., T,\ \forall k = a, b$$

Also let $(\eta^d_t(\beta_a))^T_{t=1}$ and $(\eta^d_t(\beta_b))^T_{t=1}$ denote the corresponding optimal thresholds for the dictator's problem. Then it is easy to see that the results of Lemmas B.1 – B.4 (in particular applied to the case with $\lambda = \mathbf{0}$) extend here. Moreover, the induction formula (B.14) for H extends with the term $u(p)$ replaced by $u(p) + \beta$, which when $\lambda = \mathbf{0}$ in particular implies for $k = a, b$:

$$H^k_{\mathbf{0}}(p,t) = \min\{H^k_{\mathbf{0}}(p, t+1), 0\} + \int_{p'} \max\{H^k_{\mathbf{0}}(p', t+1), 0\}G(dp'|p),\ \forall t < T \qquad \text{(B.20)}$$

We have the following observation:

Claim. $\eta^d_t(\beta_a) > \eta^d_t(\beta_b)$ for all t.

Proof for the claim. Under the full support condition in Assumption 2, by the dynamic programming result (Lemma B.2) it suffices to show $H^b_{\mathbf{0}}(p,t) > H^a_{\mathbf{0}}(p,t)\ \forall p$ for all t. This can be proved by backward induction in t. When $t = T$, we have $H^b_{\mathbf{0}}(p,t) - H^a_{\mathbf{0}}(p,t) = \beta_b - \beta_a > 0$ as desired. For any $t < T$, by the induction hypothesis we have $H^b_{\mathbf{0}}(p, t+1) > H^a_{\mathbf{0}}(p, t+1), \forall p$. Equation (B.20) then implies $H^b_{\mathbf{0}}(p,t) > H^a_{\mathbf{0}}(p,t), \forall p$.[57] □

The claim implies that the dictator's optimal policy features lower standards when β is larger. Thus the dictator's optimal policy given $\beta = \beta_b$ becomes incentive compatible at time t only if its counterpart given $\beta = \beta_a$ has become incentive compatible, which implies $\widehat{t}(\beta_b) \geq \widehat{t}(\beta_a)$ by the construction of $\widehat{t}$ in Definition 2. This proves part (a). Also by the construction of the optimal policy in Definition 2, we know that the optimal thresholds before time $\widehat{t}(\beta_a)$ must be just high enough such that the consumer's IC constraint is

[57] One can show the inequality is strict. When p is such that $H^b_{\mathbf{0}}(p, t+1) \leq 0$, we have $\min\{H^b_{\mathbf{0}}(p, t+1), 0\} > \min\{H^a_{\mathbf{0}}(p, t+1), 0\}$. When p is such that $H^b_{\mathbf{0}}(p, t+1) > 0$, we must have $G(\{p' : H^b_{\mathbf{0}}(p', t+1) > 0\}|p) > 0$, and thus $\int_{p'} \max\{H^b_{\mathbf{0}}(p', t+1), 0\}G(dp'|p) > \int_{p'} \max\{H^a_{\mathbf{0}}(p', t+1), 0\}G(dp'|p)$.

satisfied as equality. Thus we have $\eta_t^*(\beta_b) = \eta_t^*(\beta_a)$ for all $t < \widehat{t}(\beta_a)$. This proves part (b).

To show part (c), first notice that when $t \geq \widehat{t}(\beta_b)$, in both situations the optimal policy will follow the dictator's optimal policy. The fact that $\eta_t^*(\beta_b) < \eta_t^*(\beta_a)$ is then directly implied by the claim above. It thus suffices to assume $\widehat{t}(\beta_a) < \widehat{t}(\beta_b)$ and prove the result for $t \in \big[\widehat{t}(\beta_a), \widehat{t}(\beta_b)\big)$. Notice that by Proposition 4, $\eta_t^*(\beta_a)$ is strictly increasing over $t \in \big[\widehat{t}(\beta_a), \widehat{t}(\beta_b)\big)$, while $\eta_t^*(\beta_b)$ is strictly decreasing over $t \in \big[\widehat{t}(\beta_a), \widehat{t}(\beta_b)\big)$. It thus suffices to show $\eta_t^*(\beta_b) \leq \eta_t^*(\beta_a)$ for $t = \widehat{t}(\beta_a)$. This is true because the assumption that $\widehat{t}(\beta_a) < \widehat{t}(\beta_b)$ implies that $\eta^*_{\widehat{t}(\beta_a)}(\beta_b)$ must be just high enough to satisfy the consumer's IC constraint (given the distribution of $p_{\widehat{t}(\beta_a)}$, which is the same under the optimal policy in both situations by part (b) already proved). Since $\eta_{\widehat{t}(\beta_a)} = \eta^*_{\widehat{t}(\beta_a)}(\beta_a)$ is already incentive compatible, we cannot have $\eta^*_{\widehat{t}(\beta_a)}(\beta_b) > \eta^*_{\widehat{t}(\beta_a)}(\beta_a)$. This concludes the proof for part (c). *Q.E.D.*

B.10 Proof for Proposition A.1

In the following proof, I assume $\{Q_z\}_{z \in Z}$ and the conditional distributions of s_i $(i \geq 1)$ conditional on $\tilde{\theta}$ are all continuous distributions, so the dominating measure for their densities is chosen as the Lebesgue measure. In the general case, the proof remains the same with Lebesgue measure replaced by proper dominating measures on $\mathbb{R}$ (e.g., counting measure for discrete distributions).

Proof. For any $z \in Z$ and $s \in S$, I define $\psi_{(z,s)}$ as a probability density over $\mathbb{R}$ such that

$$\psi_{(z,s)}(\theta) = \frac{q_z(\theta)\ell(s|\theta)}{\int q_z(\theta)\ell(s|\theta)d\theta}$$

That is, $\psi_{(z,s)}$ is the density function of the posterior about $\tilde{\theta}$ computed from Bayes rule given prior Q_z and consumption-generated signal realization s.

Claim *(a).* For any $x, y \in Z$ and $s_a, s_b \in S$, we have $Q_y \geq_{LR} Q_x$ and $s_b \geq s_a$ together imply $\psi_{(y,s_b)} \geq_{LR} \psi_{(x,s_a)}$.

Proof for Claim (a). Assume $Q_y \geq_{LR} Q_x$ and $s_b \geq s_a$. By the definition of ψ, we have:

$$\frac{\psi_{(y,s_b)}(\theta)}{\psi_{(x,s_a)}(\theta)} = \frac{q_y(\theta)}{q_x(\theta)} \cdot \frac{\ell(s_b|\theta)}{\ell(s_a|\theta)}$$

Since $Q_y \geq_{LR} Q_x$ and $\ell(\cdot|\theta)$ increases in likelihood-ratio order in θ, both fractions on the right-hand-side are increasing in θ. Thus $\psi_{(y,s_b)} \geq_{LR} \psi_{(x,s_a)}$. □

Now, I show the following observation:

Claim *(b).* Assume λ_t is non-increasing over t. Then, for any $x, y \in Z$, we have $Q_y \geq_{LR} Q_x \Rightarrow H_\lambda(y,t) \geq H_\lambda(x,t)$ for all t.

Proof for Claim (b). I show by backward induction in t. The result holds with $t = T$ since $H_\lambda(z, T) = (1+\lambda_T)\int \theta dQ_z(\theta)$. Now, assuming the result holds for all periods since time $t+1$, I show it for period t. Recall that equation (B.14) derived in Section B.4 implies

$$\begin{aligned} H_\lambda(z,t) =& (\lambda_t - \lambda_{t+1})u(z) + \min\{H_\lambda(z,t+1), 0\} \\ &+ \int_{z'\in Z} \max\{H_\lambda(z', t+1), 0\}[\rho G(dz'; z) + (1-\rho)D(dz'; z)] \\ =& (\lambda_t - \lambda_{t+1})u(z) + \rho \min\{H_\lambda(z,t+1), 0\} + (1-\rho)H_\lambda(z, t+1) \\ &+ \rho \int_{z'\in Z} \max\{H_\lambda(z', t+1), 0\} G(dz'; z) \end{aligned}$$

(Since we have random consumer arrivals with arrival rate ρ, the transition kernel G in equation (B.14) is replaced with $\rho G + (1-\rho)D$.)

Pick any $x, y \in Z$ s.t. $Q_y \geq_{LR} Q_x$. We obviously have $(\lambda_t - \lambda_{t+1})u(y) \geq (\lambda_t - \lambda_{t+1})u(x)$ given the assumption that λ_t is non-increasing in t. Moreover, the induction hypothesis implies $\min\{H_\lambda(y, t+1), 0\} \geq \min\{H_\lambda(x, t+1), 0\}$ and $(1-\rho)H_\lambda(y, t+1) \geq (1-\rho)H_\lambda(x, t+1)$. It then suffices to show $\int_{z'\in Z} \max\{H_\lambda(z', t+1), 0\}G(dz'; y) \geq \int_{z'\in Z} \max\{H_\lambda(z', t+1), 0\}G(dz'; x)$ below.

Notice in the current setting, state z matters only through the belief it represents. With slight abuse of notation, I write $H_\lambda(q_z, t+1) = H_\lambda(z, t+1)$. Then, we have:

$$\begin{aligned} \int_{z'\in Z} \max\{H_\lambda(z', t+1), 0\}G(dz'; y) &= \int_\theta \Big[\int_s \max\{H_\lambda(\psi_{(y,s)}, t+1), 0\}\ell(s|\theta)ds \Big] q_y(\theta) d\theta \\ &\geq \int_\theta \Big[\int_s \max\{H_\lambda(\psi_{(x,s)}, t+1), 0\}\ell(s|\theta)ds \Big] q_y(\theta) d\theta \\ &\geq \int_\theta \Big[\int_s \max\{H_\lambda(\psi_{(x,s)}, t+1), 0\}\ell(s|\theta)ds \Big] q_x(\theta) d\theta \\ &= \int_{z'\in Z} \max\{H_\lambda(z', t+1), 0\}G(dz'; x) \end{aligned}$$

where the two equalities hold by the definition of $\psi_{(z,s)}$. The first inequality holds due to the induction hypothesis and that Claim (a) above implies $\psi_{(y,s)} \geq_{LR} \psi_{(x,s)}$. To see the second inequality, notice Claim (a) implies that $\psi_{(x,s)}$ increases in likelihood-ratio order in s. Together with the induction hypothesis, this implies that $\max\{H_\lambda(\psi_{(x,s)}, t+1), 0\}$ increases in s, which further implies that $\int_s \max\{H_\lambda(\psi_{(x,s)}, t+1), 0\}\ell(s|\theta)ds$ increases in θ since $\ell(\cdot|\theta)$ increases in likelihood-ratio order in θ. The inequality is hence implied by $q_y \geq_{LR} q_x$. □

Now, I slightly strengthen both the condition and the conclusion in Claim (b).

***Claim** (c).* Assume λ_t is non-increasing over t. Then, for any $x, y \in Z$, we have $Q_y \geq_{LR} Q_x$ and $\int_\theta \theta dQ_y(\theta) > \int_\theta \theta dQ_x(\theta)$ together imply $H_\lambda(y, t) > H_\lambda(x, t)$ for all t.

Proof for Claim (c). I show by backward induction in t. The result holds with $t = T$ since $H_\lambda(z,T) = (1+\lambda_T)\int \theta dQ_z(\theta)$. Now, assuming the result holds for all periods since time $t+1$, I show it for period t. By the same argument as in the proof of Claim (b), we have

$$H_\lambda(z,t) = (\lambda_t - \lambda_{t+1})u(z) + \rho \min\{H_\lambda(z,t+1),0\} + (1-\rho)H_\lambda(z,t+1) + \rho \int_{z' \in Z} \max\{H_\lambda(z',t+1),0\}G(dz';z)$$

and that for any x and y satisfying the conditions in the claim: (i) $(\lambda_t - \lambda_{t+1})u(y) \geq (\lambda_t - \lambda_{t+1})u(x)$; (ii) $\min\{H_\lambda(y,t+1),0\} \geq \min\{H_\lambda(x,t+1),0\}$; (iii) $\int_{z'\in Z} \max\{H_\lambda(z',t+1),0\}G(dz';y) \geq \int_{z'\in Z} \max\{H_\lambda(z',t+1),0\}G(dz';x)$. Moreover, the induction hypothesis directly imply that $(1-\rho)H_\lambda(y,t+1) > (1-\rho)H_\lambda(x,t+1)$ for $\rho < 1$. These together imply $H_\lambda(y,t) > H_\lambda(x,t)$ as is desired. □

Now, define functions V_λ and H_λ in the same way as in Appendix B.4, but with p_t replaced with z_t. Then the dynamic programming result – Lemma B.2 – still applies to the current setting, because its proof only relies on property (P1) in Lemma 1, which has its counterpart in Lemma A.1. Given the result of Claim (c) and Lemma A.2, Lemma B.2(b) implies that any solution to the Lagrangian optimization $\max_\phi \mathcal{L}(\phi;\lambda^*)$ (with λ^* solves the dual problem) is almost surely equivalent to some ϕ^* satisfying the property specified in Proposition A.1. The proposition hence holds by the duality result in Lemma 2.

Q.E.D.

References

Altman, E. (1999). *Constrained markov decision processes: stochastic modeling*. Routledge.

Bahar, G., Smorodinsky, R., & Tennenholtz, M. (2015). Economic recommendation systems. *arXiv preprint arXiv:1507.07191*.

Bahar, G., Smorodinsky, R., & Tennenholtz, M. (2021). Recommendation systems and self motivated users. *arXiv preprint arXiv:1807.01732*.

Ball, I. (2019). Dynamic information provision: Rewarding the past and guiding the future. *Available at SSRN 3103127*.

Banerjee, A. V. (1992). A simple model of herd behavior. *The quarterly journal of economics*, *107*(3), 797–817.

Beutler, F. J., & Ross, K. W. (1985). Optimal policies for controlled markov chains with a constraint. *Journal of mathematical analysis and applications*, *112*(1), 236–252.

Beutler, F. J., & Ross, K. W. (1986). Time-average optimal constrained semi-markov decision processes. *Advances in Applied Probability*, *18*(2), 341–359.

Bikhchandani, S., Hirshleifer, D., & Welch, I. (1992). A theory of fads, fashion, custom, and cultural change as informational cascades. *Journal of political Economy*, *100*(5), 992–1026.

Bolton, P., & Harris, C. (1999). Strategic experimentation. *Econometrica*, *67*(2), 349–374.

Che, Y.-K., & Hörner, J. (2018). Recommender systems as mechanisms for social learning. *The Quarterly Journal of Economics*, *133*(2), 871–925.

Chen, B., Frazier, P., & Kempe, D. (2018). Incentivizing exploration by heterogeneous users. In *Conference on learning theory* (pp. 798–818).

Durrett, R. (2019). *Probability: theory and examples* (Vol. 49). Cambridge university press.

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

Ely, J. C., & Szydlowski, M. (2020). Moving the goalposts. *Journal of Political Economy*, *128*(2), 468–506.

Feinberg, E. A., & Piunovskiy, A. B. (2000). Multiple objective nonatomic markov decision processes with total reward criteria. *Journal of mathematical analysis and applications*, *247*(1), 45–66.

Glazer, J., Kremer, I., & Perry, M. (2021). The wisdom of the crowd when acquiring information is costly. *Management Science*, *67*(10), 6443–6456.

Hörner, J., & Skrzypacz, A. (2017). Learning, experimentation and information design. *Advances in Economics and Econometrics*, *1*, 63–98.

Immorlica, N., Mao, J., Slivkins, A., & Wu, Z. S. (2019). Bayesian exploration with heterogeneous agents. In *The world wide web conference* (pp. 751–761).

Kamenica, E., & Gentzkow, M. (2011). Bayesian persuasion. *American Economic Review*, *101*(6), 2590–2615.

Keller, G., Rady, S., & Cripps, M. (2005). Strategic experimentation with exponential bandits. *Econometrica*, *73*(1), 39–68.

Komiyama, J., & Noda, S. (2021). Deviation-based learning. *arXiv preprint arXiv:2109.09816*.

Kremer, I., Mansour, Y., & Perry, M. (2014). Implementing the "wisdom of the crowd". *Journal of Political Economy*, *122*(5), 988–1012.

Küçükgül, C., Özer, Ö., & Wang, S. (2022). Engineering social learning: Information design of time-locked sales campaigns for online platforms. *Management Science*, *68*(7), 4899–4918.

Liang, Y., Sun, P., Tang, R., & Zhang, C. (2020). *Efficient resource allocation contracts to reduce adverse events* (Tech. Rep.). Working Paper.

Lorecchio, C. (2021). Pursuading crowds. *unpublished manuscript*.

Lorecchio, C., & Monte, D. (2021). Dynamic information design under constrained communication rules. *American Economic Journal: Microeconomics*.

Luenberger, D. G. (1997). *Optimization by vector space methods.* John Wiley & Sons.

Mansour, Y., Slivkins, A., & Syrgkanis, V. (2020). Bayesian incentive-compatible bandit exploration. *Operations Research*, *68*(4), 1132–1161.

Mansour, Y., Slivkins, A., Syrgkanis, V., & Wu, Z. S. (2016). Bayesian exploration: Incentivizing exploration in bayesian games. *arXiv preprint arXiv:1602.07570*.

Müller, A., & Stoyan, D. (2002). *Comparison methods for stochastic models and risks* (Vol. 389). Wiley.

Myerson, R. B. (1986). Multistage games with communication. *Econometrica: Journal of the Econometric Society*, 323–358.

Orlov, D., Skrzypacz, A., & Zryumov, P. (2020). Persuading the principal to wait. *Journal of Political Economy*, *128*(7), 2542–2578.

Papanastasiou, Y., Bimpikis, K., & Savva, N. (2018). Crowdsourcing exploration. *Management Science*, *64*(4), 1727–1746.

Rayo, L., & Segal, I. (2010). Optimal information disclosure. *Journal of Political Economy*, *118*(5), 949–987.

Renault, J., Solan, E., & Vieille, N. (2017). Optimal dynamic information provision. *Games and Economic Behavior*, *104*, 329–349.

Smith, L., & Tian, J. (2018). Informational inertia. *unpublished mimeo*.

Smolin, A. (2021). Dynamic evaluation design. *American Economic Journal: Microeconomics*, *13*(4), 300–331.