Essays in Marketing Innovations

by

Keyan Li

B.Econ., Economics, Fudan University (2015)
M.S., Theoretical Economics, Tsinghua University (2018)
S.M., Management Research, Massachusetts Institute of Technology (2022)

Submitted to the Department of Management
in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY IN MANAGEMENT

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

May 2024

Authored by: Keyan Li
Department of Management
April 25, 2024

Certified by: Duncan Simester
Department of Management
Thesis Supervisor

Certified by: Juanjuan Zhang
Department of Management
Thesis Supervisor

Accepted by: Eric So
Professor, Global Economics and Finance
Faculty Chair, MIT Sloan PhD Program

Essays in Marketing Innovations

by

Keyan Li

Submitted to the Department of Management
on April 25, 2024 in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY IN MANAGEMENT

ABSTRACT

The dissertation consists of three chapters on understanding marketing innovations, including targeted marketing and new product development.

The first chapter studies a novel targeting problem that many firms face and develops a new method for targeting experimentation. Adaptive learning policies that guide how firms trade off acquiring new information to improve a current targeting policy, versus exploiting the current policy to harvest, typically focus on settings in which customers arrive individually, in a frequent sequence. However, in practice, firms often conduct marketing campaigns in batches, in which they target a large group of customers with personalized marketing actions together. This has an important implication for how firms resolve the tradeoff between acquiring new information and exploiting the current policy. The large number of customers in each batch (campaign) introduces an information externality: the incremental information contributed by a single customer depends upon the assignment decisions for other customers in the batch. We investigate how to optimally acquire and coordinate information in these settings. The algorithm we propose uses Gaussian processes to estimate the value of incremental information, while accounting for the information externality between customers in the same batch. Findings are validated using data from a field experiment.

The second chapter studies customer demand in a non-market-oriented economy. The economics and marketing literature has primarily focused on market economies and studied factors such as price and advertising when analyzing customer demand. However, in non-market-oriented economies, social factors like corruption can have a significant influence on customer decisions. In particular, this paper focuses on the demand for luxury products, which are widely used for gift-giving and even bribery in emerging markets. One possible mechanism is that when the relative size of non-market-oriented sectors in the local economy increases, luxury products can be used to identify those who have a higher willingness to pay for scarce resources. As a result, the demand for luxury products moves together with the degree of corruption. By leveraging natural experiments of top-down anti-corruption campaigns that temporarily halt this channel, an empirical study is performed using a comprehensive dataset that covers the sales of all cigarette brands and the local social environment in China. The results suggest that these social factors can have an unanticipated impact on the demand for luxury products.

The third chapter studies how customer search can stifle product innovations. Conventional wisdom suggests that when an incumbent fails to innovate, there is a greater risk to the incumbent of competition from other innovators. I show conditions in which the opposite is true: by delaying innovation, an incumbent can create entry barriers that deter innovation by competitors. Consequently, both competition and innovation are suppressed. The key insight driving these outcomes is that customer search is endogenous, and absence of innovation today can disincentivize customers from searching in the future. Since customers need to search to discover innovations, when they search less, it both creates entry barriers for competitors, and reduces the competitors' incentives to innovate. Postponing innovation can benefit incumbents if it motivates customers to search less, and thus competitors to innovate less. Notably, I show that searching less is a rational customer response.

Thesis supervisor: Duncan Simester
Title: NTU Professor of Marketing; Professor, Marketing

Thesis supervisor: Juanjuan Zhang
Title: John D. C. Little Professor of Marketing; Professor, Marketing

Acknowledgments

Pursuing a doctoral degree at MIT has been an incredible journey, filled with challenges and triumphs, and I am deeply grateful to the many people who have supported and guided me along the way.

First and foremost, I want to express my heartfelt gratitude to my advisors, Duncan Simester and Juanjuan Zhang. I am indebted to their investment in me. They shaped my research tastes, standards, and perspectives. They guided me (and I am still learning) on how to own an insight and develop a dream paper from it. Duncan and Juanjuan have been there for me at every stage of my doctoral study, helping me go beyond my own limits and become the best version of myself. Duncan was always available and extremely sharp whenever I needed help, and he taught me how to identify first-order problems that matter to both academics and businesses. Juanjuan was the reason I chose MIT; her philosophy, poetry, passion, and precision have been inspiring and encouraging me since my first day in the program. She taught me the art of research and always fought her way to support me.

I also want to thank my dissertation committee member, Tony Ke, for the numerous delightful brainstorming discussions and his invaluable, unbiased career advice. And to Birger Wernerfelt, my academic grandfather and role model, thank you for your unwavering support in all possible ways. I will never forget your words of wisdom and your confidence in me, and I will try my best to pass your academic thoughts and legacy on to the next generation.

My academic journey wouldn't have begun without Qiaowei Shen and Alex White. I am very fortunate to have Qiaowei and Alex as my mentors and close friends. Qiaowei introduced me to the field of marketing. Alex taught me how to think like an economist.

I extend my gratitude to all the other MIT Marketing faculty and doctoral students. My research projects have greatly benefited from the insightful comments and discussions from Rahul Bhui, Dean Eckles, John Hauser, Drazen Prelec, Dave Rand, and Catherine Tucker. I also thank Sharmila Chatterjee and Renee Gosline for their help in developing my teaching skills and experience, and Jim Orlin for offering me my first teaching assistantship at MIT.

Boston has become my second home. James, Yuting, Shuyi, Magie, Jinlin, Tianshuo, Chang, Lei, Jenn, and Xinyu, thank you for sharing countless unforgettable moments with me and making my memories of the city so lovely.

Finally, I want to thank my parents and my partner for their boundless love and support. You have been my rock throughout this journey, and I am forever grateful.

Contents

List of Figures

List of Tables

Chapter 1

Targeted Marketing with Large Batches

Abstract

Adaptive learning policies that guide how firms trade off acquiring new information to improve a current targeting policy, versus exploiting the current policy to harvest, typically focus on settings in which customers arrive individually, in a frequent sequence. However, in practice, firms often conduct marketing campaigns in batches, in which they target a large group of customers with personalized marketing actions together. This has an important implication for how firms resolve the tradeoff between acquiring new information and exploiting the current policy. The large number of customers in each batch (campaign) introduces an information externality: the incremental information contributed by a single customer depends upon the assignment decisions for other customers in the batch. We investigate how to optimally acquire and coordinate information in these settings. The algorithm we propose uses Gaussian processes to estimate the value of incremental information, while accounting for the information externality between customers in the same batch. We validate our findings using data from a field experiment.

1.1 Introduction

Many firms send personalized marketing actions to a targeted set of customers. After each campaign, the firm can use the new information from that campaign to improve the policy it uses in future campaigns.[1] This introduces an adaptive learning problem. For the customers in the current campaign, the firm needs to balance *earning* more today (exploiting), with *learning* more today (exploring; to improve the policy in the next campaigns).

In each marketing campaign, a firm typically engages with a batch of heterogeneous customers. Firms personalize a marketing action for every customer in the batch *simultaneously*, and customer responses only arrive after decisions are made for the entire batch. The firm cannot learn from any one customer before making decisions for all other customers in the same batch. This is different from the classic adaptive learning setting, where the firm interacts with individual customers sequentially, and learns the response *before* deciding upon the action for the next customer.

Operating and learning in batches is a feature of many targeted marketing campaigns, in both non-digital and digital channels. One explanation for the use of batches is seasonality. For example, many retailers conduct back-to-school campaigns at the end of the summer, or Holiday campaigns at the end of the calendar year, with many (if not all) of the customers arriving close in time. Similarly, get-out-to-vote political campaigns, are generally conducted in batches coinciding with each election season. A second explanation is that there is a fixed cost of planning and implementing campaigns, and batches help to lower these fixed costs.[2] Finally, technological challenges may force firms to learn in batches. The alternative is to engage in automated online machine learning, which requires sophisticated infrastructure to observe new outcomes, combine new outcomes with past data, update the targeting policy,

[1] There is room for policy improvement because estimated policy has uncertainty that can be alleviated. We discuss it in Section 1.3.

[2] This appears to be the explanation for why the wholesale membership club, which provided the data we use in our empirical validation, conducts a spring campaign and a fall campaign to prospect for new customers.

and then execute on the updated policy on the next customer, all in real time. Many firms lack the connected automated infrastructure required to accomplish this. Batch learning does not require that these operations occur in real time, and so does not require the same level of sophisticated infrastructure.

The batch feature of the learning process has an important implication: the incremental information the firm learns from one customer depends upon the actions it takes with similar customers. For example, assume there are three customers in this period's batch, and the firm sends a promotion to all three of them. When the three customers are similar (in terms of features or behaviors) and receive an identical marketing action, the information the firm learns from one customer overlaps with the information that it learns from the other two customers. If the firm is focused on acquiring new information, it may learn more or save opportunity costs by avoiding this information duplication, and varying its actions across the three customers. This highlights an externality between customers in the same batch; the incremental information that you learn from taking an action with a customer depends upon the actions you take with similar customers. We label this an "*information externality.*"

The information externality leads to a combinatorial problem. The value of eliciting one type of information from any customer depends on what types of information have been elicited from all other similar customers. The action assignments are therefore interdependent, and need to be jointly optimized.

A classic adaptive learning solution does not account for this information externality because it is designed to optimize actions for customers individually, and fails to consider externalities between customers in the same batch.[3] Identifying and accounting for this information externality between similar customers is the primary contribution of this paper, and distinguishes the paper from related research in marketing, computer science and other fields.

[3]For example, multi-armed contextual bandits [1] and knowledge gradient algorithms [2].

Different Problems Magnify Different Externalities

We will refer to this class of targeted marketing problems with large batches of customers as "batch targeting" problems. When comparing them with other adaptive learning problems, we need to distinguish between two types of externalities: inter-temporal and within-batch. First there is an externality across periods; actions and outcomes for a current customer provide information that can help design policies for future customers. Second, when there are multiple customers in each batch, information externalities arise between customers in the same batch. Both externalities are complex, because they depend upon similarities between customers and the actions taken with different customers. It is generally infeasible to fully solve both externalities at the same time, so we need to prioritize which externality to focus on.

For some problems, the information externality (within-batch) is less important. An extreme example is that, if a single customer arrives each period, there is no information externality between customers within the period. In contrast, if there are a large number of customers each period, the importance of the information externality is magnified.

In our batch targeting problem, we approximate the inter-temporal problem by assuming that the firm looks only one batch ahead. This one-step ahead approximation is commonly used in the Bayesian Optimization literature.[4] The approximation also appears reasonable in marketing practice. For example, when demand is seasonal, there is typically a long time-interval between campaigns, which discounts inter-temporal externalities in distant future periods. In addition, if there is a risk that customer responses to a firm's marketing actions will be non-stationarity, then the information learned in the current period maybe relevant next period, but may not be relevant in the distant future.

Moreover, in many firms, managers are regularly rotated between roles. As a result, while a firm may care about outcomes in future periods, its managers may not. A manager's

[4]See for example [2]–[5]. Classic Bayesian optimization solutions only have the current and next period rewards in their respective acquisition functions.

investment in exploration this season may help to satisfy the manager's personal performance goals next season. However, if the manager expects to be rotated into a new role after next period, then the manager has little incentive to forgo performance this season in order to increase performance when they are no longer in the role.

Classic adaptive learning methods are designed to focus solely on inter-temporal externalities. In contrast, in batch targeting problems the focus shifts from fully resolving inter-temporal externalities to addressing information externalities within a batch.[5]

Proposed Solution

We formalize our proposed solution as an algorithm, which we label "One-step Look Ahead Targeting" (`OLAT`). We estimate the information value of exploring any customer with a nonparametric Bayesian approach that uses Gaussian processes, which allows for quantifying uncertainty pointwisely. The algorithm can in theory provide an exact solution to the one-step look ahead batch targeting problem. However, doing so requires solving a combinatorial problem; if a batch has 1,000 customers, and the firm is choosing from just two actions for each customer, the solution includes $2^{1,000}$ potential solutions. We need to solve the problem jointly, and cannot simply decompose the problem into 1,000 binary problems. The `OLAT` algorithm overcomes this complexity by grouping neighboring customers into clusters, and focuses on the information externalities within clusters.

It might appear that we also need to be concerned about selection; because this period's assignments vary with noise in past period's outcomes, and are not purely random. However, the selection decisions are based upon observed covariates, and the proposed framework uses Bayesian approach to model the data-generating process. As a result, we are able to show that the algorithm is not susceptible to selection issues.

[5] While we have distinguished the batch targeting problem from classic adaptive learning problems, other research in marketing and computer science has extended the research on multi-armed bandits to batch settings. In the literature review (Section 1.2), we will review these papers, and contrast them with the current paper.

Contribution

Our first contribution is to identify the batch targeting problem and describe why it is conceptually different than other classic adaptive learning problems. In particular, this is the first paper to identify information externalities within an adaptive heterogeneous batch and to recognize that the importance of these externalities depends upon the similarities between customers.

Our second contribution is to model the value of information and account for these information externalities. We start by providing a general model of the problem, and then present a quantifiable version using an explicit model of uncertainty.

The third contribution is to derive an algorithm, `OLAT`, which solves this batch targeting problem. `OLAT` has several important features. It uses Gaussian processes to estimate our model of the value of information with information externalities. The algorithm approximates the batch targeting problem as a one-step look ahead problem. It overcomes a combinatorial problem of jointly optimizing across a large batch of customers, by creating clusters of neighboring customers, and focusing on externalities within a cluster. The algorithm avoids potential selection issues that can arise when assignments are not random. Moreover, rather than just minimizing estimation errors, the algorithm optimizes expected cumulative profit across all customers. We use data from a direct mail experiment to validate the algorithm.

Organization of the Paper

The paper proceeds as follows. Section 1.2 reviews the literature. Section 1.3 formally describes the batch targeting problem, and proposes a model of information when there are information externalities. Section 1.4 introduces a model of uncertainty, and discusses selection issues. Section 1.5 presents the proposed algorithm. Section 1.6 provides empirical evidence to validate the algorithm, using data from a field experiment.

1.2 Literature Review

We contrast the batch targeting problem that we study with related problems in marketing and computer science, including off-policy batch learning, multi-armed contextual bandits, parallel bandits, and Bayesian optimization (BO). We summarize these comparisons in Table 1.1, and begin the discussion by distinguishing the batch targeting problem from off-policy batch learning.

We use the label "off-policy batch learning" to describe problems that learn from a *single* batch of data. Their goal is to use this training sample to construct an off-policy (a policy that is different from the policy used to generate the sample) for an implementation sample. There is no further sample collection, and so off-policy *only exploits* the implementation sample and does not explore. A typical example of off-policy batch learning in marketing is to learn a personalization policy (the off-policy) from a single customer experiment or observational data. Recent examples include coupon optimization [6], personalization of mobile advertisements [7], personalizing promotions for prospective customers [8], [9], and personalizing prices [10]. Although these problems share the "large batch" feature with batch targeting problems, in off-policy batch learning the firm does not collect new information and so there is no exploration-exploitation trade-off.[6]

There is a branch of the online learning literature that uses batch data, where this logged data was generated from an unobserved sequential process [6], [11], [12]. Although some of these papers learn contextual policies with online learning algorithms [12], or even learn a dynamic policy [6], [11], their setting considers only a single batch of data, and do not allow for new information (from collecting new samples) after the policy is learned from the current batch. In contrast, our problem involves multiple batches of data, including new information that arrives after learning an initial policy.

[6] Using the timeline that we introduce at the start of Section 1.3, the planning horizon only includes two periods, corresponding to "past year" and "this year". There is no "next year", so the firm does not need to explore for future policy improvement.

The second learning problem that we consider is the classic multi-armed bandit problem, which we discussed in the Introduction. We focus on contextual bandits as they allow for learning a personalized policy [1].[7] The arrival process effectively results in a problem with a single customer each period, and little time between arrivals. In contrast, in the batch targeting problem, there is a large batch of customers per period and the time intervals are long. Examples of common contextual bandit problems include digital advertising or search advertising. The exploration-exploitation tradeoff for this problem is generally solved using multi-armed bandit approaches, which include a broad range of methods. These include the Gittins index [14], the Upper Confidence Bound algorithm (UCB) [15], [16], and Thompson sampling [17]. Some of these methods use Gaussian processes to model uncertainty [4], [15]. These approaches have demonstrated good performance on many online learning tasks. Examples within marketing include [18], which considers a firm that has multiple versions of an advertisement and wants to decide which versions to use. [19], who also build on Gaussian processes like the current paper, study demand estimation with bandit experiments. Both proposed algorithms use variations of Thompson sampling.[8]

Within the multi-armed bandit literature, Bayesian Optimization uses a Bayesian approach to directly quantify uncertainty and information values. Unlike the heuristic approach, it is optimal (to its specific setting) by construction and does not provide regret bounds. This class of methods is generally referred to as either knowledge gradient (KG) [2], [4], [5] or expected value of information (EVI) [3]. Recent research has combined contextual bandits with KG and EVI [24], [25]. These methods are approximately equivalent to the IE policy that we introduce in Section 1.3 and use as a benchmark in Section 1.6. Like this paper, KG and EVI also adopt a one-step look ahead approximation in the formulation of their acquisition functions. Rare exceptions have extended to two-step look ahead approximations [26].

In the marketing literature, there are also examples of multi-armed bandit papers that

[7] Contextual bandits are sometimes labelled “active learning” or “adaptive learning” in marketing [13].
[8] There are other examples of multi-armed bandit models in marketing [20]–[23].

adopt the one-step (or two-step) look ahead approximation and propose heuristics to analyze customer dynamics [21], [27]. Most of these papers find that myopic solutions usually perform as well as complete forward-looking solutions [21].[9]

Recall from our earlier discussion that multi-armed bandit models prioritize inter-temporal externalities over within-period information externalities between customers. In [19], the focus is on learning the response to a single price in each period, and so there are no information externalities within a period. In the problem studied by [18], there are multiple customers within a period (within a batch), and so information externalities do arise, though their approach does not account for these externalities.

The [18] problem is an unusual bandit problem as there are multiple customers each period, and the decision-maker personalizes a policy for each customer. Most bandit problems have a single customer each period. A small class of computer science papers study parallel bandit problems, in which the decision maker solves a bandit problem with multiple observations each period. They have either proposed heuristics [33]–[36] or models [37]–[39] to encourage diversity in which arms are pulled within a batch. For the heuristic approach, the goal is to minimize cumulative regret, which usually works better in a highly sequential setting. Similar to [37], [38] and [39], our proposed method optimizes the information value exactly, and does not try to minimize cumulative regret.

[9] In our problem, the firm looks forward, but customers do not. This is somewhat standard approach in marketing, where applications include designing conjoint experiments [28], adapting website design [20], survey design [13], ad sequencing [29], eliciting consumer risk preferences [30], modelling consumer experiential learning [21], pricing experiment [19], and optimizing catalog mailing [31]. However, there have been studies that explicitly consider forward-looking consumers. A notable early example is [32], who study direct mail targeting.

Table 1.1: Comparison of Learning Problems

	More Information Collection	**Customers in a Period**	**Type of Policy**	**Examples**
Batch targeting problem	Yes	Many	Personalized	This paper
Off-policy batch learning	No	Many	Personalized	[7], [8], [10], [12]
Contextual multi-armed bandits (including BO)	Yes	One	Personalized	[13], [18], [19], [21]
Parallel bandits (including BO)	Yes	Many	Uniform	[33], [34], [36], [39]

In general, the research on parallel bandits in computer science are not well-suited to the batch targeting problem. The methods they propose are generally designed for settings where only a handful of arms can be pulled each period, and multiple batches occur close together in time. They derive uniform policies, rather than personalizing marketing actions to different customers within a batch.[10] Moreover, even though there are examples of recent methods that consider linear contextual parallel bandit problems [35], [42], they do not model how the customers' locations in covariate space contribute to information externalities. If customers are close neighbors (in covariate space), then the information we learn from them when exploring is likely to be more duplicative than customers located far apart.

The focus of our problem is learning in batches, and we focus on maximizing firm profit. Other research studying sequential learning with batches has focused on online advertising [18], conjoint analysis [43], and policy selection [40], [41]. Notably, all of these papers, except [18], focus on accurately estimating treatment effects rather than maximizing profit. While accurate estimation of treatment effects can contribute to higher firm profits, the two things are not the same. If we consider two policies, it is possible that the policy that yields more

[10] See for example [39]–[41].

accurate estimates of treatment effects will not be the policy that maximizes expected profits [44]. We also note there is a branch of literature that focuses on determining the size of each experiment batch [45], [46]. This question is outside the scope of our paper; we assume the batch sizes are exogenously given.

In the next section we formally define the batch targeting problem, and decompose the learning component into two elements; the expected cost of information, and the expected value of information.

1.3 The Batch Targeting Problem

Before presenting a formal description of the problem, we first provide a conceptual introduction that describes the timing of the firm's actions.

Suppose a retailer conducts a marketing campaign annually. Each period the retailer faces a new batch of customers, and the firm decides which marketing action to assign to each of them. This setting fits into many campaign settings, such as holiday marketing campaigns, nondigital marketing campaigns, and many political campaigns (get-out-to-vote). It also aligns with the business practice when firms delay processing data and learn from (delayed) batches of customer responses. In particular, the distinction feature of this batch targeting problem is that the firm faces batches of customer data when it uses the data to learn.

We label the focal period "Period 1." Under the one-step look ahead assumption (discussed in the Introduction), the batch targeting problem can be reduced to a three-period problem. In particular, if the firm only looks one period ahead, the firm's decision in the focal period requires observing or anticipating information from three time periods:

Period 0 (*past periods*): when making assignment decisions in Period 1, the firm observes the actions assigned to Period 0 customers and the outcomes from Period 0 customers.

Period 1 (*focal period*): when making assignment decisions in Period 1, the firm uses data from Period 0 (could be all past periods). The firm implements the assigned actions on Period 1 customers, and anticipates their outcomes at the end of Period 1.

Period 2: when making assignment decisions in Period 1, the firm anticipates that in Period 2 it will use data combined from Periods 0 and 1 to train a policy that only exploits profits from Period 2 customers, and will implement this policy on all Period 2 customers.

It is important to keep in mind that the firm only makes assignment decisions in (Period 1). In practice, this means that the firm solves a batch targeting problem separately at each period. When doing so, the firm observes decisions and outcomes from the previous period, and anticipates actions and outcomes that will occur next period.

When making decisions in Period 1, the Period 0 assignment decisions are sunk, and cannot be changed. However, the Period 0 actions and outcomes serve as an important input to the Period 1 decisions. Note that Period 0 can more generally be interpreted as all periods prior to Period 1. We also emphasize that the policy assigned to Period 0 customers need not be randomized, nor does it need to be optimized. The only requirement is that we know the Period 0 policy assignment, and can observe any data used to train that policy.

Consistent with the one-step look ahead assumption, when making Period 1 assignment decisions, the firm believes that Period 2 is the terminal period, and so it anticipates it will use all the available information (from Periods 0 and 1) to fully exploit in Period 2. The firm anticipates how the information contributed by its Period 1 decisions are expected to change the performance of Period 2's policy. It is this information that is susceptible to information externalities. The incremental improvement in Period 2's policy contributed by observing a single Period 1 customer, depends upon the actions assigned to all customers in Period 1, together with the location of customers in Periods 1 and 2.

1.3.1 Problem Description

We now introduce a model to characterize this retailer's problem. Our notation convention uses upper case for variables (e.g. $A_{i,t}$), lower case for data values (e.g. n_t), italics to denote both singulars and functions (e.g. $Y_{i,t}$), bold roman to denote vectors and matrices (e.g. $\mathbf{X}_{i,t}$), and script to identify sets (e.g. $\mathcal{N}_t$). We summarize the notation used in the main text at the end of this section in Table 2 (additional notation used in the Appendix A are summarized in Table A.1).

There are three periods in the model, denoted as $t \in \{t_0, t_1, t_2\}$. We will refer to the periods as *past period* (t_0), *this period* (t_1), and *next period* (t_2). In each period t, there is a batch of n_t customers, and this batch is denoted by $\mathcal{N}_t$. For each customer $i \in \mathcal{N}_t$, the firm observes characteristics (such as age), contained in a vector of targeting covariates, $\mathbf{X}_{i,t} \in \mathcal{X} \subseteq \mathbb{R}^m$. The retailer's action space, denoted by $\mathcal{A}$, is assumed to be a finite set, and the same across all periods. The action assigned to customer i in period t is $A_{i,t} \in \mathcal{A}$.

The single period outcome for customer i is the realized individual profit: $Y_{i,t} \in \mathcal{Y} \in \mathbb{R}$. The cost of implementing the action is included in $Y_{i,t}$, and the firm only observes customer i's outcome after implementing actions to all customers in that period's batch.

We assume that the outcome $Y_{i,t}$ is conditioned on both the observed covariates and the firm's action, and can be characterized by a stochastic response function r, such that $r : \mathcal{X} \times \mathcal{A} \mapsto \mathcal{Y}$. For the purposes of this section, we can maintain r in general form, but will introduce a specific functional form for r at the start of the next section.

We use history $\mathbf{H}_t$ to denote the information available at the start of period t:

$$\mathbf{H}_{t_1} = \{\mathbf{X}_{t_0}, \mathbf{X}_{t_1}, \mathbf{X}_{t_2}, \mathbf{A}_{t_0}, \mathbf{Y}_{t_0}\}$$
$$\mathbf{H}_{t_2} = \{\mathbf{X}_{t_0}, \mathbf{X}_{t_1}, \mathbf{X}_{t_2}, \mathbf{A}_{t_0}, \mathbf{A}_{t_1}, \mathbf{Y}_{t_0}, \mathbf{Y}_{t_1}\}$$

We make three clarifications of the formal model. First, we assume that customer response functions are stationary across all periods, and a customer's response does not depend on the past actions they received. The marketing literature commonly considers two types of dynamics. Customer responses to the same firm action may vary over time. Alternatively, the firm may face a learning problem, so that changes in the firm's information change the optimal firm action over time. In this paper, we do not consider any dynamics on the customer side. Instead, we focus solely on dynamics introduced by the firm's learning problem.

Second, we assume that the firm could observe all customers' covariates $(\mathbf{X}_{t_0}, \mathbf{X}_{t_1}, \mathbf{X}_{t_2})$ in t_0, which aligns with the practice that all customer data is available in the firm's CRM system in t_0. This assumption allows us to improve the uncertainty estimation with a technique called transduction introduced later in Section 1.4. Note that our model can be easily generalized to the case where we only know the empirical distributions of future targeting covariates. We could simply replace expectation operations over these covariates with the empirical distributions.

Third, we do not consider the endogenous formulation of customer batches and treat them as exogenously given. This aligns with the practice of holiday marketing campaigns or nondigital marketing campaigns; firms need to make assignment decisions (not sending a promotion is also considered an assignment) for most of their customers.

Policies and Objective Function

There are two types of policy in this problem. The first type of policy is denoted by p_t. This is the *(myopic) optimal* policy if the firm is solely focused on exploiting period t's customers (it is trained without regard to exploration). This myopic "exploit policy" is optimal with respect to history $\mathbf{H}_t$ and thus subscripted by t. Notice that the history changes over time, and so the exploit policy this period (p_{t_1}) will generally not be the exploit policy next period (p_{t_2}). Formally, the exploit policy is a mapping from the covariate space to the action space, with information in $\mathbf{H}_t, p_t : \mathcal{X} \mapsto \mathcal{A}$:

$$p_t(\mathbf{x}) \in \operatorname*{argmax}_{a \in \mathcal{A}} \mathbb{E}[r(\mathbf{x}, a) | \mathbf{H}_t], \forall \mathbf{X} \in \mathcal{X}.$$

In practice the firm can use any supervised learning model to train the exploit policy (the choice of this training process and the off-policy evaluation problem is outside the scope of this paper).[11] The exploit policies serve two roles in the model. The first role is in the current period (Period t_1), where p_{t_1} is used to measure the cost of exploring. In particular, we measure this cost as the expected profit when assigning the action recommended by p_{t_1} compared to the expected profit when exploring by deviating from p_{t_1}. The second role is anticipating next period's actions when making assignment decisions this period; the firm anticipates that in t_2 it will implement p_{t_2} and focus solely on exploiting t_2 customers.

The second type of policy is the actual assignment policy π in Period t_1. Unlike p_{t_1}, when choosing this policy, the firm balances exploration and exploitation by considering how this period's assignments will improve next period's policy. It is the actual assignment rule that the firm uses to assign marketing actions to every customer in Period t_1 and is denoted by $\pi : \mathcal{N}_{t_1} \mapsto \mathcal{A}$.[12] When designing π the firm maximizes the cumulative expected profit from this period and next period (Periods t_1 and t_2):

$$V_{t_1}(\pi) = \sum_{i \in \mathcal{N}_{t_1}} \mathbb{E}^{\pi} [\![Y_{i,t_1} \mid \mathbf{H}_{t_1}, \mathbf{A}_{t_1}]\!] + \sum_{i \in \mathcal{N}_{t_2}} \mathbb{E}^{\pi} \left[\!\!\left[\mathbb{E}_{t_2}[Y_{j,t_2}; p_{t_2} \mid \mathbf{H}_{t_2}] \,\middle|\, \mathbf{H}_{t_1}, \mathbf{A}_{t_1} \right]\!\!\right] \tag{1.1}$$

This is the firms objective function for the batch targeting problem. The first term describes the expected profit this period (Period t_1), while the second term describes the expected profit next period (Period t_2).

The inner expectation in the second term, $\mathbb{E}_{t_2}[Y_j; p_{t_2} | \mathbf{H}_{t_2}]$, assumes that we know next period's history $\mathbf{H}_{t_2}$ (which includes this period's actions $\mathbf{A}_{t_1}$ and outcomes $\mathbf{Y}_{t_1}$). If we know

[11] See for example the methods described in [8] and [47].

[12] Notice that π is not subscripted by t because this type of policy only arises in Period t_1.

HH_{t_2}, p_{t_2} is also determined, and so we can form expectations about next period's outcomes (Y_j). The uncertainty in the inner expectations is solely due to stochasticity in next period's response function (r).

The outer expectations of both terms, $\mathbb{E}^{\pi}[\![\cdot|\mathbf{H}_{t_1}, \mathbf{A}_{t_1}]\!]$, recognize that the outcomes for this period's customers are unknown. Even if we know this period's history ($\mathbf{H}_{t_1}$) and this period's actions ($\mathbf{A}_{t_1}$), we do not know this period's outcomes. The uncertainty in the outer expectations is solely due to stochasticity in this period's underlying response function.

While the objective function in Equation 1.1 describes the firm's objective, it does not illustrate the tradeoffs that the firm faces in optimizing this objective. We next introduce a model of these tradeoffs, and show that an assignment policy that optimizes these tradeoffs also optimizes the firm's objective in Equation 1.1.

1.3.2 The Cost and Information Value of Deviating This Period

A decision to deviate from the exploit policy this period reflects a tradeoff between two quantities. The *opportunity cost of information* is the expected opportunity cost of deviating from the exploit policy this period. The *expected value of information* is the additional profit expected next period, due to the additional information learned from this deviation. We start by discussing the cost and value of information for an individual customer.

Consider this period's (t_1) problem. For a customer i, an expected opportunity cost arises if the firm deviates from the exploit policy p_{t_1}. The opportunity cost of deviating from p_{t_1} by assigning action A_i to customer $i \in \mathcal{N}_{t_1}$ is equal to:

$$IC_{t_1}(\mathbf{X}_i, A_i) \equiv \mathbb{E}\left[r\left(\mathbf{X}_i, p_{t_1}(\mathbf{X}_i)\right) - r(\mathbf{X}_i, A_i) \,\middle|\, \mathbf{H}_{t_1}\right] \geq 0. \tag{1.2}$$

The "IC" label for this function stands for "information cost" function. The expectations are over the outcomes in Period t_1 (due to stochasticity in the response function r). Because p_{t_1} is the policy that fully exploits in t_1, we know that the information cost is weakly positive.

If $A_i = p_{t_1}(\mathbf{X}_i)$, then the assignment for customer i is the same action assigned by p_{t_1}, and the IC-function equals zero for this customer.

We next consider the expected value of information from customer i. Exploring with this customer in t_1 can improve the exploit policy in t_2. This introduces an inter-temporal externality, which is widely recognized as the exploration exploitation tradeoff. However, in the batch targeting problem, it also introduces an information externality across t_1 customers.

The "information value" (IV) function measures the expected incremental t_2 profit if customer i deviates from p_{t_1} in t_1:

$$
\begin{aligned}
IV_{t_1}(\mathbf{X}_i, A_i|\mathbf{A}_{-i}) \equiv & \sum_{j\in\mathcal{N}_{t_2}} \mathbb{E}\left[\mathbb{E}_{t_2}\left[r(\mathbf{X}_j, p_{t_2}(\mathbf{X}_j)) \,\middle|\, \mathbf{H}_{t_2}(A_i;\mathbf{A}_{-i})\right] \,\middle|\, \mathbf{H}_{t_1}, A_i, \mathbf{A}_{-i}\right] \\
& - \sum_{j\in\mathcal{N}_{t_2}} \mathbb{E}\left[\mathbb{E}_{t_2}\left[r(\mathbf{X}_j, p_{t_2}(\mathbf{X}_j)) \,\middle|\, \mathbf{H}_{t_2}(p_{t_1}(\mathbf{X}_i);\mathbf{A}_{-i})\right] \,\middle|\, \mathbf{H}_{t_1}, p_{t_1}(\mathbf{X}_i), \mathbf{A}_{-i}\right]
\end{aligned} \tag{1.3}
$$

The two terms both measure the expected profit in t_2. In the first term, the expected t_2 profits are conditioned on the firm assigning A_i to customer i in t_1, while in the second term, they are conditioned on the firm assigning the action recommended by $p_{t_1}(\mathbf{X}_i)$ to customer i in t_1. In both terms, the actions assigned to the remaining t_1 customers are denoted by $\mathbf{A}_{-i}$. [13]

Just like Equation 1.1, the outer expectations are over the uncertainty in t_1 responses, while the inner expectations are over the uncertainty in t_2 responses. We will explain in the next section how we will evaluate these expectations.

The IV function is positive when p_{t_1} (the exploit policy in t_2) has a higher expected profit if the firm deviates from p_{t_1} and assigns A_i to customer i in t_1. However, the IV function can also be negative, because the expected profit from p_{t_2} maybe lower when deviating from p_{t_1} in t_1.

As with the IC-function, this function focuses on customer i. It represents the value of

[13] We use the notation $\mathbf{H}_{t_2}(A_i;\mathbf{A}_{-i})$ to describe the history at the start of t_2, after actions $(A_i;\mathbf{A}_{-i})$ in t_1. Analogously, $\mathbf{H}_{t_2}(p_{t_1}(\mathbf{X}_i);\mathbf{A}_{-i})$ is the history at the start of t_2, after actions $(p_{t_1}(\mathbf{X}_i);\mathbf{A}_{-i})$ in t_1.

the information obtained by varying this period's action for customer i. It does not measure the aggregate information from varying the actions for other customers this period. However, unlike the IC-function, which is completely separable (and independent) between individual t_1 customers, the IV-function is not separable. In particular, p_{t_2} depends upon not just the action assigned to customeri this period, it also depends upon the actions assigned to other customers in t1 (denoted by $\mathbf{A}_{-i}$). This is the *information externality* that we discussed in the Introduction.

The IV function measures the value of deviating from p_{t_1} in Period t_1 (with customer i), while the IC function measures the cost of these deviations. To balance the trade-off between exploring and exploiting, we calculate the difference between the IC and IV-functions:

$$EE_{t_1}(\mathbf{X}_i, A_i|\mathbf{A}_{-i}) \equiv IV_{t_1}(\mathbf{X}_i, A_i|\mathbf{A}_{-i}) - IC_{t_1}(\mathbf{X}_i, A_i). \tag{1.4}$$

We label this the EE-function, representing the "Explore-Exploit" function. We use the EE-function as the objective function in Period t_1:

$$\pi^* \in \underset{A_i \in \mathcal{A}}{\operatorname{argmax}}\, EE_{t_1}(\mathbf{X}_i, A_i|\mathbf{A}_{-i}; \mathbf{A}_{-i} \in \pi^*),\ \forall i \in \mathcal{N}_{t_1}. \tag{1.5}$$

The IV-function and the IC-function are both individual-level functions (for each customer in t_1), and so the EE-function is also an individual-level function. Moreover, because the IV-function is not separable between this period's customers, the EE-function is also not separable.

Our first result recognizes that *jointly optimizing* the EE-function across all t_1 customers will also maximize the firm's total expected profits.

Result 1 (Value function maximization) *Any solution that jointly optimizes the EE-function (Equation 1.5) also optimizes the firm's total expected profits (Equation 1.1).*

Proof. See Appendix A.2.

Intuitively, when optimizing the EE-function, it is only profitable to deviate from p_{t_1} if the value of the information obtained (considering information externalities) outweighs the cost of that information. Thus, the EE-function explicitly measures the trade-off between new information (exploration) and existing knowledge (exploitation). By jointly maximizing the EE-function across this period's customers, we identify the combination of deviations from p_{t_1} that jointly maximize the total expected profits in t_1 and t_2. This combination of deviations is this period's optimal assignment policy π^*.

Result 1 confirms that the EE-equation (in Equation 1.5) describes the firm's objective function in t_1. Solving the EE-function solves the batch targeting problem each period (recall that t_1 identifies the focal period).

This result also speaks to the combination nature of handling the information externalities in the batch targeting problem. Customers in a batch and across different batches are heterogeneous, and so different combinations of customer-action pairs lead to different total profits. The information externality between customers is larger when these customers are more similar. This means that only finding the right "amount" of exploration is insufficient; we also need to find the right "identities" of customers to explore. For example, when a customer appears to be an interpolation between two other customers, this customer could generate higher information value. Exploring a new action with this customer will be more likely to also suggest the responses of the other customers receiving the same action.

Discussion: One-Step Look Ahead Approximation

Reader might wonder how we extend our model to allow for more depth forward looking than just one period. Practically, it is intractable and computationally prohibitive to directly extend to some variant of Bellman-formulation that is perfectly forward-looking; this is even the case for standard Bayesian optimization algorithms that does not deal with the combinatorial problem introduced by information externalities. However, using a "reduced form" way to go deeper into the future, we can add a factor on the IV function that discounts the expected value of information, if we anticipate more exploration (than pure exploitation) incurring higher opportunity cost in next period t_2. On the other side, if we hope to multiply the importance of future earnings when anticipating more future periods, we can also set this factor to be greater than one.

This same approach can also be used to handle suspected customer nonstationarity, which is a special type of forward looking. If we anticipate customer response function will change a bit in the future, we can also discount the estimated value of information by setting this IV-function factor to be smaller than one.

In the remainder of this section, we compare the EE-function with a benchmark function that ignores information externalities between t1 customers.

1.3.3 Benchmark: Individual Optimization

As a benchmark, we consider a problem in which the firm ignores information externalities in the IV-function. In this benchmark, the IV-function (information value) treats each t_1 customer separably. The IC-function (information cost), is already separable and so does not change under this benchmark.

However, under this benchmark the IV-function is adjusted as if there was only one customer in the t_1 batch. In particular, the information value for each customer in t_1 as if the focal customer was the only customer in t_1:

$$\begin{aligned} i\text{-}IV_{t_1}(\mathbf{X}_i, A_i) \equiv & \sum_{j \in \mathcal{N}_{t_2}} \mathbb{E}\left[\mathbb{E}_{t_2}\left[r(\mathbf{X}_j, p_{t_2}(\mathbf{X}_j)) \,\middle|\, \mathbf{H}_{t_2}(A_i)\right] \,\middle|\, \mathbf{H}_{t_1}, A_i\right] \\ & - \sum_{j \in \mathcal{N}_{t_2}} \mathbb{E}\left[\mathbb{E}_{t_2}\left[r(\mathbf{X}_j, p_{t_2}(\mathbf{X}_j)) \,\middle|\, \mathbf{H}_{t_2}(p_{t_1}(\mathbf{X}_i))\right] \,\middle|\, \mathbf{H}_{t_1}, p_{t_1}(\mathbf{X}_i)\right] \end{aligned} \tag{1.6}$$

We label this equation i-IV for "individual Information Value" and to distinguish it from the IV-function (Equation 1.3). Notice that unlike Equation 1.3, in Equation 1.6 the inner and outer expectations are not conditioned on the actions of other customers in t_1.

We optimize the difference between the i-IV-function and the IC-function for a focal customer i by:

$$A_i^I \in \underset{A_i \in \mathcal{A}}{\operatorname{argmax}}\, IE_{t_1}(\mathbf{X}_i, A_i) \equiv i\text{-}IV_{t_1}(\mathbf{X}_i, A_i) - IC_{t_1}(\mathbf{X}_i, A_i),\ \forall i \in \mathcal{N}_{t_1}. \tag{1.7}$$

We label this function the IE function for "Individual Exploration Exploitation" function. Equation 1.7 is separable for each t1 customer. In contrast, the EE-equation in Equation 1.5 is not separable, and instead requires finding a "fixed point," in which the assignment for each customer is optimal given the optimality of assignments for other customers this period. The difference between those two approaches is that the information externality is included in Equation 1.5, but not in Equation 1.7.

Without specifying an explicit process for training the period t_2 exploit policy (p_{t_2}), it is difficult to explicitly characterize the relationship between the IE and EE functions. For example, we would generally expect that information externalities lead to at least some duplication of information, so that the IE-function overstates the incremental expected t_2 profit contributed by i. However, this will depend upon the nature of the information externalities, the actions chosen for other customers, and the process used to train the t_2 exploit policy.

We do know that when information externalities are present, the EE-function correctly considers them, while the IE-function incorrectly ignores them. This means that if the optimal solution to the EE-function is different than the optimal solution to the IE-function, then the optimal EE-function policy (π^*) will yield a higher expected profit than the solution to the IE-function. By excluding information externalities when evaluating the value of information (Equation 1.6), the IE-function is subject to a distortion that can result in a sub-optimal solution. We will empirically compare the EE-solutions and IE-solutions in Section 1.6, where we solve both problems using data from a field experiment.

1.3.4 Summary

In this section we introduced the EE-function to reconcile the exploration-exploitation trade-off by measuring the expected (information) value of deviating from the current exploit policy. The EE-function quantifies the expected value of information together with the expected cost of this information. The expectations reflect uncertainty about the way customers will respond to the firm's actions. We introduce a model of this uncertainty in the next section, and show how it can be used to compute the EE-function.

Table 1.2: Notations Introduced in Section 1.3

t	Subscript identifying time: past period (t_0), this period (t_1), next period (t_2)
i	Subscript identifying customers
$\mathcal{N}_i$	The set of customers in period t
n_t	The number of customers in period t
$A_{i,t}$	Action implemented for customer i in period t
$\mathbf{A}_t$	Vector of actions implemented for customers in period t
$\mathbf{X}_i, \mathbf{X}_t$	Vector of covariates for customer i; covariates for customers in period t
$\mathbf{x}$	A realized value of covariates
$\mathbf{A}_{-i}, \mathbf{X}_{-i}$	The actions or covariates for all customers except customer i in a batch
Y_i	The single period profit earned (outcome) from customer i
$\mathcal{A}, \mathcal{X}, \mathcal{Y}$	The value spaces of actions, covariates and outcomes
$\mathbf{H}_t$	History of data observed at the start of period t
$r(\mathbf{x}, a)$	The response function with covariates $\mathbf{x}$ and action a
π	The assignment policy (of this period)
p_t	The exploit policy given $\mathbf{H}_t$
IC_t	Information cost function given $\mathbf{H}_t$
IV_t	Information value function given $\mathbf{H}_t$
EE_t	Explore-exploit function given $\mathbf{H}_t$
$i\text{-}IV_t$	Individual information value function given $\mathbf{H}_t$(no information externalities)
IE_i	Individual explore-exploit function given $\mathbf{H}_t$ (no information externalities)

We use upper case for variables, lower case for data values, italics to denote both singulars and functions, bold roman to denote vectors and matrices, and script to identify sets.

1.4 Uncertainty and Unobservables

1.4.1 Uncertainty Model: Gaussian Process

We use a nonparametric Bayesian approach to model uncertainty in the targeting response function. Specifically, for customer i in period t, we assume that the realized profit $Y_{i,t}$ is determined by the following equation:

$$Y_{i,t} = r(\mathbf{X}_{i,t}, A_{i,t}) + \epsilon_{i,t},\ \epsilon_{i,t} \sim_{\text{ind.}} \mathbb{N}(0, \tau^2). \tag{1.8}$$

In this expression, $r(\cdot,\cdot)$ is the response function and is stationary across periods. $\epsilon_{i,t}$ is a zero-mean unobservable term, which is normally and independently distributed across customers and periods. This noise term recognizes the unobserved information that is not captured by the individual covariates and actions. The existence of this unobservable term may raise endogeneity concerns, which we will address later in this section.

We model the response function r with *Gaussian processes* (GPs). The GP approach offers many benefits [48]. Most importantly, GP is a nonparametric approach and takes a function space view, and directly imposes a prior distribution on the function r. This quantifies uncertainty at each covariate value with a posterior distribution. In comparison, a standard parametric model only estimates one standard error for a parameter, and fails to measure the uncertainty of customers with different covariate values. However, in order to estimate our decision model, we need to pointwise estimate the standard error of the predicted profit in response to any covariate values. This is straightforward with our GP-based approach, but it requires techniques like a large number of bootstraps for a parametric model to work. If you consider the computation in the joint action assignment problem, the computational cost of parametric models is high.

In addition, GP also allows for easy quantification of uncertainty, which will be critical for evaluating the EE-function. It has a conjugate prior and parsimonious representation, which generates reasonable computational efficiency.

More generally, there are many benefits of taking a Bayesian perspective. First, the Bayes decision function is admissible and constitutes a complete class [49], which means that the Bayesian framework provides an attractive theoretical guarantee for decision making. Second, Bayesian inference has an embedded regularization in the likelihood, which helps to avoid overfitting. Finally, in Subsection 1.4.2, we will also show that Bayesian inference helps to address a selection problem.

Formally, we impose a Gaussian process prior directly on the response function r to model its distribution, allowing for covariance between outcomes associated with different actions:

$$r \sim \mathcal{GP}(\mu, k).$$

Following the GP convention, a GP prior is specified by (μ, k), where μ denotes the mean function, and k is the kernel (covariance) function that controls the curvature of this GP. Formally, for any two sets of inputs $(\mathbf{X}, A), (\mathbf{X}', A')$, the mean function (taken to be zero, following the GP convention) and kernel function are defined as:

$$\begin{aligned}
&\mu(\mathbf{X}, A) \equiv \mathbb{E}[r(\mathbf{X}, A)] = 0 \\
&k\left((\mathbf{X}, A), (\mathbf{X}', A')\right) \equiv \operatorname{cov}\left(r(\mathbf{X}, A), r(\mathbf{X}', A')\right), \ \forall \mathbf{X}, A, \mathbf{X}', A' \in \mathcal{X} \times \mathcal{A}.
\end{aligned}$$

Conditional on having observed some history, we can directly use the GP model to characterize a posterior distribution of the profit function. With this distribution, we know the mean value and the level of uncertainty (pointwise) for any inputs. Since a Gaussian process evaluated at any point is a Gaussian distribution, the posterior distribution of the profit function evaluated at a focal covariate and action pair also follows a Gaussian distribution. Both the mean and the variance of a known Gaussian distribution have closed-form expressions and are easy to compute. Therefore, by modelling the response function r using GP, we now can easily quantify the uncertainty at any point on the function.

In our application, we update the GP posterior as follows. We start from the history this period (t_1), which consists of two parts: inputs and outcomes. We represent the inputs (targeting covariates and assigned marketing actions) as $(\mathbf{X}_{t_0}, \mathbf{A}_{t_0})$, and the outcomes (individual profits) as $\mathbf{Y}_{t_0}$. Suppose we have a new input value $(\mathbf{X}_i, A_i)$, with the model $Y_i = r(\mathbf{X}_i, A_i) + \epsilon_i$. For compact representation, we define some useful covariance matrices, evaluated at specific inputs, as

$$
\begin{aligned}
& K_{ii} \equiv k\left((X_i, A_i), (X_i, A_i)\right), \; K_{i0} \equiv k\left((X_i, A_i), (\mathbf{X}_{t_0}, \mathbf{A}_{t_0})\right), \\
& \mathbf{K}_{00} \equiv k\left((\mathbf{X}_{t_0}, \mathbf{A}_{t_0}), (\mathbf{X}_{t_0}, \mathbf{A}_{t_0})\right).
\end{aligned}
$$

The predictive posterior distribution of the new outcome Y_i, corresponding to this new input $(\mathbf{X}_i, A_i)$, is denoted by $P\left(Y_i \,|\, (\mathbf{X}_i, A_i), (\mathbf{X}_{t_0}, \mathbf{A}_{t_0}), \mathbf{Y}_{t_0}\right)$, and expressed as

$$
Y_i|(\mathbf{X}_i, A_i), (\mathbf{X}_{t_0}, \mathbf{A}_{t_0}), Y_{t_0} \sim \mathbb{N}(\mu_i, \Sigma_i). \tag{1.9}
$$

The mean and variance of this predictive distribution can be analytically given by,

$$
\begin{aligned}
& \mu_i \equiv \mathbf{K}_{i0}(K_{00} + \tau^2 \mathbf{I})^{-1} \mathbf{Y}_{t_0} \\
& \Sigma_i \equiv K_{ii} - \mathbf{K}_{i0}(\mathbf{K}_{00} + \tau^2 \mathbf{I})^{-1} \mathbf{K}_{i0}^{\top}.
\end{aligned}
$$

A direct observation is that the posterior variance only relies on the inputs but not the outcomes. This feature is particularly helpful for t_1's assignment decisions. If we know the covariates of t_1's customers (X_{t_1}), then we can estimate the uncertainty about next period before observing t_1's outcomes.

This predictive posterior distribution can be used to construct a generative model of individual responses $R(Y)$, which we use to simulate outcome samples. Formally,

$$
R(Y|\mathbf{X}, A) \sim P\left(Y_i \,|\, (\mathbf{X}, \mathbf{A}), (\mathbf{X}_{t_0}, \mathbf{A}_{t_0}), \mathbf{Y}_{t_0}\right), \; \forall x \in \mathcal{X}, a \in \mathcal{A}. \tag{1.10}
$$

This generative model is especially useful when we evaluate the IV-function (Equation 1.3) and i-IV-function (1.6). For example, in Equation 1.6, we need to evaluate $\mathbb{E}_{t_2}[\cdot|\mathbf{H}_{t_2}(A_{i,t_1})]$ in t_1, which is the expectation of the t_2 outcomes conditional on the (unrealized) t_2 history $\mathbf{H}_{t_2}$, when A_{i,t_1} is assigned to i in t_1. Therefore, we need the distribution of $\mathbf{H}_{t_2}$ given $\mathbf{H}_{t_1}$ and

A_{i,t_1}. This is a hard problem because we do not observe Y_{i,t_1}. The generative model solves this problem by modelling $R(Y_{i,t_1}|\mathbf{X}_{i,t_1}, A_{i,t_1})$ according to Equation 1.10. This characterizes the probabilistic distribution of Y_{i,t_1}, which can then be used to construct an empirical distribution of $\mathbf{H}_{t_2}$. We use this empirical distribution to build a simulated estimator of $\mathbb{E}_{t_2}[\cdot|\mathbf{H}_{t_2}(A_{i,t_1})]$, which we will discuss in more detail in Section 1.5 (and Appendix A).

Finally, we add a few remarks on inference. First, to estimate the profit function, we need to use the marginal likelihood of observed outcomes. This marginal likelihood is given by:

$$P(\mathbf{Y}_{t_0}|\mathbf{X}_{t_0}, \mathbf{A}_{t_0}) = \int P\left(\mathbf{Y}_{t_0}|\mathbf{R}, (\mathbf{X}_{t_0}, \mathbf{A}_{t_0})\right) P(\mathbf{R}|\mathbf{X}_{t_0}, \mathbf{A}_{t_0})\, \mathrm{d}\mathbf{R}.$$

In this expression, $\mathbf{R} \equiv r(\mathbf{X}_{t_0}, \mathbf{A}_{t_0})$ represents the (predicted) response function values using training inputs from past period (t_0).[14] Both the likelihood $P\left(\mathbf{Y}_{t_0}|\mathbf{R}, (\mathbf{X}_{t_0}, \mathbf{A}_{t_0})\right)$ and the prior $P(\mathbf{R}|\mathbf{X}_{t_0}, \mathbf{A}_{t_0})$ follow Gaussian distributions, $\mathbb{N}(\mathbf{R}, \tau^2\mathbf{I})$ and $\mathbb{N}(0, \mathbf{K}_{00})$ by construction. Moreover, following the applied GP literature, we use a square exponential (SE) function, as the covariance function. [15] To find the optimal hyperparameters, we follow the convention in Bayesian inference literature, and use empirical Bayes to optimize the likelihood function. In the importance sampling procedure for computing marginal likelihood, we use the inverse action assignment probabilities as weights (see for example, [50]).

1.4.2 Selection on Observables

To ensure identification of the outcome function $r(\mathbf{X}, A)$, we need to satisfy the *selection on observables* condition.[16] For example, if the firm is choosing between two actions, `mail` or

[14] It is also possible to generalize our current profit function inference procedure to allow for doubly robust inferences. Doing so requires parametrized variables in the distributions to denote the predicted responses. However, our GP framework is nonparametric. As an alternative, we could use the predicted response values evaluated at inputs for parametrization.

[15] For an input $(\mathbf{X}, A)$, the SE function is given by $k((\mathbf{X}, A), (\mathbf{X}', \mathbf{A}')) = \exp\left(-\frac{||(\mathbf{X},A)-(\mathbf{X}',A')||^2}{2l^2}\right)$.

[16] We also need to satisfy the "Overlapping" condition, which requires that a customer of any covariate type has strictly positive probability of receiving any action assignment. We can easily verify that this assumption is satisfied by recalling the intuition of our assignment policy design. It is designed to balance exploration and exploitation. That is, for any covariate that we are uncertain about the outcome associated with some action, there is also a positive probability of assigning that action to it.

`not mail`, we need the potential outcome for `mail` and the potential outcome for `not mail` to be independent of the actual assignment (after conditioning on the covariates $\mathbf{x}$). This needs to hold for all customers each period, and we will discuss each separately.

If past period's assignment decisions ($\mathbf{A}_{t_0}$) were randomized (as in our empirical application), then the assumption is satisfied (for that period). However, it is also satisfied if t_0's assignments were made based upon observed covariates. In contrast, if the assignments were based on an unobserved covariate, then the variation in the observed outcome may be caused by variation in the unobserved covariate(s). This is the reason that in Section 1.3 we stipulated that $\mathbf{A}_{t_0}$ assignments were either randomized, or made using a policy trained on observed covariates ($\mathbf{X}_{t_0}$).

In t_1, the assignments are clearly not randomized (by construction). Instead, the assignment policy is designed based upon the *observed* covariates. Fortunately, the condition allows for selection on observed variables, and only requires independence conditional on covariates included in the model. In our case, the variables used in determining assignments are known to the firm. Therefore, conditional independence is satisfied, because all of the variables that influence the assignment design can be included as covariates when estimating the profit function.

When training a model in t_2, we use data pooled from t_1 and t_0. This may introduce a different (though related) concern. Past period's outcomes ($\mathbf{Y}_{t_0}$) appear in this period's history ($\mathbf{H}_{t_1}$) and next period's history ($\mathbf{H}_{t_2}$), and are an input to this period's assignments (Equation 1.5). However, $\mathbf{Y}_{t_0}$ is a function of observables ($\mathbf{X}_i, A_i$) and unobservables (ϵ_i):

$$Y_i = r(\mathbf{X}_i, A_i) + \epsilon_i.$$

The unobservable features in $\mathbf{Y}_{t_0}$ contribute to t_1 assignments, and could introduce autocorrelation (see for example [51], [52]). Our Bayesian perspective resolves this risk. In Bayesian inference, the learning object is regarded as a *random variable*. Based on the Likelihood Principle [51], [52], inference is based upon the *likelihood* of data conditional on that object. In our case, we estimate this likelihood using a Gaussian process to characterize the outcome r. Since the prior is external and does not contribute to selection, we only need to ensure that there is no selection risk in the likelihood.

In Result 2, we formally prove that our Bayesian approach is not susceptible to selection assuming the batch targeting campaigns last for T periods.

Result 2 (Free from Selection) *When learning the response function (r) in an adaptive batch targeting problem using Bayesian inference, the selection on observables condition is satisfied.*

Proof. See Appendix A.2.

1.4.3 Summary

In the previous section (Section 1.3), we introduced the EE-function to model the tradeoff between the opportunity value of old knowledge and the expected value of new information. In this section, we introduce a Bayesian inference framework using Gaussian Processes to model the firm's information and uncertainty. We also discuss an important identification issue; we show that our Bayesian inference approach overcomes potential selection issues. In the next section we develop an algorithm that allows the firm to evaluate and optimize the EE-function.

1.5 EE-function Evaluation and Optimization

In this section, we describe how to both evaluate the EE-function and find π, the optimal assignment policy in t_1. Because of the information externality, the evaluation and optimization of the EE-function are two interdependent tasks. The optimal assignment for each customer this period (t_1) depends upon the assignments for other customers. As a result, the optimization is not separable across customers. Instead, the firm must solve a combinatorial problem and jointly optimize t_1's assignments for every customer. We formalize this procedure as an algorithm: One-step Look Ahead Targeting (OLAT).

The algorithm also optimizes the expected total profit (Equation 1.1) instead of merely minimizing estimation errors. The algorithm includes three key components. First, we use Gaussian processes to model the firm's information and uncertainty in each period (Section 1.4). Second, we evaluate the EE-function based on the GP prior. Third, we solve the joint optimization problem overcoming the combinatorial challenge.

We begin the section by first proposing an algorithm to evaluate the EE-function. We then move to optimization, and describe the characterization and quantification of externalities. Finally, we show how to use the algorithm to find the optimal assignment policy π.

1.5.1 EE-function: Evaluation

To find the optimal assignments for this period's customers we need to evaluate the EE-function. The IC-function (Equation 1.2) can be directly estimated from the posterior means of the response function. For the IV-function (Equation 1.3),[17] as mentioned in Subsection 1.4.1, in t_1 we need to evaluate $\mathbb{E}_{t_2}[\cdot|\mathbf{H}_{t_2}(A_{i,t_1}, \mathbf{A}_{-i,t_1})]$. One difficulty is that the firm does not observe the outcomes for this period's customers before making these assignments, yet these outcomes will contribute to the assignments for next period's customers, and need to be estimated to evaluate the IV-function. Therefore, the firm needs to extrapolate one step ahead to anticipate how its assignments will change next period depending upon this period's outcomes.

To solve this extrapolation challenge, we construct artificial trajectories, and leverage the pointwise normality property of our GP framework. This largely eliminates the integration challenge when computing the expectations. In addition, we use a simulated estimator for computing the expectations in the IV-function (Equation 1.3).

For the purpose of discussion, we focus on the action evaluation for a focal customer.

[17] We use the same method to estimate the i-IV-function and thus the IE-function, defined in Equations 1.6 and 1.7, for the individual optimization approach.

Specifically, with an interim assignment policy π (the policy being evaluated) for all other t_1 customers, we draw a batch of t_1's outcome samples $\widetilde{Y}_{t_1}$ from the generative model in Equation 1.10 (based on the posterior distribution of the response function r). As discussed in Subsection 1.4.1, these draws are based on:

1. the t_1 covariate values X_{t_1},

2. assignment to the focal customer A_{i,t_1}, and

3. assignments to other t_1 customers assigned by the interim policy π.

We then construct an artificial history, $\widetilde{\mathbf{H}}_{t_2}$, combining the observed history in t_1 with these artificial samples (we use $\sim$ to denote synthetic data and measures). Finally, based on these artificial histories, we re-learn an *artificial response function*, $\widetilde{r}(\mathbf{x}', a')$, as if we are in t_2. An associated t_2 *artificial exploit policy* $\widetilde{p}_{t_2}\left(\mathbf{x}' \middle| \widetilde{\mathbf{H}}_{t_2}\right)$ is also derived from optimizing this artificial response function $\widetilde{r}(\mathbf{x}', a')$. We repeat the above process, and use a simulated estimator to compute expectations at $\mathbf{H}_{t_2}$.

Suppose we have K artificial trajectories, and the kth artificial trajectory $\widetilde{\mathbf{H}}_{t_2}^{(k)}$ gives an artificial t_2 response function $\widetilde{r}^{(k)}(\mathbf{x}', a')$, which gives the maximum posterior means of t_2 customers.[18] The simulated estimator of the inner expectation is just the simple average of the maximal posterior means obtained from the K artificial trajectories. We defer the details of this simulated estimator to Appendix A.1.

The simulation samples (artificial trajectories) are drawn from the posterior distribution of the most recently learned profit function $r(\mathbf{x}, a)$. An important feature of the GP model is that it is easy to draw samples from the exact posterior distribution. Notably, we do not require an MCMC model, which reduces computation requirements (and would introduce an additional approximation).

The information gain from this procedure is twofold. First, GP inference provides a predictive posterior distribution for any covariate location. With this simulated estimator,

[18] The maximum of all posterior means with respect to action.

we directly leverage the quantified uncertainty in the GP model to guide t_1's assignments.

Figure 1.1: Summary of EE-Evaluation Algorithm

EE-function evaluation for a cluster of customers

1 Compute the opportunity cost of information directly from the existing profit function.

2 Compute the expected value of information for the two scenarios respectively:

(a) Assign the focal customer with an experimental action.

(b) Exploit the focal customer with the current exploit policy.

It is computed using Step 3 through 9.

3 **repeat** the simulation many times, and for each round of simulation:

4 From the posterior distribution of response function, simulate artificial outcomes.

5 Re-learn the response function using data with the artificial outcomes for this period.

6 Derive an artificial exploit policy from optimizing the learned response function.

7 Assume, that next period the firm will assign actions according to the artificial exploit policy.

Compute the expected profits next period with the artificial response function.

8 **end repeat**

9 Compute the expected value of information using a simulated estimator with these computed expected next period profits.

10 **return** EE-function values by adding together the opportunity cost and expected value of information.

Second, GP posteriors, defined in Equation 1.9, allow us to make use of the information in t_2's covariate values $\mathbf{X}_{t_2}$, in the evaluation of the GP kernel functions [48]. Therefore, we can directly evaluate the projections of t_2's policy and projected outcomes at precisely these covariate values $\mathbf{X}_{t_2}$. The idea that knowledge of future covariate data can improve performance is known as transduction in the transfer learning literature [53], [54]. Although knowledge of $\mathbf{X}_{t_2}$ helps to improve the performance of the proposed method, this knowledge is not required. Our proposed algorithm works with any assumed or empirical distribution of t_2 customer covariates.

We label the EE-function evaluation algorithm "`EE-Evaluation`" and summarize the algorithm in Figure 1.1. We provide detailed pseudo-code for the algorithm in Appendix A.3.

1.5.2 Externalities and Externality Metrics

In this subsection, we discuss how to incorporate externalities when making assignment decisions for t_1 customers. The major challenge is the combinatorial problem that we discussed in the Introduction. Because the optimization is not separable across customers, we need to jointly consider every (customer, action) pair. The most straightforward method is to take an "*enumeration*" approach; we fully enumerate the interim assignments $\mathbf{A}_{t_1}$, and iterate over combinations until converging to a fixed point that yields the desired policy.

However, this enumeration approach has two issues. First, it does not allow for meaningful interpretation of the externalities. Specifically, for any two customers in t_1, we do not see how the assignment to one affects the other.[19] Second, since the optimization is a combinatorial problem, this enumeration approach becomes computationally infeasible when covariates are continuous, when there are many covariates, or when the action space is large. Without a clear spatial proximity map of t_1's customers, it is hard for the fixed-point algorithm to reach convergence.

To address these issues, we propose a clustering approach. We give an illustrative example in Appendix A.1. With this approach, a customer i is clustered to Cluster $g \in \mathcal{G}$.[20] The interim assignment vector of Cluster g is given by an externality metric $\mathbf{E}^g$, which is a vector of length $|\mathcal{A}| - 1$. The a-th element of $\mathbf{E}^g$ represents the number of Cluster g customers

[19] This can be partially addressed by separately solving the problem with the joint optimization approach (across all of t_1's customers) versus the individual optimization approach (see the earlier discussion in Subsection 1.3.3). In particular, it indicates how much similar information has been explored or under-explored, and the relative importance of information from the focal customer.

[20] These clusters $\mathcal{G}$ can be either single layer clustering, as given by a k-Means algorithm, or some more sophisticated nested-clustering, which preserves more information about all t_1 customers. In this proof of concept, we use single layer clustering.

assigned action a under the interim assignments. The covariate space is gridded to a discretized space by the clustering $\mathcal{G}$ and the metric $\mathbf{E}^g$. The clustering is based on customer's targeting covariates, which could also include the customer's past responses.

We assume there is no information externality across clusters, which allows us to solve the optimization problem within each cluster without regard to other clusters. Specifically, we use $\mathbf{E}^g$ to guide the search for assignments $\mathbf{A}_{t_1}$ in the action optimization iteration. In the evaluation steps, to ensure we get exact estimates of information values, we still use customers' *own* covariates $\mathbf{X}^g_{t_1}$ as their covariate inputs, including the inference of the posterior distributions and the estimation of the expected value of information. Because clustering only affects the optimization step and not the evaluation step, the estimated value of information is still exact (given an assignment policy).

The benefits of clustering (covariate space gridding) are two-fold. First, it quantifies externalities from different sources, and provides a clear interpretation of how these externalities affect action assignments. The firm also knows the extent to which customer similarity affects each other's assignments.

Second, it breaks down the joint optimization problem among all of this period's customers to many smaller joint optimization problems among similar customers (within each cluster), making the algorithm more tractable. Specifically, the firm does not need to jointly optimize the assignment decisions for all of t_1's customers. Instead, it can jointly optimize across the subset of customers within a cluster. Moreover, when assuming no information externality across clusters, the optimization can be parallelized (across clusters) during computation.

Third, we can compare clustering to the benchmark case (classic adaptive learning solutions) of not considering information externalities, where each customer in a batch can be seen as their own cluster using our framing. Then, this clustering approach practically improves the performance by allowing for more information externalities across clusters. These included information externalities are also the most prominent ones, since customers grouped in the same cluster are also the most similar customers. This is based on the model analysis in Section 1.3; the information externality is larger when the associated customers are most similar.

Finally, our model implies that the information externality needs to be large enough to affect any assignments compare to the individual optimization approach. By assuming out the relatively smaller information externalities across clusters, which might not change any assignments anyway in the exact (ideal) solution, it also helps to avoid overfitting the decision model.

1.5.3 EE-function: `OLAT` Algorithm

In this subsection, we propose a local improvement algorithm, `OLAT`, to jointly evaluate the EE-function and learn t_1's optimal assignment policy. Because the EE-function is conditioned on an interim assignment policy, the `OLAT` algorithm is a loop that iterates between evaluation and optimization. In each loop, the algorithm proposes an interim assignment policy, and uses the `EE-Evaluation` algorithm to evaluate the EE-function value of this interim policy. It then proposes a new assignment policy based on the current evaluation of the EE-function and the cluster structure. The algorithm iteratively improves the interim assignment policy and will reach convergence. The output includes an estimate of the EE-function together with t_1's optimal assignment policy. The (local) fixed point property ensures that the (local) maximum of the EE-function and the (nearly) optimal assignment policy coincide.

Figure 1.2: Summary of `OLAT` Algorithm

OLAT: One-step Look Ahead Targeting

```
parallel the optimizations of each cluster
    while not converge or within iteration limit
        Evaluate the EE-function using EE-Evaluation (Figure 1), based on this period's most
        recent interim assignment policy.
        The key steps in this procedure include:
            Simulate artificial outcomes and construct histories from the response function.
            Re-learn the artificial response function and the artificial exploit policy.
            Compute EE-function values.
        Determine the search direction and propose a new interim assignment policy with the
        updated EE-function.
    end while
end parallel
return This period's assignment policy
```

We summarize the `OLAT` algorithm in Figure 1.2 and provide a formal pseudo-code in Appendix A.3. We establish the convergence properties of the algorithm in Result 3.

Result 3 (Convergence of evaluation algorithm) *Within each cluster, the `OLAT` algorithm converges to a fixed point at which this period's assignment policy locally maximizes the EE-function of that cluster.*

Proof. *See Appendix A.2.*

Practically, for faster convergence and robustness, we use ε-greedy [55] directed search to almost always search on the direction containing the cluster with the highest EE-function value improvement, with a small probability of randomly selecting a direction to search. Cluster information is used to assist the search, and we gradually sample more customers from the cluster aligned with the direction of choice, and accept the new policy deviation using a similar ε-greedy rule.

By leveraging the clustering procedure, we break the joint assignment optimization problem for this period's customers into smaller parallel problems (see discussion in Subsection 1.5.2). This makes the `OLAT` algorithm computationally more efficient.

1.5.4 Summary

We have proposed a nested combination of two algorithms that jointly estimate the expected costs and benefits of the explore-exploit tradeoff, and iteratively optimize the assignment policy in t_1. In the next section we implement the algorithm using data from a large field experiment conducted to help a membership wholesale club prospect for new customers.

1.6 Empirical Validation

1.6.1 Data Description

In this section, we provide empirical evidence to validate the `OLAT` algorithm, using data from a field experiment. This data is obtained from a single large scale direct mail targeting experiment, conducted by [8] in collaboration with a major retailer. This experiment was conducted in spring 2015 with a wholesale membership club, and was designed to recruit new members. We will focus on two experimental conditions in the experiment: a free $25 paid membership (`Mail`), and a no mail control condition (`No Mail`).[21]

We observe the treatment assignments, 13 targeting covariates, and an outcome variable measuring the profit earned from each household in the 12-months after the treatments. The profit measure includes mailing costs, membership fees, and profits earned in the membership club. The targeting covariates were purchased by the retailer from a third-party data provider. The targeting covariates are standardized, and the profit variables are logarithmized with signs preserved. As a preliminary step, we regressed the outcome measures on the covariates and identified three covariates that are significant at the 5% level: Age, Past

[21] The experiment had a total of six conditions.

Response Rate, Single Family Home.[22] We will restrict attention to these covariates in our analysis.

1.6.2 Three-Wave Experiment Construction

We assume that all customers within a carrier route will receive the same marketing action. Carrier routes are created by the USPS and literally represent the routes used by individual mail carriers. Each carrier route includes approximately 400 postal mailing addresses, located in the same neighborhood. Because all customers with a carrier route receive the same action, we aggregate the household level data to the carrier route level. More precisely, within the same carrier route, we *separately* aggregated the outcomes and covariates across households that received a specific treatment (using a simple average). Aggregating and targeting at the carrier route level offers an important advantage; within each carrier route we observe an outcome for each of the two treatments. This allows us to evaluate any carrier route level targeting policy.

The carrier route-level data consists of 5,379 unique carrier route observations. We treat each carrier route as a different "customer," and randomly group the carrier routes into three "batches" of equal size. We treat each of these batches as a "period." Using these batches, we can construct history exactly as given in Equation 1.1. One batch is assigned to represent "past year" (t_0), a second batch is assigned to "this year" (t_1), and the final batch is assigned to "next year" (t_2).

We use this aggregated dataset as the "ground truth," because the counterfactual outcomes are complete and known with respect to any marketing actions. During the validation process, we simply select the outcome (among the potential outcomes) associated with the assigned action.

Notice that an important benefit of constructing our validation using a single field experiment is that we abstract away from non-stationarity problems.[23] This focuses the validation

[22] The outcome of this regression is reported in Appendix A.

[23] We can also flexibly control the amount of non-stationarity introduced into the three-wave by introducing

on the algorithm itself, rather than introducing external confounds.

1.6.3 `OLAT` and Benchmark Implementations

We separately evaluate the choice of the $25 paid membership promotion (`Mail`) versus `No Mail` promotion. We first cluster the carrier routes based on their covariate values into 10 clusters using K-Means.

The firm has three batches of data, and its decision problem starts from past year: it wants to target this year's customers, and knows that it will target customers next year. The firm's objective (Equation 1.1), is to maximize its total profit from this year and next year's customer. We use the data in the following way:

Past Year past year's customers received randomly assigned actions. Specifically, they were equally likely to receive the `Mail` and `No Mail` actions, and these actions and outcomes are the same for all of the benchmark policies.

This Year we use the `OLAT` algorithm (Section 1.5) to find an assignment policy π for this year's customers. We also separately implement the assignment policies recommended by each of the four benchmark policies.

Next Year we train an exploit policy using past year and this year's actions and outcomes. Notice that this year's actions and outcomes are different for each benchmark policy. We then implement this exploit policy on next year's customers.

The four benchmark policies using to assign actions to this year's customers include:

`Explore (Random)` this policy randomly assigns actions this year. It uses a random policy with probability $q \in [0, 1]$ of assigning the focal action `Mail` as this year's assignment

(known) covariate shifts or noise.

policy. We use $q = 0.3$ in the results reported in this section, and report the performance of this policy for a wider range of assignment probabilities in Appendix A.[24]

`Exploit` this is the exploit policy (p_{t_1}) trained using past year's actions and outcomes.

`IE` this policy is based on the individual EE-function optimization, which is the individually optimal counterpart of `OLAT` (Equation 1.7).[25] It does not consider any externalities between customers within the same batch. `IE` is used as the assignment policy for this year's customer and is learned using past year's data and this year's targeting covariates.

`Thompson` this is the classic Thompson sampling (posterior sampling) algorithm [17]. It is a heuristic, which maximizes expected profits using parameters obtained through sampling. It is one of the most popular bandit algorithms and is shown to perform well in many adaptive learning domains. It is a representative algorithm in the classic adaptive learning that overlooks information externalities.

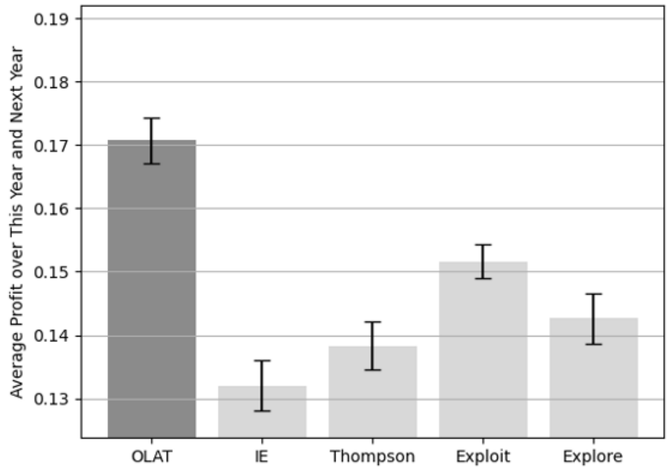
1.6.4 This Year's and Next Year's Performance

In Figure 1.3, we report the aggregate performance of `OLAT` and the four benchmark policies, measured by the average profit per customer (cumulated across this year and next year's batches).

[24] We use $q = 0.3$ because 30% of past year's carrier routes have a positive lift from the treatment. In this respect, the `Explore` policy uses information from past year's customers. Notice in Appendix A that $q = 0.3$ performs better than the evenly randomized policy ($q = 0.5$).

[25] IE is also comparable to the KG [4] and EVI algorithms [3] in the Bayesian optimization literature.

Figure 1.3: Average Profit over This Year and Next Year

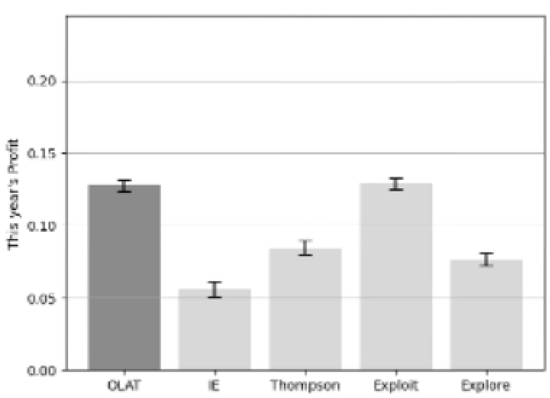


This figure reports the average profit per customer earned this year and next year from each method. Error bars indicate 95% confidence intervals.

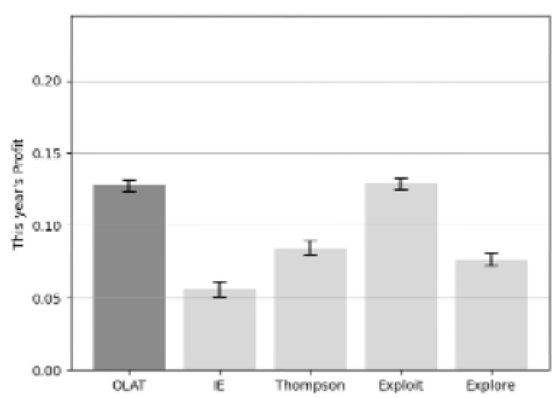
The results in Figure 1.3 confirm that `OLAT` outperforms all of the benchmark policies. The reasons are better illustrated by decomposing the average profit into the average profits earned this year and next year (see Figure 1.4). The `OLAT` solution results in a very similar profit this year compared to fully exploiting (`Exploit`). However, it then generates a substantially higher next year's profit than `Exploit`.

Figure 1.4: Decomposing Profits into This Year's and Next Year's Average Profits

(a) This year's profit



(b) Next year's profit

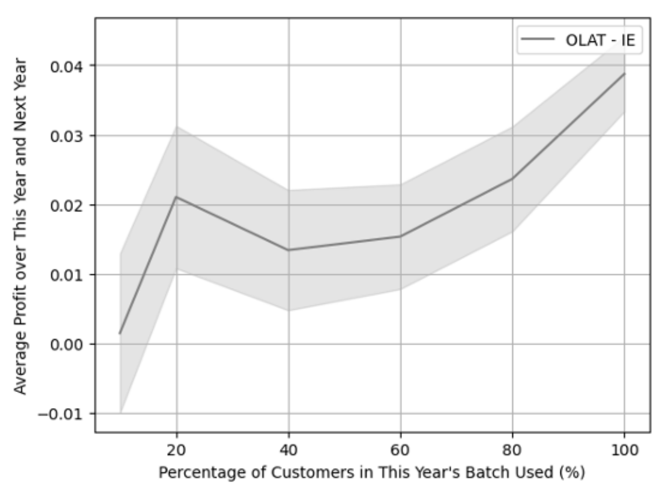


This figure reports the average profit per customer earned this year (subfigure 1.4a) and next year (subfigure 1.4b) from each method. Error bars indicate 95% confidence intervals.

In a later figure (Figure 1.6), we report the actual number of deviations this year for each policy. This reveals that the `OLAT` policy deviates from the `Exploit` policy for just 8% of this year's customers. Intuitively, by choosing the right customers to deviate with, the `OLAT` policy sacrifices only a small amount of profit this year, and generates significantly higher profits next year.

Other methods also sacrifice profit improvements this year, but they do not resolve the explore-exploit tradeoff as efficiently. In our experiment setting, the return to exploring is low for many customers, but high for a select group of customers. The `OLAT` algorithm out-performs the other policies by doing a better job of identifying the right customers to experiment with.

Figure 1.5: The Dominance of the Joint Optimization Method (`OLAT` policy)



This figure reports the difference in the average profits per customer earned this year and next year between two policies (`OLAT` and `IE`), when varying the sample size in this year's batch. Shaded regions are 95% confidence intervals.

Figure 1.5 compares the performance of the `OLAT` policy with the performance of the individually optimal `IE` policy, when varying the number of customers in this year's batch. As we have already seen, the `OLAT` policy dominates the `IE` policy. In Figure 1.5, we see that this dominance generally grows as the size of this year's batch increases. This is because the information externality becomes more pronounced when the sample size in this year's batch is larger. Intuitively, there is a greater likelihood that independently exploring with observations within a cluster will result in duplication of information, because in larger samples, observations tend to be closer together (in covariate space). Joint optimization of the information value becomes more important as the density of customers within a cluster increases.

1.6.5 Rebalancing Exploration and Exploitation

We use EE-function to directly measure how well each of the benchmark methods manage the exploration-exploitation tradeoff. Recall that this is an individual level function, which measures the information value of taking an action, minus the cost of that action (compared using this year's exploit policy p_{t_1}). The `OLAT` algorithm is explicitly designed to jointly maximize the EE-function. As we would expect, this policy achieves the best performance. Thompson sampling (`Thompson`) is also designed to balance this tradeoff. However, Thompson sampling is a heuristic, and the findings in Figure 1.4 confirm that the policy produced by this heuristic is not as good at resolving the exploration-exploitation tradeoff as the `OLAT` policy. One explanation for this is that Thompson sampling does not account for information externalities within a cluster. This can result in too many deviations from the current optimal policy among similar customers, reducing the incremental information learned from each deviating customer.[26]

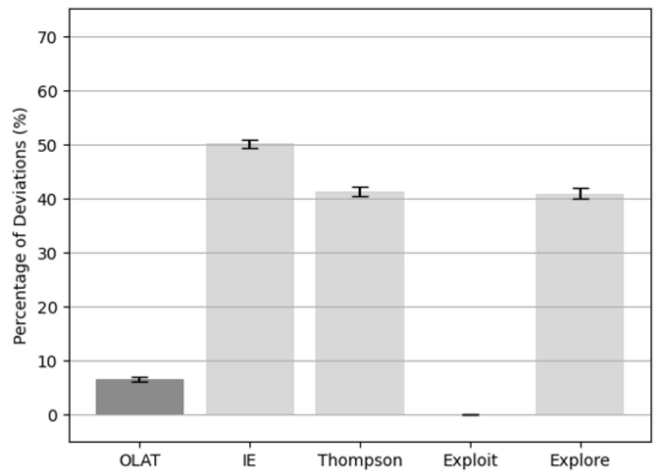
The individual EE-function optimization (`IE`) suffers from the same limitation. Like the policy produced by the `OLAT` algorithm, the `IE` policy is explicitly designed to maximize an information value function. However, the `IE` policy optimizes for each customer individually, and does not consider the information externalities between customers. This will also tend to result in too many deviations from the current optimal policy among similar customers.

Figure 1.6 demonstrates the perils of over-exploration. The figure reports the percentage

[26] Another possibility is that the `IE` and `Thompson` policies lead to too much exploiting among similar customers. Information externalities can lead to a policy underestimating the incremental information learned from each deviating customer.

of deviations from the current optimal policy in this year's batch.[27] The `IE` and `Thompson` policies both deviate more often than the `OLAT` policy. This is evidence of over-exploring, and can be attributed to the `IE` and `Thompson` policies ignoring information externalities between neighboring customers. They recommend too many deviations among similar customers.

Figure 1.6: Percentage of Deviations from Current Optimal Policy



This figure reports the percentage of deviations this year from this year's optimal policy of different policies. Error bars indicate 95% confidence intervals.

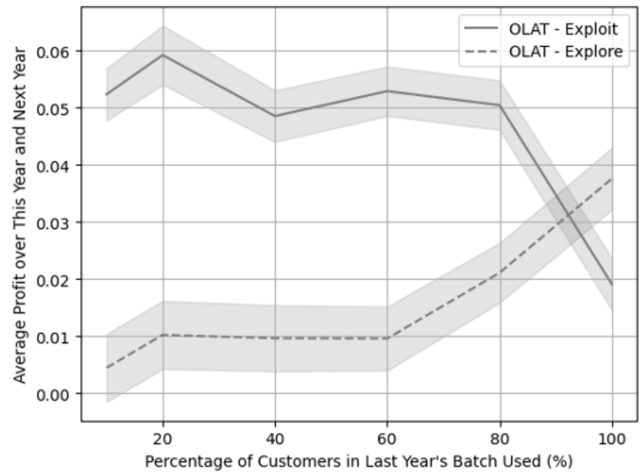
An alternative explanation for the under-performance of the `IE` and `Thompson` policies is that they do not choose the right customers to explore with. The optimal policy chooses not only how many, but also which customers to explore with.

1.6.6 Tradeoff Between Existing Knowledge and New Information

Resolving the exploration-exploitation tradeoff also depends on the amount of existing information. To investigate the impact of existing knowledge, we vary past year's batch size. When past year's batch is larger, there is more existing information, and less need to explore this year. In Figure 1.7, we compare the `OLAT` policy with the and `Explore` and `Exploit` policies. We see that the dominance of `OLAT` over these benchmarks depends upon the size of past year's batch.

[27]The `Explore` policy plotted in Figure 1.6 uses $q = 0.3$ (as the probability of receiving action `Mail`).

Figure 1.7: Varying Past Year's Batch Size



This figure reports the differences in the average profits (over this year and next year) per customer of three policies (`OLAT` versus `Explore`, and `OLAT` versus `Exploit`), with different batch sizes of past year. Shaded regions are 95% confidence intervals.

When we have very little existing knowledge, exploitation this year is relatively unprofitable. At the other extreme, when we have a lot of existing knowledge, exploration is no longer needed, and exploitation becomes more profitable. This is reflected in Figure 1.7. The `OLAT` policy is more profitable relative to `Exploit` when there were fewer customers in the past year, and the reverse is true when comparing `OLAT` with `Explore`.

1.7 Concluding Remarks

In this paper, we study the batch targeting problem. This is a common problem faced by firms conducting marketing campaigns in both digital and non-digital channels. Firms need to assign marketing actions to a large number of customers in each campaign, and there is delay in updating the model between campaigns. We propose a method for solving this problem that balances the costs and benefits of experimentation. The method evaluates the incremental value of information provided by each customer, conditional on the information provided by other customers.

This is the first paper to identify information externalities in exploration-exploitation problems when customers arrive in batches. Our model and algorithm explicitly account for variation in customer locations both within and across batches. Because information externalities are magnified among neighboring customers, controlling for variation in customer locations is important in many marketing settings.

Although we focus on targeted marketing campaigns in batches, information externalities generalize to many experimental design settings that use sequences of experiments. Hidden costs are incurred when ignoring information overlaps between experiments. By anticipating these information externalities and coordinating designs across experiments, firm may be able to improve their decisions and/or reduce the size of their experiments.

We acknowledge several limitations in our research. First, we treat the size of batches as exogenously given, and do not consider the optimal design of each batch size. Batch sizes affect the expected value of information and ultimately the design of assignment policies. Second, we keep the methods used in each element of our `OLAT` algorithm as simple as possible. These methods could be improved. For example, we can use more sophisticated directed local search algorithms, or even deep graph networks for clustering and characterizing information externalities. Third, we do not consider any customer dynamics in the exploit (targeting) policy design. All of these limitations suggest promising paths for future research.

Chapter 2

The Invisible Hand behind Luxury Consumption

Abstract

The economics and marketing literature has primarily focused on market economies and studied factors such as price and advertising when analyzing customer demand. However, in non-market-oriented economies, social factors like corruption can have a significant influence on customer decisions. In particular, this paper focuses on the demand for luxury products, which are widely used for gift-giving and even bribery in emerging markets. One possible mechanism is that when the relative size of non-market-oriented sectors in the local economy increases, luxury products can be used to identify those who have a higher willingness to pay for scarce resources. As a result, the demand for luxury products moves together with the degree of corruption. By leveraging natural experiments of top-down anti-corruption campaigns that temporarily halt this channel, an empirical study is performed using a comprehensive dataset that covers the sales of all cigarette brands and the local social environment in China. The results suggest that these social factors can have an unanticipated impact on the demand for luxury products.

2.1 Introduction

The economics and marketing literature has mainly focused on economic variables such as price and advertising, as well as individual heterogeneity, when analyzing customer demand. However, customer behavior can also be influenced by broader social and cultural factors. In contrast to the extensive literature that quantifies the impact of various economic variables, empirical research on the effect of sociocultural factors on product demand is limited, especially for symbolic goods or products with well-recognized social meanings. Using data from China's tobacco industry, this paper seeks to provide evidence for the significant impact of sociocultural factors, particularly gift-giving and bribery, on luxury consumption. We also explore how relationship between the importance of these sociocultural factors and the size of the non-market-oriented economy.

China is the world's largest tobacco-producing and cigarette-consuming country, accounting for 44% of the world's total cigarette consumption in 2017, according to Euromonitor data. Despite the global trend of declining tobacco consumption in the last decade, cigarette sales in China remain high, with the growth of the high-end sector outpacing that of the low-end sector in recent years.

China has a unique cigarette culture, where cigarette sharing and gifting are ubiquitous [56]. A survey conducted in six large Chinese cities in 2006 found that cigarettes were the top choice for gift-giving, accounting for about 30% of gifting occasions. Additionally, 63% of survey respondents indicated that gift-giving was the primary reason they purchased high-end cigarettes. [1] Cigarettes can be considered a social currency [57] that enhances interpersonal connections, with expensive cigarettes signaling higher value. This is particularly true when it comes to gaining an advantage in resource allocation; high-end cigarettes can both build connections with officials and signal high self-quality.

While economic and demographic factors are often considered the main drivers of luxury

[1] The six cities are Beijing, Shanghai, Guangzhou, Chengdu, Shenyang and Hangzhou. Survey results are available at .

and, in particular, high-end cigarette consumption [58], [59], the prevalent gift-giving culture in China leads us to conjecture that the demand for high-end cigarettes in this market is also linked to the importance of personal connections in economic activities. The need for personal relationships is more prominent in non-market-oriented economies with heavy government involvement than in more transparent and market-oriented economies. Previous literature shows that Chinese firms are motivated to establish and maintain good relationships with government officials for favorable treatment and more resources [60], [61].

Therefore, we hypothesize that the demand for high-end cigarettes is likely to be higher when the size of non-market-oriented sectors in a local economy is larger or when the government plays a more significant role in economic activities. To test our hypotheses, we compile a comprehensive dataset containing city-level sales information over time for all cigarette brands that are allowed to be sold in China. We complement the sales data with city characteristics, including variables indicating local economic development, such as GDP per capita, and data related to government-oriented and market-oriented economic activities at the local level, collected from multiple sources. Our empirical analysis reveals significant and robust evidence that regions with higher government involvement in the economy are associated with higher demand for premium cigarettes, while regions with larger market-oriented sectors see lower demand, after controlling for local economic and demographic factors.

When government officials have discretion over resource allocation and are highly involved in economic activities, corruption is more likely to occur. Premium cigarettes, especially top brands, are often reported as bribery gifts to government officials. Therefore, we further conjecture that the demand for premium cigarettes is related to the level of bureaucratic corruption. A direct test of this conjecture is not straightforward because corruption is illicit and difficult to measure [62]. To identify this effect, we resort to indirect evidence of corruption and take advantage of recent anti-corruption campaigns in China to investigate the effect of corruption on cigarette demand. These staggered top-down anti-corruption campaigns serve as natural experiments that temporarily halt corrupt activities. The central government's objective when designing the campaign agenda is to create surprise; this coincides with our need for identification with an exogenous shock, allowing us to identify the effect of corruption on luxury consumption with a difference-in-differences design. Our analysis shows that when the degree of corruption in the local area goes up, the sales of high-end cigarettes indeed also increases.

The institutional and cultural factors that drive high-end cigarette demand are intertwined. Relying on a large-scale cross-regional family survey in China, we find direct evidence that regions with a more prominent gifting culture, as measured by higher amounts of reported gift-giving or gift-receiving, are associated with higher sales of premium cigarettes. Interestingly, we also find evidence that regions with more democratic local governance and a more inclusive mindset are associated with lower sales of expensive cigarettes.

By compiling data from various sources and utilizing different measures, we consistently find that the demand for high-end cigarettes in China is affected not only by economic and demographic factors but also by sociocultural factors, and these sociocultural factors are more prominent when the size of non-market-oriented sectors is larger.

Our research highlights the importance of social and cultural factors in driving luxury consumption in non-market-oriented economies. While it is well-recognized that social context affects human behavior and customer needs, few studies have attempted to quantify such effects on demand. One challenge is the measurement of these factors. Unlike economic and demographic variables, which have clear definitions and reliable measurements, social and cultural factors are not directly observable and have largely relied on surveys in previous studies, such as Hofstede's five cultural dimensions [63]. Another challenge arises from identification. Past research typically leverages cross-national variation to identify the effect of culture on product adoption [64]–[66]. In our study, we address these challenges by leveraging an exogenous policy shock that temporarily changes the way people interact and by using regional variation to identify the impact of sociocultural factors on luxury consumption. Our results show that cultural and institutional factors play an important role in explaining the differences in demand across regions.

Our research contributes to the literature on cigarette consumption in several ways. First, previous studies have traditionally focused on the effect of advertising on cigarette sales [67]–[70], the effectiveness of anti-smoking campaigns and regulations on tobacco control [71]–[73], and the consequences of tax increases on cigarette consumption [74]–[77]. In contrast, our study focuses on the impact of sociocultural factors, which has received little attention in the existing literature.

The remainder of this paper is structured as follows. Section 2.2 introduces the industry background and the cigarette data. Section 2.3 through Section 2.5 conduct empirical analysis of institutional factors, corruption and sociocultural environment on the demand for premium cigarettes. Section 2.6 concludes the paper.

2.2 Industry Background and Data

2.2.1 Industry Background

The tobacco industry in China is highly regulated. The state-owned China National Tobacco Corporation (CNTC) enjoys the status of a virtual monopoly, with 98% of the domestic market share. The structure of the organization is vertical. Under the corporation, there are provincial level branches that supervise the production and sales of tobacco products in the local area. The CNTC contracts out production to licensed local tobacco manufacturers, and it is also responsible for the distribution and sales of tobacco products. Approximately 40 cigarette manufacturers which are located in every province of China except Tibet. Tobacco and cigarette production contributes significantly to local fiscal revenue through tax.

While the production of cigarettes is local, the distribution is nationwide and responds to market demand. The wholesale price of cigarettes is the same across regional markets. Retail prices may vary across retail outlets, yet there are guiding retail prices announced by the

State Tobacco Monopoly Administration (STMA), the government agency that monitors the industry. Retail stores need to obtain a license from the STMA to sell cigarettes. Advertising of cigarettes in print, radio or on TV is banned in China.

The cigarette tax in China combines a specific excise tax and an *ad valorem* tax. Before May 2009, the specific tax was the same across cigarettes while the *ad valorem* tax was two-tiered: 30% for cigarettes with producer price less than 5RMB per pack and 45% for those equal to or above this price level. In May 2009, China adjusted the *ad valorem* tax rate schedule: the tier 1 rate was increased to 36% and the tier 2 rate was increased to 56%.[2] The tax policy change occurs during our sample period and we use year dummy to control for such effect on cigarette demand in late analysis.

2.2.2 Sales Data

We obtain cigarette sales data from the STMA. The data contain detailed information on the sales of cigarette of all the brands across 273 cities at the year level from 2007 to 2014. Monthly data are available from January 2013 to July 2015. Each observation is a sales record of a cigarette product and includes the following information: the name of manufacturer, brand, SKU, price category, wholesale price, city and sales volume. We identify high-end cigarettes by price category. There are 9 price categories identified in the data from high to low. Category 1 (supreme) refers to products whose wholesale price is RMB 600 or above per carton (10 packs); category 2 (high class) refers to wholesale price in the range of RMB 500 to 600 per carton. Figure B.1 shows the distribution of the number of SKUs that fall into each price category. We define a product as high-end if it belongs to the first two price categories.[3]

We construct two variables that are of main interest to our analysis. The first is the total

[2] The 2009 tax adjustment also re-classified the two tiers. Cigarettes with producer price 7 RMB per pack or above were subject to the tier 2 tax rate while those below 7 RMB per pack were subject to the tier 1 rate. In other words, more expensive cigarettes face a higher tax increase with the tax adjustment.

[3] We also run robustness checks using alternative definition of high-end cigarettes: (1) products in the highest price tier; (2) Well-recognized premium brands such as Zhonghua. The results are qualitatively the same.

sales of high-end cigarettes in a city during a year or month. The second is the percentage of high-end cigarette sales out of the total cigarette sales in a city during a year or month. The second variable measures the intensity of the demand for premium cigarettes, which is more comparable across regions. Figure B.2 shows the ratio of high-end cigarettes at the city level in 2013 on the map of China. There is heterogeneity in the high-end cigarette demand across provinces, and such a difference cannot be fully explained by the state of economic development or income effect. Figure B.3 shows the level of GDP per capita in 2013 in each city on the map. Some areas, such as the southwest region of Yunnan Province, were behind in terms of GDP per capita yet they saw a high proportion of premium cigarette sales. Such contrast motivates us to further explore the factors driving the demand for high-end cigarettes.

2.2.3 Cigarettes as Gifts

Gift-giving facilitates establishing and maintaining interpersonal relationships [78] and is an important ritual in Chinese culture that emphasizes harmony and social connections [79]. Cigarettes, particularly expensive ones, are commonly used in gift-giving contexts in China. Cigarettes are ideal gifts not only because there is a large population of smokers but also because cigarettes are easy to sell or exchange at stores that recycle gifts [80].

One indication of cigarettes as gifts is the strong seasonal pattern of cigarette sales. Using the monthly data from 2013 to mid-2015, we plot the national sales of premium cigarettes over the 12 months for each year in Figure B.4. The pattern is clear and consistent: the highest sales occur in January followed by September. The spike in January is due to the celebration of Spring Festival (Chinese New Year), which usually falls between late January and mid-February. Spring Festival is the most important holiday for the Chinese and the time for family reunion. Gift-giving is a tradition for this holiday. For products in gift categories, the sales from this season account for a significant portion of the total annual year sales. The second highest sales volume of cigarettes occurs in September. This coincides with the Moon Festival, another Chinese festival that calls for family reunion. Gift-giving during the festivals not only occurs among relatives, friends and acquaintances, but also extends to a larger social network.

Aside from being among the top gift choices for festivals, cigarettes are commonly used as gifts to establish and strengthen personal connections in everyday life. In the following section, we further explore this particular nature of cigarettes and investigate its impact on cigarette sales across regions.

2.3 Institutional Settings and Luxury Consumption

2.3.1 Hypothesis and Measurements

Personal connections, or *guanxi* in Chinese, are extremely important in navigating a system with heavy bureaucratic presence. At the personal level, one may need connections with government agents to increase the efficiency of dealings with the government. At the business level, firms need to establish and maintain good relationships with government officials to obtain favorable tax treatment and better government services [60], or to overcome competitive and resource disadvantages [61]. When institutional settings allow government officials more discretionary power, the need for relationship is more prominent. When the government plays a more significant role in the economy, government officials have more opportunities to engage in corruptive behavior. High-end cigarettes, along with expensive liquor and other luxury items, have been repeatedly reported as choice gifts to officials and signs of corruption.[4]

These observations lead to our hypothesis: all else being equal, the demand for high-end cigarettes is higher in regions where the government has stronger influence in the economy; the demand for high-end cigarettes is lower in regions with a more developed market economy.

[4]Example articles: http://www.chinadaily.com.cn/opinion/2017-02/17/content_28234071.htm; https://www.bloomberg.com/news/articles/2012-03-04/cigarettes-the-most-stable-international-currency.

To test this hypothesis, we need to find measures that reflect the magnitude of government involvement in the economy.

The first measure is related to the fiscal spending of the local government. Public spending reflects the involvement of the government in the local economy. Previous literature also shows that public spending is closely related to corruption [81]. We collected data from the Ministry of Finance Report and use the yearly fiscal expenditure per capita of a city as the focal measure. The variable is log-transformed as it is highly skewed.

Second, we consider the power of government agencies in a city as an indicator of government presence in the local economy. While a direct measurement is not available, we are able to collect from public sources the number of job openings posted by a local government as well as the number of applications for these openings for the year 2017.[5] We compute the ratio of the number of applications to the number of government job slots in a city as our measurement. This reflects the intensity of competition for government employment, which implies the attractiveness of government positions in a region. The variation of government job attractiveness across regions, controlling for local economic conditions, may reflect the implicit rent-seeking opportunities or power of government positions in the region.

Personal relationships become less important in a more transparent and market-oriented economy. The size of foreign investment in a region is often considered to reflect the openness of the region and the market oriented part of the economy. As a result of the open-door policy, foreign direct investment (FDI) has increased rapidly in China since the 1980s. The major types of FDI are wholly foreign-owned enterprises and joint ventures. Foreign capital typically seeks regions with a better market infrastructure and business environment, such as regions with protection of intellectual rights and agglomeration economies [82], [83]. Previous research shows that FDI has positive spillover effects on local industry and enhances the integration of the local market with the international market [84], [85]. We use the proportion

[5] Ideally we would like to collect the government job posting information for the years that match the cigarette sales data, yet data for past years are not available. The current measures are the proxy for the local demand of government jobs, which we argue can be correlated over time.

of foreign-invested enterprises' output in the gross industry output as a measure of the local market-oriented economy. The higher the share of foreign-invested enterprises is, the more likely the prevailing rules in business operation are market oriented instead of relationship-based. Therefore, we expect a negative correlation between the concentration of foreign-invested enterprises and the demand for high-end cigarettes.

In addition to collecting the above data from statistical yearbooks, we search for existing surveys that contain information about our measures of interest. First, we utilize the survey results in 200 cities from the 2009 Urban Household Income and Expenditure Survey (UHIES), conducted by the National Bureau of Statistics. The survey covers all the provinces in China and is considered nationally representative.[6] One question in the survey is directly related to our discussion above: the number of family members working in state-owned enterprises (SOE). SOEs are known for emphasizing *guanxi* or personal ties in operation and promotion. Since our analysis unit is at city level, we construct the variable *percentage of family members working in an SOE* at the city level, which is the percentage of family members working in SOEs averaged across the sampled households in a city.[7]

The second survey used is the City Public Governance Survey [87]. The survey collected data from over 6000 respondents from 220 cities in 2013. One question in the survey asked whether the respondent relied on personal relationships in dealing with affairs in government sectors. We compute the percentage of respondents that answered yes in a city as another measure for the importance of relationships in the local area.

We use these measures constructed from various sources to test our hypothesis. The list of variable definition and the data source is in Table B.1.

2.3.2 Empirical Specifications

Our main interest is to test the impact of cultural and institutional factors on high-end cigarette demand. Since the size of the smoking population in different regions varies and

[6][86] use the UHIES data from 1992 to 2003 to study the income and consumption inequality in China.
[7]The sample size of each city in the survey is 300-400.

such information is not available, we use the ratio of high-end cigarette sales out of the total sales in a city in year t, r_{it}, to measure the local demand intensity. We model r_{it} to be a function of the local characteristics, X_{it}, and institutional or cultural factors, R_{it} :

$$r_{it} = \frac{\exp(\alpha + \beta R_{it} + \gamma X_{it} + \varepsilon_{it})}{1 + \exp(\alpha + \beta R_{it} + \gamma X_{it} + \varepsilon_{it})}. \tag{2.1}$$

The functional form ensures that the dependent variable is between 0 and 1. Transformation of the above equation leads to the following specification:

$$\ln\left(\frac{r_{it}}{1 - r_{it}}\right) = \alpha + \beta R_{it} + \gamma X_{it} + \varepsilon_{it}. \tag{2.2}$$

This is the main regression model we estimate. The dependent variable, y_{it} is the log transformation of the ratio of high-end cigarette sales to other cigarette sales. Instead of using the log ratio as the dependent variable, we also estimate a model with an alternative dependent variable $\ln(S_{it})$, which is the log transformation of the total sales volume of high-end cigarettes in a city in year t.

The control variable X_{it} includes the following variables: GDP per capita (in log form), ratio of the service industry in GDP and population density of city i in year t. GDP per capita captures the income effect on premium cigarette sales. The share of the service industry reflects the economic development of a city. A higher share of the service sector indicates a more developed area, which is likely to lead to higher demand for premium cigarettes. We also include a dummy variable that indicates whether the city has high-end cigarette manufacturers. There could be a local effect [88] that promotes the consumption of premium cigarettes.

Another set of variables that may affect the cigarette demand are related to the need for gift-giving and *guanxi* in economic activities. These are summarized by R_{it}., which have been discussed in the last section. Since cultural and institutional factors do not change dramatically from year to year, our main identification is from the regional heterogeneity controlling for basic economic conditions. The summary statistics of the variables are provided in Table B.2.

2.3.3 Results

We first run a baseline model with only the set of basic control variables X_{it}. The results are reported in Table B.3. Column (1) uses the log-ratio of premium cigarettes as the dependent variable (Equation 2.2), and Column (2) uses the log sales as the dependent variable. We also use year dummies to capture factors such as tax policy change.

As expected, GDP per capita has a significant and positive effect on the demand for high-end cigarettes. We also find that cities with more developed service sectors see a higher demand of premium cigarettes. These are consistent with the income effect interpretation. In terms of the effect of local premium cigarette production, we find evidence that it is associated with higher sales of premium cigarettes. However, the ratio of high-end cigarette sales relative to other cigarette sales in these cities is lower. A possible explanation is that local production of premium cigarettes increases the total demand for cigarettes in the region and that the increase in the demand for regular cigarettes exceeds that for premium cigarettes.

To test our hypothesis, we further include variables introduced in Subsection 2.3.1 into the regression. The first two columns of Table B.4 report the result with local government spending per capita (in log form) and the demand-to-supply ratio of local government job openings as additional explanatory variables. These variables reflect the degree of government influence in a local area. We find both coefficients to be positive and significant. In other words, all else being equal, cities with higher fiscal spending and higher demand for government jobs are associated with higher demand for premium cigarettes, in both the relative sense (relative to the sales of other cigarettes) and the absolute sense (total sales volume). These results support the hypothesis proposed, that the demand for premium cigarettes increases where the government is more involved in the economy.

In Columns (3) and (4) of Table B.4, we control for the share of foreign invested firms' output in the gross industry output in a city, a variable that reflects the degree to which the economy in the region is market oriented. Consistent with our conjecture, its coefficient is negative and significant. Our interpretation is that the need for gift-giving and personal connection is lower in a more transparent and market-driven economy. Therefore, the demand for high-end cigarettes is reduced in these areas.

Columns (5) and (6) show the results when controlling for all three measures. The results indicate that the demand for high-end cigarettes is higher in regions with more government involvement, but lower in regions with a more developed market economy. This is because relationship-building and gift-giving are more prevalent in an economy with stronger government involvement.

Tables B.5 and B.6 report the regression results using the measures derived from surveys. Table B.5 utilizes information from the UHIES conducted in 2009. The variable of interest is the ratio of family members working in SOEs. Another variable derived from the survey is the self-reported total value of gifts in eating and drinking received by a family in the year. Note that both variables are computed as the average of the sampled households in a city to match our unit of analysis. We find that a higher share of SOE employment in a city is associated with a higher demand for premium cigarettes, indicating that interpersonal relationship building is more important in the state-controlled economy. We also find evidence that a higher value of gifts received is associated with higher demand for high-end cigarettes, an indication that sales of high-end cigarettes are related to gift-giving.

Table B.6 draws information from the City Public Governance Survey in 2013. The key variable constructed from the survey is the percentage of respondents in a city who indicated that they have relied on personal relationships when interacting with government sectors. The coefficient of this variable turns out to be positive and significant, suggesting that the demand for premium cigarette sales is higher in cities where personal connections are more often resorted to in dealing with the government. This is again consistent with our hypothesis.

2.4 Corruption and Luxury Consumption

The empirical evidence from the last section suggests that the demand for high-end cigarettes is higher in regions with more government intervention. This leads to the further hypothesis that the demand for premium cigarettes is related to corruption. Premium cigarettes have been repeatedly reported as gifts in bribing government officials, although they are of relatively low price compared to other luxury products used in bribery [89], [90].

It is challenging to test directly whether corruption leads to higher demand for premium cigarettes. The primary reason is that corruption is illicit and difficult to measure [62]. Many empirical studies have detected corruption by focusing on specific scenarios or activities that involve corrupt behavior [89], [91]–[95]. Following the literature, we resort to indirect evidence of corruption to investigate the effect on cigarette demand. Our hypothesis is the following: all else being equal, the demand for high-end cigarettes is higher in regions with more corruption among government officials; the demand for high-end cigarettes is lower in regions with stronger anti-corruption efforts.

We first take advantage of the anti-corruption campaigns in recent years to detect the impact of corruption on premium cigarette demand. The Communist Party of China (CPC) initiated a far-reaching campaign against corruption in late 2012 that is considered the largest in the history of the party's rule in China. The Central Commission for Discipline Inspection (CDI) is the highest internal-control institution of the CPC that enforces regulations and combats corruption. There are also provincial CDIs that take charge at the provincial level. Since 2013, the central commission as well provincial commissions have visited many cities to investigate government officials of various ranks. The inspection visits have a significant effect on curbing bribery and corrupt behavior [89], [96]. If high-end cigarettes are frequently used in bribing government officials, then the inspection visits are likely to reduce such gift-giving and lower the demand for premium cigarettes. We therefore conjecture that the demand for high-end cigarettes decreases in cities with anti-corruption inspections.

To test this hypothesis, we collected information on the timing of inspection visits by either the central CDI or provincial CDIs to cities through news reports dated back to January 2013. The newspapers we searched include major newpapers at both the state level and the provincial level. Given the monthly data of cigarette sales from 2013 to 2014, we run the following analysis:

$$y_{it} = \alpha + \beta_0 I_{it} + \sum_{s=1}^{S} \beta_s I_{i,t-s} + \gamma X_{it} + \tau y_{i,t-1} + \varepsilon_{it} \tag{2.3}$$

where y_{it} is the measure of the demand for high-end cigarettes in city i at period (month) t. I_{it} is the indicator of whether the central CDI or a provincial CDI paid an inspection visit to the focal city at period t. Since the effect may not take place immediately, we also control for the lagged effect of inspection on the demand for premium cigarettes. $I_{i,t-s}$ stands for whether there is an inspection s months before period t. In our empirical analysis, we allow for up to 5 periods of lag effect. X_{it} include city level characteristics and month dummies. We also control for the lagged dependent variable (cigarette sales) to account for state-dependence in the demand.

The estimation results are presented in Table B.7. The first two columns report the effect of inspections from the central CDI on the ratio of high-end cigarette sales or sales volume. We find that the effect of the inspection of government officials by the CDI started to appear two months after inspection began. The demand for high-end cigarettes in the focal city, by either measure, significantly decreases two months after the visit by the central CDI. The effect attains the highest level three months later and then recedes. It provides some evidence that the inspection curbs potential corrupt behavior and discourages gift-giving using premium cigarettes. Columns (3) and (4) show the results controlling for both central CDI inspections and provincial CDI inspections. It seems that inspection from the provincial

CDI has little impact on premium cigarette demand. In comparison, one may conclude that inspections from the central CDI are more effective in combating corruption, significantly dampening the demand for high-end cigarettes. The result is also consistent with [96], who find that the announcement of the CDI lowers the expected return of luxury-goods producers using data from the stock market.

In addition to utilizing the inspection events, we collected data from news media on the corruption charges of high-ranking government officials at the city and provincial levels from 2013 to 2015. Similar to the inspections, the corruption charges of high level officials send a warning signal against bribery behavior. We conjecture that areas with more reported cases, which imply stricter execution of the anti-corruption campaign, would see lower demand for premium cigarettes.

We search for only bribery cases involving government officials who assumed high level positions in the administration team and Party Committee of a city or province, including mayor / vice mayor, secretary of municipal Party Committee / vice secretary, province governor / vice governor secretary, secretary of provincial Party Committee and chief officer of important government department or bureau. These positions exist for all the cities and provinces, regardless of the size of the local government. Each record shows the name of the government official accused, the month and year when the case was publicly announced, the position that the person assumed, and the city or the province of the position. We then compute the number of corruption charges revealed in each month in each city in our dataset and construct two variables. One is the number of corruption charges on municipal officials of the city, and the other is the number of corruption charges against the provincial officials in the province where the city is located.

We run a similar regression as Equation 2.3 above, replacing the indicator of CDI inspections with the number of corruption cases. The results are presented in Table B.8. Column (1) shows that the corruption charges of provincial officers have an immediate effect on curbing the demand for high-end cigarettes, as reflected by the negative and significant coefficients of first and second month lags. However, the effect disappears for deeper time lags, implying that the pressed demand rebounds.

2.5 Social Environment and Luxury Consumption

Our empirical analysis in the last two sections, using evidence from various sources, consistently shows that the demand for premium cigarettes is positively correlated with the extent of government involvement or nonmarket orientation in the economy. Apart from economy, such correlation has deeper roots in the culture and social environment. Given the geographical dispersion and historical background, different regions in China not only experience different paces of economic reform and development, but also exhibit different cultural and social norms. The economic and cultural factors are often intertwined. More market-oriented regions are typically associated with a more open culture and more breakaway from traditions, while the less developed areas are often more conservative and emphasizing more of relationship. Following this logic, we have the following hypothesis: all else being equal, the demand for high-end cigarettes is lower in regions with a more open culture and higher in regions with more traditional relation-based culture.

To find measures that reflect the local cultural and social environment, we resort to a large scale cross-region family survey conducted in China, i.e., the China Family Panel Studies (CFPS). The CFPS is a nationally representative and biannual survey of Chinese families since 2010. The survey collects a wealth of data including both economic and social information of individuals and families.[8] The data we draw on come from the 2010 survey, which has a sample size of approximately 15,000 families nationwide. We collect the responses to the following questions that could potentially be relevant for our study:

1. How many times has your family sent gifts to others during the last year, including

[8]The CFPS is funded by the Chinese government through Peking University. It is considered the most comprehensive survey data of Chinese families and communities and is widely used in Chinese social science research. See link http://opendata.pku.edu.cn/dataverse/CFPSformoreinformation.

presents (e.g. cigarettes, wine, tea and jewels) and money (e.g. lucky money during Chinese New Year) ?

2. What is the value of the gifts, including both presents and money, that your family received during the last year?

3. Do you think learning English is important in communication? Rate on a scale from 1-5 where 1 means not important at all and 5 means extremely important.

4. Have you been to Hong Kong/Macao or Taiwan? (Yes/No)

5. Has your region implemented direct election? (Yes/No) [9]

The first set of questions (1 and 2) regards the gifting behavior itself. Questions 3 and 4 are related to communication with the outside world. The last question reflects the degree of democracy in local governance. The answers to these questions can provide information concerning the local gifting culture and the openness of the local society. We consider a region as more open if it has a higher percentage of residents who travel outside, consider foreign language as important, or has more democracy in politics. For questions (1) - (3), we compute the average value of the responses from a specific region. For question (4) and (5), we compute the percentage of respondents that answered 'Yes'. We then run the analysis parallel to Equation 2.2. The key explanatory variables are the five variables extracted from the survey.[10] The results are reported in Table B.9.

We find direct evidence that regions with more prominent gifting culture (as measured by higher amounts of self-reported gift-giving or gift-receiving) are associated with higher sales of premium cigarettes, as reported in Columns (1) and (2). Interestingly, we also find from Columns (3) and (4) that everything else being equal, regions that implement direct election and that have more residents valuing foreign language in communication or travelling outside

[9]Since 1998, various trials of direct election at the town level have taken place in many regions of China. The reform is considered to represent important democratic progress [97].

[10]Although we have data for only one year, we believe these measures are relatively stable. Our identification relies on the cross-sectional variation.

mainland China are associated with a significantly lower demand for high-end cigarettes. The results lend support to our hypothesis that the demand for expensive cigarettes is positively correlated with gifting culture but negatively correlated with local openness and democracy.

2.6 Conclusion

Economy, culture and institutional settings are intertwined and shape each other. Using various measures reflective of these underlying forces, our analysis reveals that these factors have significant influence on the sales of high-end cigarettes in China. To further evaluate their overall impact, we run an overarching analysis that include both institutional and culture factors in one regression. Recall that we have three key variables that reflect the degree of government involvement in the local economy and a set of variables that proxy for gifting culture and local openness. We report the results in Table B.10. Comparing to the baseline results in Table B.3 which only controls for economic conditions, we find that measures of culture and institutional factors can explain additional 15% of the variation in the demand of premium cigarettes.

Our identification strategy takes advantage of the heterogeneity of the Chinese market across geographical regions. The variation in economic development, local culture and institutional factors, along with local demographics, gives rise to the differences in demand and consumption. We utilize data from various sources to construct measures of such contextual factors and our empirical analysis consistently indicates that these factors are important in explaining the sales for high-end cigarettes.

Our research takes initial steps in quantifying the subtle contextual influence in product demand. Future research is required in at least two areas. First, although we have created different measures of culture and institution settings from various data sources, future research can improve the measures or provide more comprehensive measures of such contextual factors. Second, more research is needed to investigate the demand for products that not only satisfy personal use but also serve important social roles. The demand dynamics for such products would be correlated to the change in the social context.

Chapter 3

When Customer Search Stifles Product Innovations

Abstract

Conventional wisdom suggests that when an incumbent fails to innovate, there is a greater risk to the incumbent of competition from other innovators. I show conditions in which the opposite is true: by delaying innovation, an incumbent can create entry barriers that deter innovation by competitors. Consequently, both competition and innovation are suppressed. The key insight driving these outcomes is that customer search is endogenous, and absence of innovation today can disincentivize customers from searching in the future. Since customers need to search to discover innovations, when they search less, it both creates entry barriers for competitors, and reduces the competitors' incentives to innovate. Postponing innovation can benefit incumbents if it motivates customers to search less, and thus competitors to innovate less. Notably, I show that searching less is a rational customer response.

The world, people, institutions, whatever you want to call it, need time to adapt and think about these things (AI products) ... Our goal is not to have shock updates to the world, (but) the opposite.

– Sam Altman (CEO of OpenAI)

3.1 Introduction

In 2010, Apex Ski Boots released an innovative product designed to offer greater comfort for skiers – a longstanding issue with traditional ski boots. Despite offering a technologically groundbreaking product in a market missing innovation for over twenty years, the company failed to achieve commercial success. Apex's story is not an isolated case; many innovations that are able to meet highly anticipated, but unfulfilled, customer needs struggle to compete with incumbent firms and become profitable, and even more never reach consumer markets.

Conventional wisdom suggests that when an incumbent firm fails to innovate, it creates entry opportunities for prospective competitors to enter the market. If these entrants are more innovative and can develop superior products, their entry is likely to hurt the incumbent firm. The competition might intensify, and entrants' products might end up dominating the incumbent's. Entrants will have stronger incentives to innovate, which could ultimately also benefit customers through better products and possibly lower prices in the market. However, this conventional wisdom does not apply to the story of Apex Ski Boots, whose innovative product failed to gain traction in a market characterized by low innovation from incumbent firms.

In this paper, we argue that the opposite of this conventional wisdom may hold true under certain conditions. Rather than enabling entry, less innovation by an incumbent firm can create entry barriers for outside innovators. In addition, strategic delays in the launch of successful innovations can sometimes help incumbents maintain their market power. The overall result is a market with less innovation and competition, which may explain the struggles faced by entering innovative products like those from Apex Ski Boots.

Our argument hinges on the endogeneity of the customer search. Customers do not automatically become aware of innovative products. In fact, the more groundbreaking an innovation is, the less likely mainstream customers will be aware of something similar, increasing the difficulty of discovery [98]. The process requires incurring search costs to learn about new offerings, where customers weigh the risk of not finding desirable innovations against the costs of searching.

The key insight we present in this paper is that these search costs and discovery challenges play a crucial role in how incumbent (in)action can stifle innovation. When incumbents fail to provide innovative products, it can trigger a decline in overall customer search, making it harder for future innovations to be discovered and adopted. This insight has two implications for firms. First, it disincentivizes potential competitors from innovating. Second, it creates entry barriers, even for entrants with groundbreaking innovations, as customers are not actively searching. Consequently, incumbents may prefer to delay or withhold innovations, leading to prolonged periods of stagnation. These dynamics create a self-reinforcing cycle: firms are deterred from innovating due to reduced customer search, and customers search less because they expect fewer innovations.

To formalize these dynamics, we construct a two-period stylized model of innovation featuring an incumbent and an entrant. In the model, customers must search to discover any innovative products. Customers do not know the firms' true innovation speeds, but they can form inferences based on their search outcomes. A key finding is that total customer search declines in the second period if customers fail to discover innovative products in the first period. This decline occurs because customers rationally infer that innovation is likely to be slow, reducing their perceived likelihood of finding new products in the future.

As customer search declines, it becomes more difficult for new firms to enter the market, as there is a greater risk that customers will not discover their innovations. This creates a higher profitability threshold for firms considering launching innovative products in the second period, allowing incumbents to maintain market power even without further innovation.

(a) "cloud computing" (b) "microprocessor"

Figure 3.1: Google Trends for certain keywords

This prediction is consistent with the patterns observed in the Google Trends data. Figure 1 shows the search trends for keywords related to innovation in various domains on Google. There is an association between the time elapsed since the last disruptive innovation and reduced customer search. For example, customers search for new innovative breakthroughs in the domain of cloud computing after the launch of a disruptive innovation. However, after years of searching for new breakthroughs without major discoveries, we would expect search interest in this domain to decrease over time, which aligns with the trend depicted in Figure 3.1a. During this decline, emerging disruptive innovations in this domain might struggle to gain sufficient attention to achieve profitability, preventing these products from entering the consumer market.

Our model reveals a surprising result: incumbents can benefit from strategically delaying the launch of innovations. The incumbent may strategically withhold a successful innovation if the entrant has not yet succeeded in innovating. Intuitively, by doing so, the incumbent manipulates customer expectations, making innovation appear more infrequent than it actually is. This deceptive strategy discourages some customers from continuing to search, as they perceive a lower likelihood of discovering an innovative product.

In this equilibrium, even though customers are aware that the incumbent might be hoarding innovations, it is rational for them not to search again in the second period. This reduces total customer search in the second period and hinders the discovery of competing innovations. By delaying launching its innovations, the incumbent discourages the entrant from innovating. As a result, the incumbent maintains market power.

The incumbent makes this strategic decision when the potential risk of losing market share to an entrant with an innovative product outweighs the expected gains from launching its own innovation. Specifically, this strategic withholding of innovation is more likely to occur when entrant has a high probability of successfully developing a competing, innovative product. It is also more likely to occur when the incremental markup from selling the new product is not sufficiently higher than the markup from selling its existing basic product.

This paper makes several contributions to our understanding of the interplay between customer search, innovation, and competition. First, we highlight how customer search costs and the difficulty of discovering new products can stifle innovation. Second, we show that a lack of innovation by incumbents can create entry barriers, challenging the conventional wisdom that incumbent inaction facilitates competition. Third, we demonstrate that incumbents may strategically delay launching innovations, as this can reduce customers' propensity to search. Taken together, these findings may help to explain why many markets remain dominated by incumbent firms and experience prolonged periods of infrequent innovation. Our model provides a tractable framework for examining these dynamics and delivers novel, testable predictions.

The remainder of this paper is structured as follows. Section 2 reviews the literature. Section 3 describes the setup of the stylized model. Section 4 characterizes the evolution of customer search behavior over time. Section 5 analyzes the firms' strategies and characterizes the corresponding equilibria, discussing the implications of these equilibria. Finally, Section 6 explores potential extensions of the model and concludes the paper.

3.2 Literature Review

This paper contributes to the literature on innovation, customer search, strategic entry deterrence, and the interplay between them. We begin by briefly reviewing the literature on the economics of innovation.

Innovation has been widely regarded as the driving force for economic growth [99]. It can be converted into products with higher quality and generate more surplus; innovative disruptions from entrants can also motivate incumbent firms to innovate more in order to cope with intensifying competition, thereby increasing the overall level of innovation and ultimately driving economic growth [100], [101].

Conventional wisdom suggests that when established firms fail to innovate and adapt to new innovations, they become vulnerable, creating opportunities for new entrants to capitalize on these innovations and disrupt the market [98], [102]–[104]. [105] shows that even well-resourced incumbents can sometimes be outperformed by new entrants. In our paper, we use a game-theory model to propose a counterargument to this conventional wisdom. Our key insight is that, due to the endogeneity of customer search, the incumbent firm can strategically withhold innovation to hinder innovation progress, reducing competition and maintaining market power without innovating.

Our model builds on the premise that customers search to discover new products and expand their consideration sets. We argue that searching for awareness is particularly relevant in the context of innovation and new product launch. Previous work has shown that innovation is hard (and getting harder) and infrequent, making it challenging for customers to stay aware of the latest progress [106]–[109], especially for disruptive innovations [98]. Note that our paper focuses solely on customer learning; firm learning (its abilities) in new product launches [110] is beyond the scope of this paper.

While the search literature in economics and marketing primarily focuses on acquiring aware but uncertain information, such as heterogeneous matching values [111], [112] and prices [113], a subset of research emphasizes the discovery or awareness function of search. This includes both theoretical [114], [115] and empirical studies [116]–[118]. A branch of literature also studies the costly formation of the consideration set in customers' decision-making process [119]–[124], which is related to the consideration set expansion setting in this paper.

Classic marketing research examines how awareness affects innovation adoption in the new product diffusion literature [125]–[128]. Our paper differs from the diffusion literature by considering how the endogeneity of customer search stifles innovation through a self-reinforcing mechanism. Specifically, the perceived low innovation leads to lower search in the future, suppressing future innovations. Moreover, firms can leverage the endogenous search to further manipulate search and innovation in the future.

Shifting our focus from the customer side to the firm side, we summarize the related literature on how firms can gain an advantage by strategically manipulating customer search. When increased customer search, particularly informational search, benefits the firm, it can encourage search through two primary channels: reducing search costs or increasing search incentives. To reduce search costs, firms can offer alternative informational channels, such as advertising [117], [129], or improve search efficiency [118]. Alternatively, firms can increase customers' incentives to search by strategically reducing the informativeness of their advertising content [130], [131]. In our paper, although it may seem that the incumbent would prefer more customer search to increase product awareness, we show that the incumbent can sometimes benefit by reducing search and lowering awareness.

In our paper, firms can achieve search reduction through their (in)action, such as withholding successful innovation. The literature on how firms' (in)action can sometimes improve market performance is particularly relevant to this paper. For example, [132] show that when customers use the amount of marketing as a signal for product quality, a firm may choose to de-market its product to improve overall sales. Other examples in this domain include withholding brand advertising [133] and simplifying product line design by reducing the number of products [122].

The idea of revealing the partial truth as a manipulative signal is also related to the literature on information design. Existing work shows that the informationally superior party can design a signal (or an experiment) that mixes misleading information with true information, changing the decisions of the informationally inferior party [130], [131], [134], [135].

This paper contributes to the strategic entry deterrence literature by identifying a novel deterrence strategy. In their seminal work, [136] describe conditions under which incumbents hope to deter or accommodate entry, and show that these conditions depend upon whether firms' actions are strategic complements or substitutes. The literature offers various strategic handles to deter entry, such as prices and output levels [137], product durability [138], and overinvesting in advertising to raise awareness and build customer relationships [139], [140].

When advertising raises customers' valuation for entrants' products, incumbents may deter entry by lowering awareness of alternatives or creating market confusion. They can achieve this by decreasing advertising expenditure [141], which can sometimes reduce the effectiveness of entrants' marketing efforts [116]. Related to our paper, a few studies explore how firms manipulate endogenous customer search to deter entry, including increasing customer search cost [142] and switching cost [143]. In contrast, the incumbent in our paper directly leverages its private information about innovation speed and the signaling role of new product launches. Instead of changing costs, it can disincentivize customers from searching by withholding innovation and suppressing a signal that innovation occurs frequently.

In the context of innovation and new product launches, the literature explores how the incumbent may deter entry by strengthening its own innovation ability [100], [144] and innovating more. For example, by committing to a high level of innovation, the incumbent can signal its ability to compete aggressively [145].Firms sometimes invest in innovation solely to prevent competition [146], [147]. Moreover, institutions, such as intellectual property rights and patents, are set up to protect innovation [148]–[150]. Despite the intention to reward and motivate more innovations, these policies may actually deter later innovations [151]. Contrary to the literature, instead of innovating more to deter entry, this paper contributes to the literature by offering a deterrence strategy of innovating less. This also contributes to the literature exploring mechanisms for suppressing investment in innovation, including [152] and [153], both of which focus on firms optimizing short-term returns instead of long-term gains.

3.3 A Simple Model

3.3.1 Model Setup

We consider a two-period model with customers and firms. The two periods are denoted by $t \in \{1, 2\}$. In the model, suppose there is a continuum unit of customers with mass 1. There is one strategic firm (incumbent I) in the market. There is also a potential entrant (E).

For the two firms, I and E, suppose firm I always offers a basic product in each period. It can also engage in developing and offering new products. Firm E can only opt in to offer new products. The basic product is offered at a fixed price b, equal to the customer's outside option, and the marginal cost of producing a basic product is assumed to be 0. The new products will be offered at price $b + p$. We assume that both firms are price takers for now.

In the benchmark model, we assume that innovations are independent processes and are not controlled by the firm. If any innovation happens, the firm can choose to incur a fixed cost K, $K > 0$, in each period it tries to launch an innovative product on the market. If the product gets launched, the marginal cost of producing it is normalized to 0. If I ends up not launching any new product in a period, it can still sell the basic product at price b. For ease of exposition, we will denote the basic product and the innovative products (offered by firms I and E) as B, N_I, and N_E, respectively.

We denote the realization of innovations in Period t by h_t, $h_t \in \{00, 10, 01, 11\}$. The innovation can happen with different frequencies. We assume that the innovation frequency is governed by the state of the world, given by $\theta \in \{L, H\}$. In any period, an innovation happens with probability $\alpha_\theta \in \{\alpha_L, \alpha_H\}$. To simplify the game, we assume $\alpha_L = 0$. We further simplify the baseline model by assuming that the innovation always happens for firm I in $t = 1$. In addition, if a firm has a successful innovation in $t = 1$, the success will carry over to $t = 2$ as well.

The state will be determined prior to the beginning of the game and stays invariant throughout the timeline and for both firms. Firms are informed of the true state of the world because they are deeply involved and monitor the underlying innovation process. Customers, however, do not observe the state and have a (common) uniform prior belief of the state distributions, given by $r_L = r_H = 1/2$.

Based on the above description of the innovation process, the probability distribution of having a successful innovation is given by the tree in Figure 3.2.

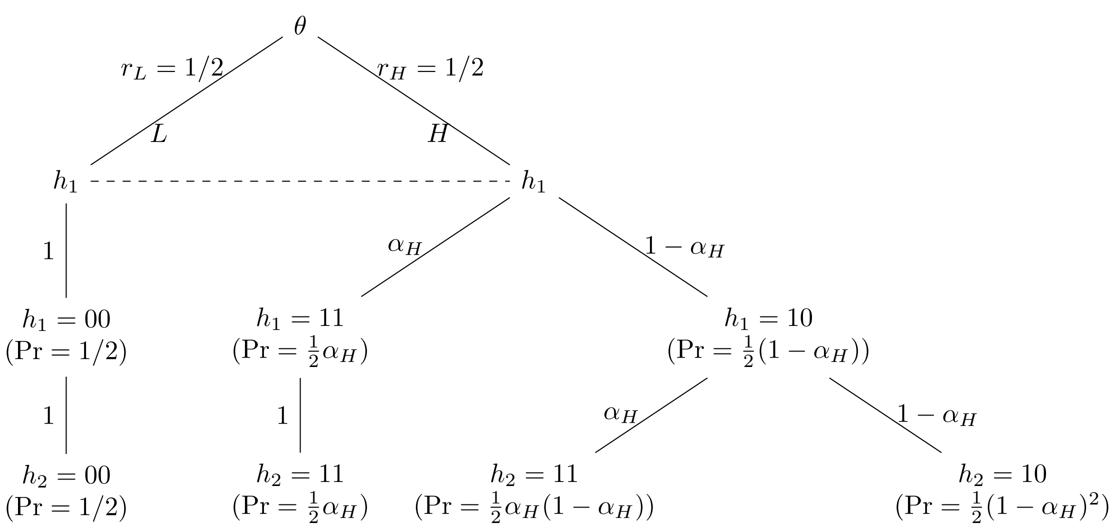


Figure 3.2: The innovation process for firms I and E

We now discuss the decision problem that customers face. Suppose customers have a unit demand and want to buy one and only one product in *each period*. Customers know the (homogeneous) value of the basic product, which is v_b. Customers also know their incremental value for the new product, v_N, $v_N \sim U[0, \bar{v}]$. Firms know both customer values, v_b and v_N.

We impose tie-breaking conditions on customers' decisions. If the innovative and basic products offer the same utility, a customer purchases the innovative product. If the entrant and incumbent both offer identical products, a randomly selected set of half of the customers buys the incumbent's product, and the other half purchases the entrant's product. In a later version of the model, we may use horizontal differentiation to model these decisions.

Customers initially do not know the existence of any new products on the market and need to incur a search cost of s, $s > 0$, to become aware. If a customer incurs the search cost and there are new innovative products offered on the market *in that period*, the customer will become aware of these innovative products and their prices. Otherwise, the customer only knows about the basic product. We denote the set of innovative products that the customer observes in Period t as m_t, where $m_t \in \{00, 10, 01, 11, \varnothing\}$, and $m_t = \varnothing$ means the customer did not search in period t. Customers (who searched) only observe the launch decisions (m_t). Firm I can observe both the realization and the launch decisions (h_t and m_t), and we use H_t to represent the firm's information set; $H_1 = \{h_1\}, H_2 = \{\theta, h_1, m_1, h_2\}$.

If a customer searches and learns the existence of a given innovative product, say, firm I's innovative product, in Period 1, she will not need to search again to know the existence of firm I's innovative product in Period 2. She will have to search again if she wants to find *other* innovative products (such as firm E's product). We impose a regularity condition that the incremental for the innovative product v_N is high enough to justify customer search cost s, $v_N > s$.

3.3.2 Timeline

We consider a two-period game with the following sequence of moves:

Period 1 Innovations happen with the given probability α_θ.

Step 1.1 Firm I privately observes the state of the world θ and the realization of innovations (for both firms) $h_1 \in \{00, 10, 11\}$.

Step 1.2 If any innovation happens, firm I can then decide whether to pay the fixed cost K to launch a new product $m_1 \in \{00, 10, 01, 11\}$.

Step 1.3 Customers observe (v_b, b). Based on their prior beliefs, they decide whether to search for any new product offerings at cost s. If they search, they learn the information m_1, and their prices.

Step 1.4 After observing their choice sets, customers make purchase decisions (buy the basic product; buy the new product if offered and preferred), and the profit is then realized for the firm.

Period 2 Innovations happen with the given probability α_θ. As mentioned before, the successful innovations in $t = 1$ will reprise the success in $t = 2$.

Step 2.1 Firms privately observes the realization of any innovations $h_2 \in \{00, 10, 11\}$.

Step 2.2 If any innovation happens, firm I can then decide whether to pay the fixed cost K to launch a new product $m_2 \in \{00, 10, 01, 11\}$.

Step 2.3 Customers observe (v_b, b). If a customer searched in $t = 1$, she has observed m_1. They then update their beliefs of both the state θ and the offering probabilities of each new product type. Then, they decide whether to search for any new product offerings.

Step 2.4 After observing their choice sets, customers make purchase decisions (buy the basic product; buy the new product if offered and preferred), and the profit is then realized for the firm.

In Period 2, when a customer decides whether to search, we assume that all customers can choose to search with cost s in $t = 2$. Moreover, if a customer searched in $t = 1$ and became aware of an innovative product offered by any firm, she remains aware of (*just*) that product in $t = 2$ (even if the product is not offered in $t = 2$). To become aware of other products, she needs to search again with the same cost s. If the customer did not search in $t = 1$, her belief is updated in $t = 2$ even if she did not search, because she knows that the probability of the product offering changes in $t = 2$ in equilibrium.

We look for pure-strategy perfect Bayesian equilibrium (PBE) in this game.

3.4 Customer's Search Dynamics

In our model, we assume customers need to engage in costly search to gain awareness of any innovative products offered in the market. As the starting point of our main argument, we first investigate how customers form their search decisions and how these decisions are influenced by the innovation frequency. Specifically, we are interested in the change in a customer's search decision in Period 2 compared to her decision in Period 1.

3.4.1 Customer's Search Dynamics in a Minimal Model

Since the focus of this section is on customers' decision making, to better demonstrate the dynamics of customer search, we first consider a simpler version of the model described in Section 3.3. We make two simplifications in this version of the model. First, we simplify the market structure and assume that there is only one firm in the market. This firm always offers a basic product, B, and engages in innovation with a success probability of $\alpha_\theta \in \{\alpha_L, \alpha_H\}$, $\theta \in \{L, H\}$, $\alpha_L = 0$. Again, we assume only the firm can observe θ and the realization of innovation in each period, h_t, and customers only observe the launched products, m_t. The revised tree structure is given in Figure 3.3.

The second simplification is that we assume the firm always launches the innovative product when it has a success in innovation; that is, the customer's information in each period is consistent with the realization of innovation in that period, $m_t = h_t$. We retain all other assumptions. Next, we turn to the customer's decision making problem.

Suppose a customer has discovered the innovative product, offering her a utility of $v_b - b + v_N - p$ compared to the utility from the basic product, $v_b - b$. She then chooses to buy the innovative product as long as $v_N \geq p$.

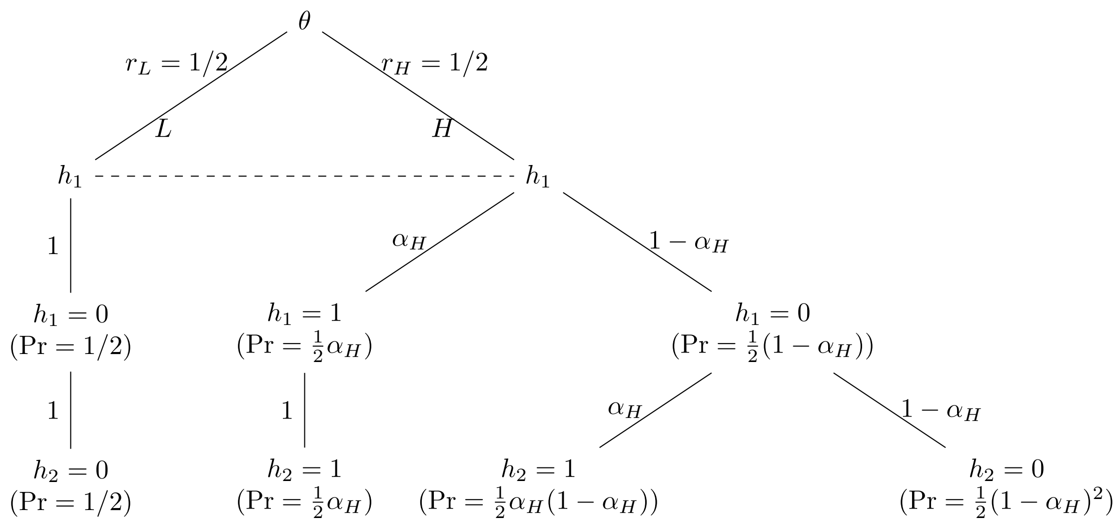


Figure 3.3: Innovation process with one innovating firm

However, if the customer has not yet discovered the innovative product, she needs to decide whether to search for it. If she searches the market, there is a possibility that she could find a new product offering that she was not aware of prior to that point, and this new product might yield a higher surplus for her. On the other hand, if she does not search, she does not incur the search cost s, but she will have to make a product purchase decision based on her current knowledge of the set of products she has already discovered, which could be just the basic product. Her decision (to search) in Period t is thus driven by her expected utility from searching, $\mathbb{E}U_t^s$, versus her utility from not searching, $\mathbb{E}U_t^0$. She searches in Period t only if $\mathbb{E}U_t^s \geq \mathbb{E}U_t^0$.

Customers' expected utilities are estimated based on their beliefs. We first establish the beliefs that a customer holds in each period. Based on her information set, the customer needs to know (1) the probability of the state being a high state, H, $\Pr(H|m_1)$, and (2) the probability of finding an innovative product, $\Pr(m_2 = 1|m_1)$. In Period 1, the customer believes $r_H = r_L = 1/2$, consistent with her prior belief, and the probability of discovering an innovative product is $1/2\alpha_H$.

In Period 2, a customer's posterior belief is updated based on her observation $m_1 \in \{0, 1, \varnothing\}$. If she observes $m_1 = 0$, she updates her beliefs about the state to $\Pr(H|m_1 = 0) =$

$\frac{1-\alpha_H}{2-\alpha_H}$ and $\Pr(L|m_1 = 0) = \frac{1}{2-\alpha_H}$, and she thinks she will find a new product with probability $\frac{\alpha_H(1-\alpha_H)}{2-\alpha_H}$ in Period 2. If she observes $m_1 = 1$, she knows $\theta = H$, and she will find the innovative product again with probability 1 in Period 2. Notably, after not observing any innovative product offered in Period 1, $m_1 = 0$, the customer thinks the state is more likely to be a low state, L, compared to her prior belief, as one can easily show $\frac{1-\alpha_H}{2-\alpha_H} < \frac{1}{2}$. If she did not search in Period 1 ($m_1 = \varnothing$), she still believes $r_H = r_L = 1/2$, and the probability of discovering at least one innovative product is $\frac{1}{2}\alpha_H + \alpha_H(1-\alpha_H) = \frac{1}{2}\alpha_H(2-\alpha_H)$. We remark here that although the customer's belief about the state remains the same, the probability distribution of discovering an innovation changes. This is because if the firm succeeds in innovating, it retains that innovation in Period 2, as indicated in Figure 3.3.

We solve the customer's search decision problem by backward induction. We start by describing the customer's problem in Period 2, which contingent on the customer's observation $m_1 \in \{0, 1, \varnothing\}$, as her beliefs are different.

If $m_1 = 1$, the customer will not search again in Period 2 because she has already discovered an innovative product. Her period utility is $\mathbb{E}U_2^0 = v_b - b + v_N - p$. If $m_1 = 0$, based on the above beliefs, the customer's period utility from searching is given by

$$\mathbb{E}U_2^s = v_b - b - s + \frac{\alpha_H(1-\alpha_H)}{2-\alpha_H}(v_N - p)^+,$$

against her utility from not searching, which is $\mathbb{E}U_2^0 = v_b - b$.

If $m_1 = \varnothing$, the customer will search only if her utility from searching $\mathbb{E}U_2^s = v_b - b - s + \frac{1}{2}\alpha_H(2-\alpha_H)(v_N - p)^+$ is greater than not searching $\mathbb{E}U_2^0 = v_b - b$.

Similar to Period 2, the customer's search decision in Period 1 reflects the utilities she can earn through searching (or not searching) in Period 1. Additionally, as the customer is forward-looking, she considers the impact of her Period 1 decision on her expected outcomes in Period 2 when making decisions in Period 1. Specifically, her utility in Period 2 may influence her decision in Period 1 through two possible channels. First, she is aware of the possibility that she may need to search in Period 2, which generates option value in Period 1. Second, if she discovers an innovative product after searching in Period 1, she reserves the option to purchase that product in Period 2 without having to search again. Therefore, when making a search decision in Period 1, the customer considers her expected utilities from both periods. The customer's utilities from searching and not searching are given by, respectively,

$$\begin{aligned}
\mathbb{E}U_1^s =& \underbrace{(v_b - b)}_{t=1:B} \underbrace{-s + \frac{1}{2}\alpha_H(v_N - p)^+}_{t=1:N} + \underbrace{(v_b - b)}_{t=2:B} + \underbrace{\frac{1}{2}\alpha_H(v_N - p)^+}_{t=2:N,m_1=1} \\
&+ \underbrace{\frac{1}{2}(2-\alpha_H)\left[-s + \frac{\alpha_H(1-\alpha_H)}{2-\alpha_H}(v_N - p)^+\right]^+}_{t=2:\text{ option value from } N,m_1=0,m_2=1} \\
\mathbb{E}U_1^0 =& \underbrace{(v_b - b)}_{t=1:B} + \underbrace{(v_b - b)}_{t=2:B} + \underbrace{\left[-s + \frac{1}{2}\alpha_H(2-\alpha_H)(v_N - p)^+\right]^+}_{t=2:\text{ option value from } N}
\end{aligned} \tag{3.1}$$

We solve the above decision-making problems in both periods, and summarize the customers who search ("searchers") and do not search ("non-searchers") in each period in the below result.

Result 4 (Customer's Search Dynamics with A Single Firm) *Customers will search for innovative products when their incremental valuation for the product v_N is sufficiently high. Specifically:*

1. *In Period 1, the searchers are $v_N \in \left[p + \frac{s}{\alpha_H}, \bar{v}\right]$, and the non-searchers are $v_N \in \left[0, p + \frac{s}{\alpha_H}\right)$.*

2. *In Period 2, the searchers in Period 2 are $v_N \in \left[p + \frac{(2-\alpha_H)s}{\alpha_H(1-\alpha_H)}, \bar{v}\right]$ upon observing $m_1 = 0$, or $v_N \in \varnothing$ upon observing $m_1 \in \{1, \varnothing\}$. The remaining customers are non-searchers.*

3. *The total amount of customer search declines in Period 2, regardless of the Period 1 observation, m_1.*

Proof. *See Appendix C.1.*

The search declines in Period 2 for two reasons. First, when $m_1 = 0$, meaning the customer searched in Period 1 but did not discover any innovative product, if her valuation for the innovative product is not sufficiently high to support a second search, she would accept the fact and not search again in Period 2. This is also the primary search decline mechanism we are interested in in this paper—failed product discovery leads to lower search incentive in the future. In Section 3.5, we focus on the firm's strategy leveraging this search dynamics.

Second, customers who have already discovered an innovative product stop searching in Period 2. Notably, the second reason does not drive the result, and we will show this in Subsection 3.4.2, where we consider a market with horizontally differentiated products.

We also discuss the demand faced by the firm in each period. In Period 1, all searchers buy the innovative product if $m_1 = 1$; non-searchers and searchers observing $m_1 = 0$ buy the basic product. In Period 2, when $m_1 = 1$, the Period 1 searchers continue buying the innovative product; when $m_1 = 0$ and $m_2 = 1$, the Period 2 searchers discover and buy the innovative product while Period 2 non-searchers buy the basic product; Period 1 non-searchers again do not search in Period 2 and buy the basic product.

Therefore, if the innovative product was launched in Period 1, although the total search declines, the demand for this product does not decline. This is because all customers who have bought the innovative product in Period 1 will buy it again in Period 2. However, if the innovative product is not launched until Period 2, its demand comes only from the Period 2 searchers who observe $m_1 = 0$ and $m_2 = 1$, and this demand is lower than the demand in the case of launching in Period 1.

3.4.2 Horizontal Innovative Products Competition

We now return to the setting where there are two firms, firm I and firm E, on the market, and they both engage in innovation. The market and information structures thus revert to the model in Section 3.3, and the innovation process goes back to the original tree structure in Figure 3.2. We still keep the assumption (from Subsection 3.4.1) that firms always launch the new products whenever they succeed in innovation.

In this horizontal differentiation model, we relax one assumption in Section 3.3 on customers' preferences for the two innovative products offered by firms I and E. Recall that we denote the innovative products from firms I and E as N_I and N_E, respectively. A customer's net utilities from buying these two products are given by $v_b + v + \tau v_N - (b+p)$ and $v_b + v + \tau(1 - v_N) - (b+p)$, respectively, where $v_N \sim U[0,1]$, $v > \tau > 0$, and $v > p$. [1] That is, we assume the two products are substitutes for customers, with v representing the common base incremental utility for the innovative product, v_N representing the heterogeneous horizontal taste, and τ controlling the degree of substitution. Put differently, using the analogy in a Hotelling line model, we assume firms E and I are located at the two ends of a linear city of length τ, and customers are uniformly distributed on the line, with a transportation cost equal to τ. We add a tie-breaking condition that when the customer's utility is the same for products N_I and N_E, she chooses N_E, and when her utility is the same for products B and any innovative product N, she chooses the innovative product. We maintain all other assumptions.

The primary objective of this subsection is to reproduce Result 4 and show that customer search declines in Period 2 when there are two firms competing on innovation and have a chance to offer horizontally differentiated products. In the rest of this subsection, we solve the customers' decision-making problem in this model.

Customer's Purchase Decisions

We start by summarizing the conditions under which the customer chooses the innovative product offered by firm I, N_I. We use $\succ$ to represent the preference order; for example

[1] To justify this assumption, note that we assume $v > \tau$, and the complete information Hotelling pricing for the case with only two new products is $p = \tau$. Prices over τ will be hard to sustain in a pricing competition equilibrium; one can also show that, in a complete information game with any one firm offering the new product, the firm would also not charge a new product premium such that $p > v$.

$N_I \succ N_E$ means the customer prefers N_I over N_E. The more interesting case is when the customer observes $m_t = 11$, she can choose one product from $\{B, N_I, N_E\}$, and she makes this decision by solving the following problem:

$$\max\{v_b - b,\, v_b + v + \tau v_N - (b+p),\, v_b + v + \tau(1 - v_N) - (b+p)\}$$

We summarize her choice rules in Result 5.

Result 5 (Customer's Preference with Differentiated Products) *In the model with horizontally differentiated innovative products, customer's preference and choice conditions are given as follows:*

1. *When the customer observes $m_t = 10$, everyone buys N_I.*

2. *When the customer observes $m_t = 01$, everyone buys N_E.*

3. *When the customer observes $m_t = 11$, $v_N \in [0, 1/2]$ customers will choose to buy N_E, and $v_N \in (1/2, 1]$ customers will choose to buy N_I.*

Proof. See Appendix C.1.

We remark here that, although it appears that customers do not buy the basic product in the presence of at least one innovative product, offering the basic product still has option value for firm I, as there is no cost to do so, and it gives the firm more strategic flexibility. Especially when $m_t = 00$, firm I can still earn revenue by selling the basic product.

Customer Beliefs and Search Strategies

Since customers' search decisions are based on their beliefs in each period, we now derive the customers' belief update process. In Period 1, a customer holds her prior belief on the state, given by $r_H = 1/2$. Her beliefs on the probabilities of each observation outcome $m_1 \in \{11, 10, 00\}$ are consistent with the tree in Figure 3.2. She thinks the probabilities of discovering N_I and N_E in Period 1 are thus $\frac{1}{2}$ and $\frac{1}{2}\alpha_H$. In addition, if she observes the off-equilibrium outcome $m_1 = 01$ by any chance, she can infer that the true realization is $h_1 = 11$ and $\theta = H$.

In Period 2, customers' beliefs are contingent on observing $m_1 \in \{11, 10, 00, \varnothing\}$. For example, when $m_1 = 10$, the customer thinks $\Pr(H|m_1 = 10) = 1$, and $h_2 = 11$ with probability α_H, so she believes she will be able to discover the product N_E with probability α_H. Note that the customer is only motivated to search again in Period 2 after discovering $m_1 = 10$ if she prefers N_E over N_I.

Similar to the method we used in Subsection 3.4.1, we use backward induction to solve customer's search decision-making problem.

Search Decisions in Period 2 We again start by describing the customer's decision problem in Period 2. The cases of $m_1 \in \{00, 11\}$ are easier to solve: When $m_1 = 00$, the customer knows that $\theta = L$ and $m_2 = h_2 = 00$, so she will not search, and her utility is $\mathbb{E}U_2^0 = v_b - b$, and everyone buys B. When $m_1 = 11$, the customer knows that $\theta = H$ and $m_2 = h_2 = 11$, so she will not search, as she has already discovered both innovative products, and her utility is

$$\mathbb{E}U_2^0 = v_b - b + \underbrace{v + \tau v_N - p}_{t=2:N_I} + \underbrace{\tau(1 - 2v_N)^+}_{t_2:|N_E - N_I|}.$$

Note that the utility difference between the two products (the extra utility a customer gets from N_E net of the utility from N_I) is $\tau(1 - 2v_N)^+$. We can easily determine that customers with $v_N \in [0, \frac{1}{2}]$ will choose to buy N_E, and customers with $v_N \in (\frac{1}{2}, 1]$ will choose to buy N_I.

When $m_1 = 10$, as discussed above, her search is only meaningful when she prefers N_E

and discover it in Period 2. Her period utility is given by

$$\begin{aligned}
\mathbb{E}U_2^s =& v_b - b + v + \tau v_N - p - s + \alpha_H \cdot \tau(1 - 2v_N)^+ \\
\mathbb{E}U_2^0 =& v_b - b + v + \tau v_N - p
\end{aligned}$$

When $m_1 = \varnothing$, her period utility is thus given by $\mathbb{E}U_2^0 = v_b - b$ when she does not search; when she searches,

$$\mathbb{E}U_2^s = v_b - b - s + \frac{1}{2}(v + \tau v_N - p) + \frac{1}{2}\alpha_H(2 - \alpha_H) \cdot \tau(1 - 2v_N)^+$$

Search Decisions in Period 1 When making decisions in Period 1, the customer need to look forward to Period 2, and consider the Period 2 utility as well, which is conditional on her Period 1 observation m_1. Looking forward to Period 2, she believes that, if she does not search in Period 1, she still reserves the option to search in Period 2, and the possibilities to discover N_I and N_E are $\frac{1}{2}$ and $\frac{1}{2}\alpha_H(2 - \alpha_H)$. The customer's search decision problem in Period 1 is given by

$$\begin{aligned}
\mathbb{E}U_1^s =& (v_b - b) - s + \frac{1}{2}(v + \tau v_N - p) + \frac{1}{2}\alpha_H \cdot \tau(1 - 2v_N)^+ \\
& + (v_b - b) + \frac{1}{2}(v + \tau v_N - p) + \frac{1}{2}\alpha_H \cdot \tau(1 - 2v_N)^+ \\
& + \frac{1}{2}(1 - \alpha_H)\left[-s + \alpha_H \cdot \tau(1 - 2v_N)^+\right]^+ \\
\mathbb{E}U_1^0 =& (v_b - b) + (v_b - b) + \left[-s + \frac{1}{2}(v + \tau v_N - p) + \frac{1}{2}\alpha_H(2 - \alpha_H) \cdot \tau(1 - 2v_N)^+\right]^+
\end{aligned} \tag{3.2}$$

We can then solve the inequality $\mathbb{E}U_1^s \geq \mathbb{E}U_1^0$ to find customers who search in Period 1. Before demonstrating the results, we first define a few extra quantities for ease of exposition.

$$
\begin{aligned}
\hat{v}_1 \equiv& \frac{1}{2}\left(1 - \frac{s}{\alpha_H \tau}\right) \\
\hat{v}_2 \equiv& \frac{1}{\tau}(p - v + 2s), \quad \hat{v}_3 \equiv \frac{1}{\tau}(p - v + s), \quad \hat{v}_4(\alpha_H) \equiv \frac{1}{\tau}\left[p - v + \frac{1}{2}(3 - \alpha_H)s\right] \\
y(\alpha_H) \equiv& \alpha_H^2 - 3\alpha_H + 1
\end{aligned}
$$

It follows immediately that $\hat{v}_3 < \hat{v}_4(\alpha_H) < \hat{v}_2$. We also impose the following regularity conditions:

Assumption 1 (Regularity Conditions) *We assume:*

1. $s - \alpha_H \tau < 0$, *which implies* $\hat{v}_1 \in \left(0, \frac{1}{2}\right)$;
2. $\hat{v}_3 > \frac{1}{2}$, *which implies* $\hat{v}_2 > \hat{v}_4(\alpha_H) > \frac{1}{2}$;
3. $\alpha_H \leq 1 - \sqrt{1 - \hat{v}_2}$ *when* $\hat{v}_2 \in (\frac{1}{2}, 1)$;
4. $2\hat{v}_4(\alpha_H) + y(\alpha_H) < 1$ *when* $\alpha_H > \frac{1}{2}(3 - \sqrt{5})$.

Under Assumption 1, we can summarize the search dynamics in this problem as Result 6.

Result 6 (Decline in Search with Product Competition) *A customer will search only if her utility from at least one innovative product is sufficiently high. The search decisions in each period are:*

1. *When* $v_N > 1/2$, *customers have the preference* $N_I \succ N_E \succ B$.
   (a) *In Period 1, searchers are* $v_N \in [\hat{v}_3, 1]$. *The non-searchers are* $v_N \in (\frac{1}{2}, \hat{v}_3)$.
   (b) *In Period 2: neither the Period 1 searchers nor the Period 1 non-searchers will not search. Notably, the result remains even without the two regularity conditions on* $\hat{v}_2$ *and* $\hat{v}_3$.

2. *When* $v_N \le 1/2$*, , customers have the preference* $N_E \succ N_I \succ B$*.*

   (a) *In Period 1:*

      i. *For customers who would not search in Period 2 after observing* $m_1 = 10$*, the searchers are* $v_N \in \left(\bar{v}_1, \frac{\hat{v}_3-\alpha_H}{1-2\alpha_H}\right]$ *when* $\alpha_H > \frac{1}{2}$ *and* $\hat{v}_1 + \hat{v}_3 < 1$*. There are no searchers otherwise.*

      ii. *For customers who would search again in Period 2 after observing* $m_1 = 10$ *(Period 2 searchers), the searchers are* $v_N \in [0, \tilde{v}]$ *with sufficiently high success rate* α_H *and sufficiently small* $\tilde{v}$*.* [2]

   (b) *In Period 2: the non-searchers will not search. The searchers will search if* $m_1 = 10$ *and* $v_N \in [0, \tilde{v}]$ *with the same sufficiently high success rate* α_H *and sufficiently small* $\tilde{v}$ *in 2(a)(ii).*

3. *The total customer search declines in Period 2 regardless of the observation in Period 1.*

Proof. *See Appendix C.1.*

We add a few remarks on the search decisions by customers who prefer N_E. First, if the probability of firm E successfully innovating (α_H) is sufficiently small, these customers perceive a very low likelihood of discovering N_E through searching; thus they would choose not to search in Period 1. Second, a Period 1 searcher will search again in Period 2, if the searcher perceives a sufficiently high probability of discovering N_E, and her valuation for N_E is sufficiently large. The latter condition implies that many Period 1 searchers will not search again in Period 2, even after observing an outcome $m_1 = 10$ that indicates a high success rate for N_E. This is because they are satisfied with the already discovered N_I and are thus unwilling to incur the search cost again.

In this section, we have shown that customer search will strictly decline if firms fail to innovate in Period 1. This positive association between the amount of innovation *right now*

[2]Formally, these thresholds are $\alpha_H > \frac{1}{2}(3 - \sqrt{5})$ and $\tilde{v} \equiv \min\left\{\hat{v}_1, \frac{\hat{v}_4}{y} + \frac{1}{2}(1 - \frac{1}{y})\right\}$.

and the amount of customer search *in the future* is robust. Next, we show that the incumbent can use this rational choice of customers to improve its market performance.

3.5 Initial Equilibrium Analysis

In this section, we return to the exact setup in Section 3.3 and solve the two-period game. Specifically, compared to the model in Subsection 3.4.2, we allow the incumbent, firm I, to make decisions on whether to launch the innovative product in each period, given there is a success. To simplify the customer preference, we use the utility functions proposed in Section 3.3: we assume customers' incremental value for the new product is v_N, $v_N \sim U[0, \bar{v}]$. When it comes to tie-breaking, we again assume half of the customers prefer firm I's product over firm E's product.

It follows immediately that, after discovering an innovative product, the customer chooses to buy it whenever $v_N \geq p$, and she buys from firm I with probability $\frac{1}{2}$ if both firms offer innovative products.

We look for the existence of pure-strategy perfect Bayesian Nash equilibria in this game. Firm I's strategy is to decide whether to launch an innovative product, conditional on its observation of the state θ, realization histories h_1, h_2, and product offering histories m_1, m_2. We denote this strategy as

$$m_t | H_t, \theta \in \{00, 10, 01, 11\}.$$

Recall that we use H_t to represent the firm's information set; $H_1 = \{h_1\}, H_2 = \{h_1, m_1, h_2\}$. Note that firm E's product offering is determined by the innovation process, not firm I. We suppose customers hold the rational belief that the firm plays the equilibrium strategy and will not introduce an extra notation here. Note that, even if customers do not observe the actual outcome of h_t, they still hold some beliefs about what firms would do in these situations. Suppose we only allow for a one-time deviation, and if firm I deviates twice, there will be a sufficiently high penalty. This is a standard assumption in the literature.

3.5.1 Truth-telling Equilibrium with Two Firms Innovating

We first show the existence of a separating equilibrium where firm I is always truth-telling: it always launches the innovative product whenever it becomes available. Since firm I is truth-telling, on the equilibrium path, customers' observation in Period 1, m_1^*, always reflects the realizations of innovation $h_1, m_1^* = h_1$. In the separating equilibrium, customers think firm I plays $m_t^* = h_t$, and they update their beliefs accordingly.

Customer's optimal decisions Similar to Subsection 3.4.2, we can derive the customers' beliefs in each period. By backward induction, we start our analysis from the customer's search decision in Period 2, which is again contingent on the customer's observation m_1, as her beliefs are different. If $m_1 = 00$, the customer will not search in Period 2, because she knows there will be no innovation in Period 2. If $m_1 \in \{10, 11\}$, the customer will also not search again in Period 2, because she has already discovered an innovative product. If $m_1 = \varnothing$, the customer will search only if her utility from searching $\mathbb{E}U_2^s = v_b - b - s + \frac{1}{2}(v_N - p)^+$ is greater than not searching $\mathbb{E}U_2^0 = v_b - b$.

In Period 1, the customer's utility is given by

$$\begin{aligned} \mathbb{E}U_1^s &= v_b - b - s + \frac{1}{2}(v_N - p)^+ + (v_b - b) + \frac{1}{2}(v_N - p)^+ \\ \mathbb{E}U_1^0 &= v_b - b + (v_b - b) + \left[-s + \frac{1}{2}(v_N - p)^+\right]^+ \end{aligned} \tag{3.3}$$

The solutions to customers' decision-making problems in Period 2, discussed above, and Period 1, defined in Equation 3.3, are the optimal decisions given firm I plays the equilibrium strategy.

Firm I's optimal decisions We denote these demand functions for products B, N_I, N_E as D_b, D_N^I and D_N^E. These demand functions will follow immediately from the purchase rule, after solving the customers' search decision problem. For firm I, the per-period profit when not launching the new product is $\Pi_t^I(m_t = 0x) = b \cdot D_b$, where x represents any number between 0 and 1. When it launches the new product, the per-period profit is

$$\Pi_t^I(m_t = 1x) = b \cdot D_b + (b + p)D_N^I - K \tag{3.4}$$

The total profit in Period 1 is thus

$$T\Pi_1^I = \Pi_1^I(m_1) + \mathbb{E}_{h_2}\left[\Pi_2^I(m_2)\middle|\, m_1, h_2\right]. \tag{3.5}$$

By definition, we can then show that this separating equilibrium exists by showing there do not exist profitable off-equilibrium path deviations for firm I under certain conditions. This truth-telling equilibrium is summarized in Result 7.

Result 7 (Truth-telling Product Launch) *A separating equilibrium exists where firm I always launches the innovative product when it succeeds in innovation, in any period. Customers whose value satisfies $v_N \in [p + s, \bar{v}]$ will search in Period 1, and no customers search in Period 2, so the total search declines in Period 2 regardless of the search outcome in Period 1. This equilibrium can be sustained if the fixed cost (K) or the search cost (s) is sufficiently small.*

Proof. *See Appendix C.1.*

Result 7 confirms that the decline in customer search still holds in an equilibrium with a strategic firm. For firm I, when the innovation is profitable enough, implied by the sufficiently small fixed cost or search cost, firm I is willing to always launch the new product whenever possible. This launching strategy allows firm I to earn a higher markup without much risk of giving up a high market share to the entrant. In addition, when firm I fails to innovate initially, this result shows that customers will not search again in Period 2, and all customers will still buy the basic product from firm I, creating an entry barrier for entrants and protecting the incumbent from competition.

3.5.2 Strategic Innovation Withholding

As discussed in the Introduction, we are interested in showing that there exists a semi-separating equilibrium, where firm I only reveals partial truth. Specifically, under the threat of a potential entrant, firm I may find it beneficial to be a monopoly of the basic product market and keep its market share. This could give firm I a higher profit compared to competing with firm E in the innovative product market, earning a higher markup with a smaller market share.

We are interested in showing that the following firm I's (semi-separating) strategy m_t^* constitutes a PBE, as summarized in Table 3.1. Note that firm I would not "tell" the truth when it observes that firm E fails innovation in Period 1, $h_1 = 10$, and when firm E fails innovation in Period 1 but succeeds in Period 2, $h_1 = 10$, $h_2 = 11$.

$h_1 \mid \theta$	Period 1 Strategy	$h_2 \mid \theta; h_1$	Period 2 Strategy
$00 \mid L$	no launch decision	$00 \mid L; 00$	no launch decision
$11 \mid H$	launch	$11 \mid H; 11$	launch
$10 \mid H$	not launch	$10 \mid H; 10$	launch
		$11 \mid H; 10$	not launch

Table 3.1: Firm I's innovation withholding strategy

Customers' search strategies Suppose the customer holds the rational belief that firm I plays this equilibrium strategy m_t^*. We can then derive her beliefs in each period. In Period 1, the customer's prior belief on the state is given by $r_H = r_L = 1/2$, and she can discover at least one innovative product in $t = 2$ with probability $\frac{1}{2}\alpha_H$.

In Period 2, customers might observe $m_1 \in \{00, 11, \varnothing, 10, 01\}$. Note that a customer might also observe a deviation outcome $\{10, 01\}$ from the off-the-equilibrium path. For each of these observations, a customer updates her beliefs according to Bayes' rule. The interesting case is when she observes $m_1 = 00$, and she updates her posterior state distribution to $\Pr(H|m_1 = 00) = \frac{1-\alpha_H}{2-\alpha_H}$, $\Pr(L|m_1 = 00) = \frac{1}{2-\alpha_H}$. The customer then believes she can discover at least one innovative product in $t = 2$ with probability $\frac{1-\alpha_H}{2-\alpha_H}$.

We now solve the customer's search decision problem. In Period 2, a customer's search decision is contingent on her observation in Period 1, so we need to look at each observation separately. The interesting case is when she observes $m_1 = 00$. The customer believes she can discover at least one product in Period 2 with probability $\frac{1-\alpha_H}{2-\alpha_H}$, and the customer's period utilities from searching and not searching are given by

$$\mathbb{E}U_2^s = v_b - b - s + \frac{1-\alpha_H}{2-\alpha_H}(v_N - p)^+$$
$$\mathbb{E}U_2^0 = v_b - b$$

We detail the rest of this discussion in the proof of Result 8.

In Period 1, the customer's period utilities from searching and not searching are given by

$$\begin{aligned}\mathbb{E}U_1^s =& \underbrace{(v_b - b)}_{t=1:\, B} \underbrace{-s + \frac{1}{2}\alpha_H(v_N - p)^+}_{t=1:N} + \underbrace{(v_b - b)}_{t=2:\, B} + \underbrace{\frac{1}{2}\alpha_H(v_N - p)^+}_{t=2:N,\, m_1=11} \\ & + \underbrace{\frac{1}{2}(2-\alpha_H)\cdot\left[-s + \frac{1-\alpha_H}{2-\alpha_H}(v_N - p)^+\right]^+}_{t=2:\text{option value from } N,\, m_1=00,\, m_2\neq 00} \\ \mathbb{E}U_1^0 =& \underbrace{(v_b - b)}_{t=1:\, B} + \underbrace{(v_b - b)}_{t=2:\, B} + \underbrace{\left[-s + \frac{1}{2}(v_N - p)^+\right]^+}_{t=2:\,\text{option value from } N}\end{aligned} \tag{3.6}$$

Similar to Subsection 3.5.1, the solutions to customers' decision-making problems are the optimal decisions given that firm I plays the equilibrium strategy.

Firm I's optimal decisions Turning to the firm side, for firm I, its demand functions will follow immediately from the purchase rule after solving the customers' search decision problem. Firm I's per-period profit functions and the total profit functions when the firm is truth-telling are given by the same Equations 3.4 and 3.5 as in Subsection 3.5.1. Notably, when firm I plays the withholding strategy, $h_1 = 10$, $m_1^* = 00$, the total profit is

$$T\Pi_1^I = \Pi_1^I(m_1^* = 00, h_1 = 10) + \alpha_H \cdot \Pi_2^I(m_2^* = 01|h_2 = 11) + (1-\alpha_H) \cdot \Pi_2^I(m_2^* = 10|h_2 = 10). \tag{3.7}$$

We can then summarize the existence of the semi-separating equilibrium, as well as its strategic withholding outcome, in Result 8.

Result 8 (Strategic Withholding) *There exist some threshold values, such that for any innovation with sufficiently low markup, measured by p/b, and any sufficiently high innovation success rate $\alpha_H > \frac{1}{2}$ that also does not exceed an upper limit,* [3] *there exists a semi-separating equilibrium where firm I plays strategies in Table 3.1. In this equilibrium, customers' search rules are given by:*

1. *In Period 1, the searchers are $v_N \in \left[p + \frac{s}{\alpha_H}, \bar{v}\right]$, and the non-searchers are $v_N \in \left[0, p + \frac{s}{\alpha_H}\right)$.*
2. *In Period 2: when $m_1^* = 11$, $v_N \in \varnothing$; when $m_1^* = 00$, $v_N \in \left[p + \frac{(2-\alpha_H)s}{1-\alpha_H}, \bar{v}\right]$; when $m_1' = 10$, $v_N \in \left[p + \frac{s}{\alpha_H}, \bar{v}\right]$; when $m_1' = 01$, $v_N \in \varnothing$. The non-searchers from Period 1 will also not search in Period 2. In equilibrium, the total customer search declines in Period 2, regardless of the customer observation in Period 1, m_1.*

Proof. See Appendix C.1.

Intuitively, firm I has the incentive to be a monopoly in the basic product market and withhold innovation when it either perceives a high risk of intensified competition or is not satisfied with the incremental markup from selling its innovative product. The risk of intensified competition is high when the probability that firm E succeeds in innovation, α_H, is sufficiently high. When both firms offer innovative products and a majority of customers are aware of both products, firm I loses customers to whom it could have sold the basic product, provided they are not aware of firm E's product. This effect will be amplified if the

[3] More formally, there exist $\tilde{\alpha} > \hat{\alpha} > \frac{1}{2}$ and $\hat{q} > 3$, such that for any $\alpha_H \in (\hat{\alpha}, \tilde{\alpha}]$ and $q \equiv b/p > \hat{q}$.

incremental markup from the innovative product, measured by p/b, is sufficiently small. In addition, the entrant's success rate cannot be too high; otherwise, customers are confident that they can discover its new product through searching, and the withholding strategy will not work.

In order to make the launch withholding a credible signal, firm I can only withhold launching its new product before firm E has an innovation, that is, when $h_1 = 10$. Otherwise, when $h_1 = 11$, even if firm I withholds its innovation, not only will all Period 1 searchers buy from firm E, causing firm I to be in an inferior position in both market share and markup, but these searchers are also able to infer that the state is H and continue to buy from firm E in Period 2. Notably, firm I's choice to withhold launching a new product again in Period 2 with histories $h_1 = 10$, $m_1^* = 00$, and $h_2 = 11$ is not a strategic decision but a simple optimization: in this specification, the demand for its innovative product is too small, so firm I is better off foregoing the innovation entirely.

The partial truth nature of firm I's strategy makes it rational for customers not to search again. Upon observing $m_1 = 00$, although customers know that the true realization of innovation could be $h_1 = 10$ and firm I might be withholding its innovation, the likelihood that the true state is L still exists, and it increases customers' risk of not discovering any new product in Period 2. Therefore, it is rational for customers with insufficiently high valuation v_N to refrain from searching again in Period 2.

3.6 Possible Extensions and Concluding Remarks

Perceptual monopoly could lead to perpetual monopoly. This paper explores the dynamics between firms' innovation processes and customer search behavior, investigating scenarios where reduced innovation benefits the incumbent firm. While conventional wisdom suggests that an incumbent's failure to innovate facilitates competition by encouraging new entrants, our two-period game theory model argues that the opposite can sometimes be true: the incumbent may prefer to forgo innovation to discourage customer search and deter entry.

These findings have important implications for both customers and firms. For customers, the model indicates that their search behavior can significantly influence the market's innovation pace; if they quickly become discouraged and stop searching for new products, they may inadvertently slow down overall innovation. For firms, the model highlights the strategic considerations in launching innovations. Incumbents may have an incentive to delay introducing new products to maintain market power, even if it means withholding valuable innovations from customers.

Finally, we propose several extensions that could enrich the initial analysis in our baseline model. These extensions aim to enhance the robustness of our findings by relaxing assumptions and exploring additional strategic dimensions.

We expect that the ability to set prices would allow firms to influence customer search and the profitability of innovation. In equilibrium, the incumbent may use prices, especially the price of its basic product, as a signal to deter entry or discourage customer search. The information contained in the signal will depend upon when customers can observe prices (before or after search), and when the incumbent can observe the entrant's prices (before or after setting its own price). In this extension, we would also anticipate that the entrant endogenously sets the price of its product.

We could extend our model to account for other types of customer heteroegeneity while also considering firms that offer differentiated innovative products. This extended model would help to explain why the incumbent firm continues to offer the baseline product, even when innovative products are on the market. Notice that the absence of the baseline product could itself act as a signal that innovative products are available.

We could extend the model to allow the entrant to strategically decide whether to invest in innovation, and/or whether to launch an innovative product. When the entrant expects few customers to engage in search, the entrant may forgo innovation. This would allow the model to describe outcomes in which an incumbent's decision to forgo innovation or withhold launch deters innovation by its competitors.

Comparing innovations of varying sizes, such as infrequent disruptive innovations versus continuous incremental innovations, could shed additional light on the relationship between innovation and customer search. The possibility of discovering a disruptive innovation that is large in magnitude could increase the motivation for customers to engage in search. However, if larger innovations take longer to develop, the long interval between innovations could reduce customers' motivation to search.

Appendix A

Appendix for Chapter 1

A.1 Additional Notations and Formulations

Table A.1: Table of notations used in Appendices A.1 through A.2

t	Subscript identifying time: past period (t_0), this period (t_1), next period (t_2)
i	Subscript identifying customers
$\mathcal{N}_i$	The set of customers in period t
n_t	The number of customers in period t
$A_{i,t}$	Action implemented for customer i in period t
Y_i	The single period profit earned (outcome) from customer i
$\mathbf{X}_{i,t}$	Vector of covariates for customer i in period t
$\mathbf{x}$	A realized value of covariates
$\mathcal{A}, \mathcal{X}, \mathcal{Y}$	The value spaces of actions, covariates and outcomes
$\mathbf{H}_t$	History of data observed at the start of period t
$g \in \mathcal{G}$	Cluster g and cluster space
$\mathbf{X}_t^g$	Period t Cluster g customers' covariates
$\mathbf{E}^g$	Externality metric
r	The response function that models the individual profit
$\widetilde{r}$	Artificial next wave profit function when assigned experimental action
	We use $\sim$ to represent synthesized terms
π_t	Period t assignment policy
p_t	The exploit policy given $\mathbf{H}_t$
$\widetilde{\mu}_t$	The maximum of all period t posterior means
IC_t	Information cost function given $\mathbf{H}_t$
IV_t	Information value function given $\mathbf{H}_t$
EE_t	Explore-exploit function given $\mathbf{H}_t$

We use upper case for variables, lower case for data values, italics to denote both singulars and functions, bold roman to denote vectors and matrices, and script to identify sets.

Additional functions and formulations

For ease of mathematical exposition, we introduce some additional functions and formulations in the Appendices. Suppose the adaptive batch targeting problem is of T periods, and the focal period (this period as in the main text) is any period t. The underlying response function is $r(\mathbf{x}, a)$, and we use r_t^* to represent the best response with respect to all actions in the response function under the information (history) of period t:

$$r_t^*(\mathbf{x}) \equiv r_t^*(\mathbf{x}; \mathbf{H}_t) \equiv \max_{a \in \mathcal{A}} \mathbb{E}[r(\mathbf{x}, a)|\mathbf{H}_t].$$

We further denote the posterior means of $r(\mathbf{x}, a)$ based on period t history $\mathbf{H}_t$ as:

$$\begin{aligned}
\mu_t(\mathbf{x}, a) &= \mathbb{E}[r(\mathbf{x}, a) \mid \mathbf{H}_t] \\
\mu_t^*(\mathbf{x}) &= \mathbb{E}[r(\mathbf{x}, p_t(\mathbf{x})) \mid \mathbf{H}_t].
\end{aligned}$$

Similarly, we can denote the posterior means of $r_{t+1}(\mathbf{x}, a)$ and $r_{t+1}^*(\mathbf{x})$ based on period t history $\mathbf{H}_t$:

$$\begin{aligned}
\mu_{t+1}(\mathbf{x}, a) &= \mathbb{E}[\mathbb{E}_{t+1}[r(\mathbf{x}, a) \mid \mathbf{H}_{t+1}] \mid \mathbf{H}_t] \\
\mu_{t+1}^*(\mathbf{x}) &= \mathbb{E}[r_{t+1}^*(\mathbf{x}) \mid \mathbf{H}_t].
\end{aligned}$$

Moreover, we use $p_t(\mathbf{x}) \equiv \max_{a \in \mathcal{A}} \mu_t(\mathbf{x}, a)$ to represent the best action implied by the exploit policy p_t in period t.

Going back to the time subscript $t \in \{t_0, t_1, t_2\}$, the IC-function and the IV-function of period t_1 now can be simplified with these additional notations as

$$\begin{aligned}
IC_{t_1}(\mathbf{X}_i, a) \equiv& \mathbb{E}\left[r_{t_1}^*(\mathbf{X}_i) - r(\mathbf{X}_i, A_i) \,\middle|\, \mathbf{H}_{t_1}, A_i = a\right] \\
IV_{t_1}(\mathbf{X}_i, a|\mathbf{A}_{-i}) \equiv& \sum_{j \in \mathcal{N}_{t_2}} \mathbb{E}\left[r_{t_2}^*\left(\mathbf{X}_j; \mathbf{H}_{t_2}(A_i; \mathbf{A}_{-i})\right) \,\middle|\, \mathbf{H}_{t_1}, A_i = a, \mathbf{A}_{-i}\right] \\
&- \sum_{j \in \mathcal{N}_{t_2}} \mathbb{E}\left[r_{t_2}^*\left(\mathbf{X}_j; \mathbf{H}_{t_2}(p_{t_1}(\mathbf{X}_i); \mathbf{A}_{-i})\right) \,\middle|\, \mathbf{H}_{t_1}, p_{t_1}(\mathbf{X}_i), \mathbf{A}_{-i}\right]
\end{aligned} \tag{A.1}$$

Formulation of the simulated estimator

We start with an interim assignment policy π as the policy being evaluated. It is also implemented for all other t_1 customers. We draw a batch of t_1's outcome samples $\widetilde{\mathbf{Y}}_{t_1}$ from the generative model (based on the posterior distribution of the response function r), defined in Equation 1.10, and repeat this procedure K times, to construct K sets of artificial outcomes, effectively synthesizing the posterior distribution of $\mathbf{Y}_{t_1}$. These draws are based on inputs that are at the t_1 covariate values $\mathbf{X}_{t_1}$, assignment to the focal customer being the proposed action $A_{i,t_1} = a$, and also assignments to other t_1 customers assigned by the interim policy π.

For each set of artificial outcomes $\widetilde{\mathbf{Y}}_{t_1}^{(k)}$, we then construct an artificial history, $\widetilde{\mathbf{H}}_{t_2}^{(k)} = \{\mathbf{H}_{t_1}, \mathbf{A}_{t_1}, \widetilde{\mathbf{Y}}_{t_1}^{(k)}\}$, combining the observed history in t_1 and these artificial samples. Based on each set of on these artificial histories $\widetilde{\mathbf{H}}_{t_2}^{(k)}$, we re-learn an artificial response function, $\widetilde{r}^{(k)}(\mathbf{x}', a')$, as if we are in t_2. An associated t_2 artificial exploit policy $\widetilde{p}_{t_2}^{(k)}\left(\mathbf{x}' \,\middle|\, \widetilde{\mathbf{H}}_{t_2}^{(k)}\right)$ is also derived from optimizing this artificial response function, such that $\widetilde{p}_{t_2}^{(k)}\left(\mathbf{x}' \,\middle|\, \widetilde{\mathbf{H}}_{t_2}^{(k)}\right) = \max_{a \in \mathcal{A}} \mathbb{E}\left[\widetilde{r}^{(k)}(\mathbf{x}', a') \,\middle|\, \widetilde{\mathbf{H}}_{t_2}^{(k)}\right]$.

We repeat the above process, and use a simulated estimator to compute the expectations at $\widetilde{\mathbf{H}}_{t_2}$. To evaluate the expected value of information, we can use the clustering approach to characterize the information externality in this algorithm. Since the clustering approach allows us to focus the information externalities within a cluster, it is more efficient to evaluate with GP inference (the complexity of the kernel matrices increases with the number of samples). That is, for customer i in Cluster g, we can evaluate her EE-function restricting attention to all Cluster g customers. The simulated estimator for the IV-function for Period

t customer i in Cluster $g \in \mathcal{G}$ is given by:

$$\widetilde{IV}_{t_1}(\mathbf{X}_i, a|\mathbf{A}^g_{-i}) \equiv \sum_{j \in \mathcal{N}_{t_2}} \mathbb{E}\left[\widetilde{\mathbb{E}}_{t_2}\left[r^g(\mathbf{X}_j, p_{t_2}(\mathbf{X}_j)) \,\middle|\, \mathbf{H}_{t_2}(A_i; \mathbf{A}_{-i})\right] \,\middle|\, \mathbf{H}_{t_1}, A_i = a, \mathbf{A}_{-i}\right]$$
$$- \sum_{j \in \mathcal{N}_{t_2}} \mathbb{E}\left[\widetilde{\mathbb{E}}_{t_2}\left[r^g(\mathbf{X}_j, p_{t_2}(\mathbf{X}_j)) \,\middle|\, \mathbf{H}_{t_2}(p_{t_1}(\mathbf{X}_i); \mathbf{A}_{-i})\right] \,\middle|\, \mathbf{H}_{t_1}, p_{t_1}(\mathbf{X}_i), \mathbf{A}_{-i}\right] \tag{A.2}$$

The terms under tilde ($\sim$) are synthesized (based on simulation) quantities, using information known in period t_1. We use $\mathbb{E}[\cdot]$ to denote empirical expectation, to differentiate from the quantities computed based on actually observed histories in period t_1, $\mathbf{H}_{t_1}$. We use the notations $\mathbf{X}^g_{-i}$ and $\mathbf{A}^g_{-i}$ to denote the covariates and the action assignments for other Cluster g customers.

Alternatively, when the customers in a cluster are homogeneous, we can evaluate using the externality metric $\mathbf{E}^g$ to replace full specifying $\mathbf{A}_{-i}$ in Equation A.2.

In the algorithm, we only need to estimate the first term, because the second term is invariant to the optimization problem. Suppose we have K artificial trajectories, and the kth artificial trajectory gives an artificial response function $\widetilde{r}^{(k)}(\mathbf{x}', a')$, which gives the posterior means $\boldsymbol{\mu}^{(k)}(\mathbf{X}_{t_2}, a)$. The simulated estimator of the inner expectation is just the simple average of all maximal posterior means obtained from the K artificial trajectories: for a customer j in t_2, this is given by $\frac{1}{K}\sum_{k \leq K}\sum_{j \in \mathcal{N}_{t_2}} \mu^{(k)}_{j,t_2}(\mathbf{X}_{j,t_2}, a)$.

Quantification of the information externality and its relation to existing knowledge

We can directly quantify the exact amount of information externality between a focal customer and any same batch customers, based on our value of information framework. We focus on the expected value of information (as defined in IV-function), as the IC-function treats customers separably. For i in t_1, the firm observes $\mathbf{H}_{t_1}$, and can quantify the total information (ν) and the incremental information (δ) from i (assigned action A_i), when considering a subset of customers $\mathcal{J} \subset \mathcal{N}_{t_1}$ (assigned actions $\mathbf{A}_{\mathcal{J}}$; for example, when i belongs to cluster g, we consider all other customers in g) with (omit $\mathbf{H}_{t_2}$ notations for ease of exposition):

$$\begin{aligned}
\nu_i(A_i) \equiv & \sum_{j \in \mathcal{N}_{t_2}} \mathbb{E}\left[\mathbb{E}_{t_2}\left[r\left(\mathbf{X}_j, p_{t_2}(\mathbf{X}_j)\right) \,\middle|\, A_i\right] \,\middle|\, \mathbf{H}_{t_1}\right] \\
& - \sum_{j \in \mathcal{N}_{t_2}} \mathbb{E}\left[\mathbb{E}_{t_2}\left[r\left(\mathbf{X}_j, p_{t_2}(\mathbf{X}_j)\right) \,\middle|\, \emptyset\right] \,\middle|\, \mathbf{H}_{t_1}\right] \\
\delta_{i,\mathcal{J}}(A_i, \mathbf{A}_{\mathcal{J}}) \equiv & \sum_{j \in \mathcal{N}_{t_2}} \mathbb{E}\left[\mathbb{E}_{t_2}\left[r\left(\mathbf{X}_j, p_{t_2}(\mathbf{X}_j)\right) \,\middle|\, A_i; \mathbf{A}_{\mathcal{J}}\right] \,\middle|\, \mathbf{H}_{t_1}\right] \\
& - \sum_{j \in \mathcal{N}_{t_2}} \mathbb{E}\left[\mathbb{E}_{t_2}\left[r\left(\mathbf{X}_j, p_{t_2}(\mathbf{X}_j)\right) \,\middle|\, \mathbf{A}_{\mathcal{J}}\right] \,\middle|\, \mathbf{H}_{t_1}\right].
\end{aligned}$$

And the information externalities of $\mathcal{J}$ on i is the difference between (a) information value that could have been contributed by i when treated separably $\nu_i(A_i)$ and (b) the information value that is contributed by i when $\mathbf{A}_{\mathcal{J}}$ present $\delta_{i,\mathcal{J}}(A_i, \mathbf{A}_{\mathcal{J}})$. That is,

$$\text{Information Externality}(A_i, \mathbf{A}_{\mathcal{J}}) = \nu_i(A_i) - \delta_{i,\mathcal{J}}(A_i, \mathbf{A}_{\mathcal{J}}).$$

We offer a further observation of the benefit from considering the information externalities with an example, by considering a batch of two customers 1 and 2 in t_1. Consider the values of information a customer contributes:

$$
\begin{aligned}
\nu_1(A_1) \equiv & \sum_{j \in \mathcal{N}_{t_2}} \mathbb{E}\left[\mathbb{E}_{t_2}\left[r\left(\mathbf{X}_j, p_{t_2}(\mathbf{X}_j)\right) \,\middle|\, A_1\right] \,\middle|\, \mathbf{H}_{t_1}\right] \\
& - \sum_{j \in \mathcal{N}_{t_2}} \mathbb{E}\left[\mathbb{E}_{t_2}\left[r\left(\mathbf{X}_j, p_{t_2}(\mathbf{X}_j)\right) \,\middle|\, \emptyset\right] \,\middle|\, \mathbf{H}_{t_1}\right] \\
\nu_1(A_2) \equiv & \sum_{j \in \mathcal{N}_{t_2}} \mathbb{E}\left[\mathbb{E}_{t_2}\left[r\left(\mathbf{X}_j, p_{t_2}(\mathbf{X}_j)\right) \,\middle|\, A_2\right] \,\middle|\, \mathbf{H}_{t_1}\right] \\
& - \sum_{j \in \mathcal{N}_{t_2}} \mathbb{E}\left[\mathbb{E}_{t_2}\left[r\left(\mathbf{X}_j, p_{t_2}(\mathbf{X}_j)\right) \,\middle|\, \emptyset\right] \,\middle|\, \mathbf{H}_{t_1}\right] \\
\nu_{12}(A_1, A_2) \equiv & \sum_{j \in \mathcal{N}_{t_2}} \mathbb{E}\left[\mathbb{E}_{t_2}\left[r\left(\mathbf{X}_j, p_{t_2}(\mathbf{X}_j)\right) \,\middle|\, A_1, A_2\right] \,\middle|\, \mathbf{H}_{t_1}\right] \\
& - \sum_{j \in \mathcal{N}_{t_2}} \mathbb{E}\left[\mathbb{E}_{t_2}\left[r\left(\mathbf{X}_j, p_{t_2}(\mathbf{X}_j)\right) \,\middle|\, \emptyset\right] \,\middle|\, \mathbf{H}_{t_1}\right]
\end{aligned}
$$

Observe $\mathbb{E}\left[\mathbb{E}_{t_2}\left[r\left(\mathbf{X}_j, p_{t_2}(\mathbf{X}_j)\right) \,\middle|\, \emptyset\right] \,\middle|\, \mathbf{H}_{t_1}\right] = \mathbb{E}\left[r\left(\mathbf{X}_j, p_{t_1}(\mathbf{X}_j)\right) \,\middle|\, \mathbf{H}_{t_1}\right] = \mu^*_{t_1}\left(\mathbf{X}_j, p_{t_1}(\mathbf{X}_j)\right)$. This implies that the expected t_2 profit evaluated at p_{t_2} degenerates to it evaluated at the p_{t_1} without new information. To interpret the above measurements, for example, $\nu_1(A_1)$ is the total information value Customer 1 with A_1 has, and it is the difference between the expected t_2 profit with and without Customer 1 assigned A_1. $\delta_{12}(A_1, A_2) = \nu_{12}(A_1, A_2) - \nu_2(A_2)$ is the incremental information of Customer 1, considering Customer 2 (with A_2). Individual optimization only considers $\nu_1(A_1)$ when constructing Customer 1's value, which should be $\delta_{12}(A_1, A_2)$.

The information externality is measured by $\nu_1(A_1) - \delta_{12}(A_1, A_2)$. When it is close to zero, it means i is not heavily influenced by the information externality, and the individual optimization (only optimize $\nu_1(A_1)$) is still able to recover the optimal assignment. As it increases until passing a threshold, the information externality grows, and the assignment by the individual optimization becomes suboptimal, and we have to do joint optimization.

There are many factors causing the $\nu_1(A_1) - \delta_{12}(A_1, A_2)$ to (weakly) increase. For example, it increases in the proximity between 1 and 2.

We also consider the impact of existing knowledge on how to optimally trade off exploration and exploitation. When $\nu_1(A_1)$ is large (the firm has limited knowledge), as $\delta_{12}(A_1, A_2)$ decreases in comparison to $\nu_1(A_1)$, the information externality increases, and it is more likely that individual optimization overestimates the value of exploring A_1 with Customer 1, and *over-explores* A_1. In contrast, when $\nu_1(A_1)$ is small (the firm already has intermediate level of knowledge), as $\delta_{12}(A_1, A_2)$ decreases in comparison to $\nu_1(A_1)$, it becomes more likely that individual optimization overestimates the value of exploiting A_1 with Customer 1, and *over-exploits* (*under-explores*) A_1.

The informativeness of the information from each customer affects how the firm should explore and exploit considering information in a similar way to the existing knowledge. When each customer is very informative for learning the response function, $\nu_1(A_1)$ is large, as $\nu_1(A_1) - \delta_{12}(A_1, A_2)$ increases, it is more likely that individual optimization over-explores A_1. When $\nu_1(A_1)$ is small (Customer 1 is not so informative), as $\delta_{12}(A_1, A_2)$ decreases in comparison to $\nu_1(A_1)$ (Customer 2 becomes more informative), it becomes more likely that individual optimization under-explores A_1.

Illustrative example of the clustering approach

Consider a firm that has two possible marketing actions `mail`, `not mail`, and five (discrete) covariate values ($\mathbf{x}^1$ through x^5). We further assume that customers can be clustered into five groups using these values, and the response function for customers in one cluster is independent of the response function for customers in the other clusters. This implies that the clusters are separable, and so there are no information externalities between them. For example, for a customer with covariate $\mathbf{x}^1$, the information the firm needs to exclude externalities from the other customers this period is the (interim) assignment vector for all of the $\mathbf{x}^1$ customers.

In addition, these smaller joint optimization problem in this discrete covariate example can be further simplified. Because the total count of $\mathbf{x}^1$ customers is known (and constant), we only need one parameter to represent the number of customers that receive action `mail`, and the number of these customers that receive `not mail`. This count can also be thought of as a state variable that represents every possible state of the information externalities between $\mathbf{x}^1$ customers. In particular, if two of the $\mathbf{x}^1$ customers receive action mail under the interim assignments, it does not matter which two customers they are. The joint optimization problem is reduced to optimizing conditional on this state variable.

Notice also from this example how the size of the action space and the covariate space affects the complexity of the problem. With three possible actions, we now need two state variables to represent the externalities. In the case of continuous and correlated covariate space, depending on how similar customers are in a cluster, we can choose to either solve the exact within-cluster combinatorial problem or the reduced state variable optimization problem.

A.2 Proofs of Main Results

We use the additional functions and formulations introduced in Appendix A.1 in this section.

Proof of Result 1 (Value function maximization)

The proof shows the joint maximization problem of the EE-function, given in Equation 1.5 is equivalent to the maximization of the value function, given in Equation 1.1. Start from Equation 1.5, we have

$$
\begin{aligned}
&\max_{a_i \in \mathcal{A}} \max_{\pi'} EE_{t_1}(\mathbf{X}_i, a_i | \mathbf{A}_{-i}; \mathbf{A}_{-i} \in \pi') \\
=&\max_{a_i \in \mathcal{A}} \max_{\pi'} IV_{t_1}(\mathbf{X}_i, a_i | \mathbf{A}_{-i}; \mathbf{A}_{-i} \in \pi') - IC_t(\mathbf{X}_i, a_i) \\
\propto&\max_{a_i \in \mathcal{A}} \max_{\pi'} EE_{t_1}(\mathbf{X}_i, a_i | \mathbf{A}_{-i}; \mathbf{A}_{-i} \in \pi') - \sum_{k \in \mathcal{N}_t} IC_{t_1}(\mathbf{X}_k, a_k)
\end{aligned}
$$

$$
\begin{aligned}
&= \max_{a_i\in\mathcal{A}} \max_{\pi'} \sum_{j\in\mathcal{N}_{t_2}} \mathbb{E}\left[r^*_{t_2}\left(\mathbf{X}_j; \mathbf{H}_{t_2}(a_i; \mathbf{A}_{-i})\right) \;\Big|\, \mathbf{H}_{t_1} \right] - \sum_{j\in\mathcal{N}_{t_2}} \mathbb{E}\left[r^*_{t_2}\left(\mathbf{X}_j; \mathbf{H}_{t_2}(p_{t_1}(\mathbf{X}_i); \mathbf{A}_{-i})\right) \;\Big|\, \mathbf{H}_{t_1} \right] \\
&\quad - \sum_{k\in\mathcal{N}_t} \mathbb{E}\left[r^*_{t_1}(\mathbf{X}_i) - r(\mathbf{X}_i, a_i) \,\Big|\, \mathbf{H}_{t_1} \right] \\
&\propto \max_{a_i\in\mathcal{A}} \max_{\pi'} \sum_{j\in\mathcal{N}_{t_2}} \mathbb{E}\left[r^*_{t_2}\left(\mathbf{X}_j; \mathbf{H}_{t_2}(a_i; \mathbf{A}_{-i})\right) \;\Big|\, \mathbf{H}_{t_1} \right] - \sum_{i\in\mathcal{N}_{t_1}} \mathbb{E}\left[r^*_{t_1}(\mathbf{X}_i) - r(\mathbf{X}_i, a_i) \,\Big|\, \mathbf{H}_{t_1} \right] \\
&\propto \max_{a_i\in\mathcal{A}} \max_{\pi'} \sum_{j\in\mathcal{N}_{t_2}} \mathbb{E}\left[r^*_{t_2}\left(\mathbf{X}_j; \mathbf{H}_{t_2}(a_i; \mathbf{A}_{-i})\right) \;\Big|\, \mathbf{H}_{t_1} \right] + \sum_{i\in\mathcal{N}_{t_1}} \mathbb{E}\left[r(\mathbf{X}_i, a_i) \,\Big|\, \mathbf{H}_{t_1} \right] \\
&= \max_{a_i\in\mathcal{A}} \max_{\pi'} \sum_{j\in\mathcal{N}_{t_2}} \mathbb{E}\left[\mathbb{E}_{t_2}\left[Y_j; p_{t_2} | a_i; \mathbf{A}_{-i} \in \pi'\right] \;\Big|\, \mathbf{H}_{t_1} \right] + \sum_{i\in\mathcal{N}_{t_1}} \mathbb{E}\left[Y_i; a_i, \mathbf{A}_{-i} \in \pi' \mid \mathbf{H}_{t_1}\right] \\
&= \max_{a_i\in\mathcal{A}} \max_{\pi'} V_{t_1}(a_i, \mathbf{A}_{-i}; \mathbf{A}_{-i} \in \pi') = V_{t_1}(\pi^*) \\
&\equiv \sum_{i\in\mathcal{N}_{t_1}} \mathbb{E}^{\pi^*}[Y_i | \mathbf{H}_{t_1}] + \sum_{j\in\mathcal{N}_{t_2}} \mathbb{E}^{\pi^*}[\mathbb{E}_{t_2}[Y_j; p_{t_2} \mid \mathbf{H}_{t_2}] | \mathbf{H}_{t_1}]
\end{aligned}
$$

The second equality is because the IC-functions are separable, and thus IC-functions of other customers from this period don't affect the joint optimization problem. The fourth equality is because the second term in the IV-function does not concern a_i, and $p_{t_1}(\mathbf{X}_i)$ is invariant to the joint optimization problem. Similarly, the fifth equality is because the first term in the IC-function only concerns $p_{t_1}(\mathbf{X}_i)$, which is invariant to the joint optimization problem. The sixth equality is because we assume Y_j is determined by an unbiased response function $r(\mathbf{X}_j, A_j)$, and in period t_2, the firm uses the exploit policy p_{t_2}; in period t_1, the firm uses $A_i = a_i$ and $\mathbf{A}_{-i} \in \pi'$.

■

Proof of Result 2 (Free from Selection)

Formally, in any given period, we learn the profit function r with experiment data from all previous periods. This result says that:

$$
\ell(\Theta) \equiv P(\mathbf{A}_{\le t}, \mathbf{Y}_{\le t}(\cdot) | \mathbf{X}, \Theta) = \prod_{s=1}^{t} P(\mathbf{A}_s | \mathbf{X}, \Theta) P(\mathbf{Y}_s(\cdot) | \mathbf{X}, \Theta),
$$

where we use Θ to denote the parameter set for function r.[1] *It means that the potential outcomes and assignments are independent, conditional on all the covariate values.*

We first discuss the roadmap. We prove this proposition in two steps, using the definition of conditional independence. In the first step, we show that, if firm only uses data from a single wave, the assignments and the outcomes are conditionally independent. That is, in period s,

$$P(\mathbf{A}_s, \mathbf{Y}_s(\cdot)|\mathbf{X}, \Theta) = P(\mathbf{A}_s|\mathbf{X}, \Theta)P(\mathbf{Y}_s(\cdot)|\mathbf{X}, \Theta). \tag{A.3}$$

In the second step, we show that, the assignments and the outcomes from each wave are conditionally independent. Specifically, we show the following result,

$$P(\mathbf{A}_{\leq t}, \mathbf{Y}_{\leq t}(\cdot)|\mathbf{X}, \Theta) \propto P(\mathbf{Y}_s(\cdot)|\mathbf{A}_s, \mathbf{X}, \Theta). \tag{A.4}$$

Then, from Equation A.3, $\mathbf{Y}_s(\cdot)$ and A_s are independent conditional on Θ and $\mathbf{X}_s$. We thus have $P(\mathbf{Y}_s(\cdot)|\mathbf{A}_s, \mathbf{X}, \Theta) = P(\mathbf{Y}_s(\cdot)|\mathbf{X}, \Theta)$. Finally, since $P(\mathbf{A}_{\leq t}, \mathbf{Y}_{\leq t}(\cdot)|\mathbf{X}, \Theta) \propto P(\mathbf{Y}_s(\cdot)|\mathbf{X}, \Theta)$, we combine it with Equation A.3 again, and the conditional independence in Equation A.3 is proved.

Step 1. Consider firm only uses period s data to learn the profit function r. The likelihood of assignments and outcomes, conditional on covariates, is then given by

$$P(\mathbf{A}_s, \mathbf{Y}_s(\cdot)|\mathbf{X}_s, \Theta) = P(\mathbf{Y}_s(\cdot)|\mathbf{X}, \Theta)P(\mathbf{A}_s|\mathbf{Y}_s(\cdot), \mathbf{X}_s, \Theta).$$

To show Equation refeq:ap-selection-decompose, it suffices to show $P(\mathbf{A}_s|\mathbf{Y}_s(\cdot), \mathbf{X}_s, \Theta) = P(\mathbf{A}_s|\mathbf{X}_s, \Theta)$. In our GP framework, Θ is the sufficient statistic for learning profit function, i.e., $r \equiv r_\Theta$. Notice that $\mathbf{A}_s$ is entirely determined by history at period s, i.e., $\mathbf{A}_s = f(\mathbf{X}_{<s}, \mathbf{A}_{<s}, \mathbf{Y}_{<s}, \mathbf{X}_s)$, and thus not directly on period s outcomes $\mathbf{Y}_s$. Therefore, it remains

[1] In Bayesian inference, "parameters" Θ are treated as random variables. In nonparametric Bayesian inference, the equivalent of "parameter set" is the function values (as random variables) evaluated at inputs. We denote the function values at period t inputs as $\mathbf{R} \equiv r(\mathbf{X}, \mathbf{A})$. That said, the reader can see $\Theta \equiv \mathbf{R}$.

to show that, conditional on Θ and $\mathbf{X}_s$, period s potential outcomes and outcomes from any wave prior to period s are independent, i.e., $\mathbf{Y}_s(\cdot) \perp \mathbf{Y}_{<s}|\Theta, \mathbf{X}_s$. This conditional independence holds, because Equation 1.8 implies the potential outcome is determined by

$$Y_{i,s}(a) = r_\Theta(\mathbf{X}_{i,s}, a) + \epsilon_{i,s}, \ \forall a \in \mathcal{A}, \tag{A.5}$$

and $\boldsymbol{\epsilon}_s$ and $\boldsymbol{\epsilon}_{<s}$ are independent by construction.

Step 2. We start from writing out the joint likelihood of all the action assignments and outcomes, conditional on covariates. It is given by

$$\begin{aligned} &P(\mathbf{A}_{\le t}, \mathbf{Y}_{\le t}(\cdot)|\mathbf{X}, \Theta) \equiv P(\mathbf{A}_1, \cdots, \mathbf{A}_t, \mathbf{Y}_1(\cdot), \cdots, \mathbf{Y}_t(\cdot)|\mathbf{X}, \Theta) \\ =&\prod_{s=1}^{t} P(\mathbf{A}_s, \mathbf{Y}_s(\cdot)|\mathbf{A}_1, \cdots, \mathbf{A}_{s-1}, \mathbf{Y}_1(\cdot), \cdots, \mathbf{Y}_{s-1}(\cdot), \mathbf{X}, \Theta) \\ =&\prod_{s=1}^{t} P(\mathbf{A}_s|\mathbf{A}_{<s}, Y_{<s}(\cdot), \mathbf{X}, \Theta) P(\mathbf{Y}_s(\cdot)|\mathbf{A}_s, \mathbf{A}_{<s}, Y_{<s}(\cdot), \mathbf{X}, \Theta). \end{aligned}$$

These equalities hold because of Bayes' rule. To further simply the above expression, first recall that $\mathbf{A}_s$ is entirely pinned down by history at period s, that is, $\mathbf{A}_s = f(\mathbf{X}_{<s}, \mathbf{A}_{<s}, \mathbf{Y}_{<s}, \mathbf{X}_s)$. Therefore, $P(\mathbf{A}_s|\mathbf{A}_{<s}, Y_{<s}(\cdot), \mathbf{X}, \Theta) = P(\mathbf{A}_s|\mathbf{A}_{<s}, Y_{<s}, \mathbf{X}, \Theta) = P(\mathbf{A}_s|\mathbf{A}_{<s}, Y_{<s}(\cdot), \mathbf{X})$, as this distribution has conditioned on the entire period s history, and thus does not rely on Θ.

For the second term, we know from Equation A.5 that $\mathbf{Y}_s(\cdot)$ does not depend on past assignments or outcomes. Hence, $P(\mathbf{Y}_s(\cdot)|\mathbf{A}_s, \mathbf{A}_{<s}, Y_{<s}(\cdot), \mathbf{X}, \Theta) = P(\mathbf{Y}_s(\cdot)|\mathbf{A}_s, \mathbf{X}, \Theta)$. Then,

$$P(\mathbf{A}_{\le t}, \mathbf{Y}_{\le t}(\cdot)|\mathbf{X}, \Theta) = \prod_{s=1}^{t} P(\mathbf{A}_s|\mathbf{A}_{<s}, Y_{<s}(\cdot), \mathbf{X}) P(\mathbf{Y}_s(\cdot)|\mathbf{A}_s, \mathbf{X}, \Theta) \propto \prod_{s=1}^{t} P(\mathbf{Y}_s(\cdot)|\mathbf{A}_s, \mathbf{X}, \Theta).$$

The last step holds, because $P(\mathbf{A}_s|\mathbf{A}_{<s}, Y_{<s}(\cdot), \mathbf{X})$ does not depend on Θ, and thus have no impact on the learning of the likelihood. We have now proved Step 2, and finished the proof.

$\blacksquare$

Proof of Result 3 (Convergence of evaluation algorithm)

The EE-function optimization algorithm converges to $EE_t(\cdot, a \mid \pi^*)$ *and* $\pi^*, \mathbf{A}^* \in \mathcal{S}(\pi^*)$, *such that* $A_i^* \in \max_{a\in\mathcal{A}} EE_t(\mathbf{X}_i, a \mid \mathbf{A}_{-i}; \mathbf{A}_{-i} \in \pi^*)$, $\forall i \in \mathcal{N}_{t_1}$*; the policy* π^* *is a local maximizer of* $EE_t(\mathbf{X}_i, A_i \mid \mathbf{A}_{-i})$.

We omit the cluster $g \in \mathcal{G}$ subscripts for ease of exposition. The proof consists of two parts. First, we show that the evaluated EE-function value always weakly increase after each iteration. Then, we show that the assignment policy converges to a (local) optimum when the new assignment proposal is as good as, but no better than, the old policy.

First, consider a focal customer i with covariates $\mathbf{x}_i$. Suppose the interim assignment proposal from the last iteration is $\pi^{(n-1)} \equiv \left(A_i^{(n-1)}, \mathbf{A}_{-i}^{(n-1)}\right)$. The optimization result in this iteration is given by

$$A_i^{(n)} \equiv \underset{a\in\mathcal{A}}{\operatorname{argmax}}\, EE_{t_1}\left(\mathbf{X}_i, a \,\middle|\, \mathbf{A}_{-i}^{(n-1)} \in \pi^{(n-1)}; \widetilde{r}^{(n)}, \widetilde{\mathbf{H}}_{t_2}\right) \tag{A.6}$$

By construction of Equation A.6, $\pi^{(n)} \equiv \left(A_i^{(n)}, \mathbf{A}_{-i}^{(n-1)}\right)$ weakly dominates $\pi^{(n-1)}$, because the former leads to a weakly higher EE-function value, i.e.,

$$\underset{a\in\mathcal{A}}{\operatorname{argmax}}\, EE_{t_1}\left(\mathbf{X}_i, a \,\middle|\, \mathbf{A}_{-i}^{(n)} \in \pi^{(n)}\right) \geq EE_{t_1}\left(\mathbf{X}_i, A_i^{(n-1)} \,\middle|\, \mathbf{A}_{-i}^{(n-1)} \in \pi^{(n-1)}\right). \tag{A.7}$$

Therefore, the iteration in the optimization generates new assignment policies that always weakly improve on the existing policy.

Second, suppose the new assignment policy $\pi^{(n)}$ leads to the same value of the EE-function as the existing interim policy $\pi^{(n-1)}$ for all customers. In this case, $EE_{t_1}^{\pi^{(n-1)}} = EE_{t_1}^{\pi^{(n)}}$. Then, for any $i \in \mathcal{N}_{t_1}$, we have

$$\begin{aligned} EE_{t_1}^{\pi^{(n)}}(\mathbf{X}_i, A_i^{(n)}) \equiv & \underset{a\in\mathcal{A}}{\operatorname{argmax}}\, EE_{t_1}\left(\mathbf{X}_i, a \,\middle|\, \mathbf{A}_{-i}^{(n-1)} \in \pi^{(n-1)}\right) \\ = & EE_{t_1}\left(\mathbf{X}_i, A_i^{(n-1)} \,\middle|\, \mathbf{A}_{-i}^{(n-1)} \in \pi^{(n-1)}\right). \end{aligned} \tag{A.8}$$

And it must be the case in which $\pi^{(n)} \equiv \pi^{(n-1)}$. In the next iteration, the values will not update, and hence the algorithm is converged to a local optimum.

$\blacksquare$

A.3 Pseudo-code of Algorithms

Algorithm 1 `EE-Evaluation`: EE-function $EE_t(g)\left(\mathbf{X}_i, \cdot \mid \mathbf{A}^g_{-i}\right)$ evaluation for a Cluster g customer

Input: data $\mathbf{H}_{t_1} = \{\mathbf{X}, \mathbf{A}_{t_0}, \mathbf{Y}_{t_0}\}$, response function r, current exploit policy p_{t_1}, interim assignments for other Cluster g customers $\mathbf{A}^g_{-i}$.

Compute $IC_{t_1}(\mathbf{X}_i, a)$ for all $a \in \mathcal{A}$ using Equation 1.2.

Construct a generative model $R(Y \mid \cdot, \cdot)$ based on the predictive posterior distribution of $r(\cdot, \cdot)$, as shown in Equation 1.9.

repeat

    for $a \in \mathcal{A}$ **do**

        Construct $\widetilde{\mathbf{Y}}^{g(k)}_{t_1}$ by selecting sample $\widetilde{\mathbf{Y}}^{g(k)}_{t_1} = \left(\widetilde{\mathbf{Y}}_{t_1}(\mathbf{X}_i, a), \widetilde{\mathbf{Y}}_{t_1}(\mathbf{X}^g_{-i}, \mathbf{A}^g_{-i})\right)$.

        Use these to construct artificial history $\widetilde{\mathbf{H}}^{(k)}_{t_2}$.

        Re-learn artificial response function $\widetilde{r}^{(k)} \leftarrow r\left(\cdot, a \,\middle|\, \widetilde{\mathbf{H}}^{(k)}_{t_2}\right)$.

        Optimize $\widetilde{r}$ to get artificial exploit policy $\widetilde{p}^{(k)}_{t_2}\left(\cdot \,\middle|\, \widetilde{\mathbf{H}}^{(k)}_{t_2} \sim a\right) \leftarrow \operatorname{argmax}_a \mathbb{E}_{\widetilde{r}}[\widetilde{r}^{(k)}(\cdot, a)]$.

        Compute the expectation at t_2 using means of the posterior GP for all $j \in \mathcal{N}_{t_2}$,

        $\mu^{(k)}_{j,t_2}(\mathbf{X}_j, a) = \mathbb{E}_{t_2}[\widetilde{r}^{*g(k)}(\mathbf{X}_j) \mid a; \mathbf{A}^g_{-i}, \mathbf{E}^g]$.

    end for

until K times

Compute the expectations of $\widetilde{IV}_{t_2(g)}(\mathbf{X}_i, a \mid \mathbf{A}^g_{-i})$ at t_2, given in Equation A.1 with $\mu^{(k)}_{j,t_2}(\mathbf{X}_j, a)$ by the simulated estimator.

return EE-function values $EE_{t_1(g)}(\mathbf{X}_i, a \mid \mathbf{A}^g_{-i})$ computed using Equation 1.4 for all $a \in \mathcal{A}$.

Algorithm 2 `OLAT`: One-step Look Ahead Targeting Optimization

Input: data $\mathbf{H}_{t_1} = \{\mathbf{X}, \mathbf{A}_{t_0}, \mathbf{Y}_{t_0}\}$, response function r, current exploit policy p_{t_1}.

Initialize response function with $\widetilde{r} \leftarrow r_{t_1}$, Period t_1's EE-function $EE^{(0)}_{t_1(g)}$, and Period t_1 assignment policy with $\pi^{(0)}_{(g)}$.

while not converge **or** below iteration limit **do**

    repeat

        Simulate outcome samples $\widetilde{\mathbf{Y}}^{g}_{t_1}(\mathcal{A})$ for K times.

        parallel Cluster $g \in \mathcal{G}$

            Evaluate $EE^{(n)}_{t_1(g)}$ using algorithm `EE-Evaluation`,

                based on the assignment policy from the last iteration $\pi^{(n-1)}_{(g)}$.

            Evaluate $\max_a EE^{(n)}_{t_1(g)}\left(\cdot, a \,\middle|\, \mathbf{A}^{g(n-1)}_{-1} \in \pi^{(n-1)}_{(g)}; \widetilde{r}^{(n)}, \widetilde{\mathbf{H}}_{t_2}\right)$.

            Propose a new search direction with ε-greedy search.

            Propose new assignments $\pi^{(n)}_{(g)}$ based on the direction and externality metrics $\mathbf{E}^{g(n)}$.

        end parallel

    until M global steps

end while

return Period t_1 assignment policy π^*.

Appendix B

Appendix for Chapter 2

B.1 Figures and Tables

Figure B.1: Price Categories

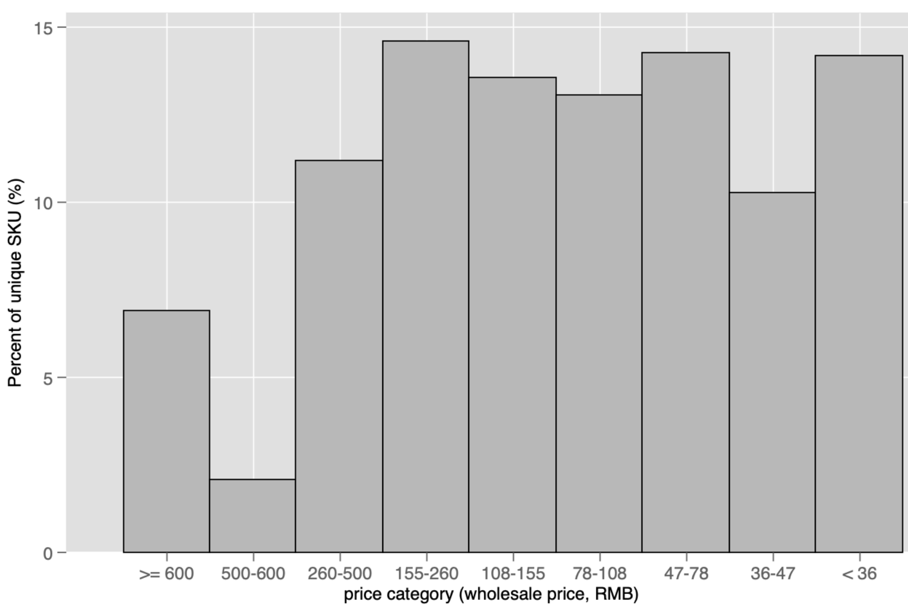


Figure B.2: Map of High-end Cigarette Sales Ratio in 2013

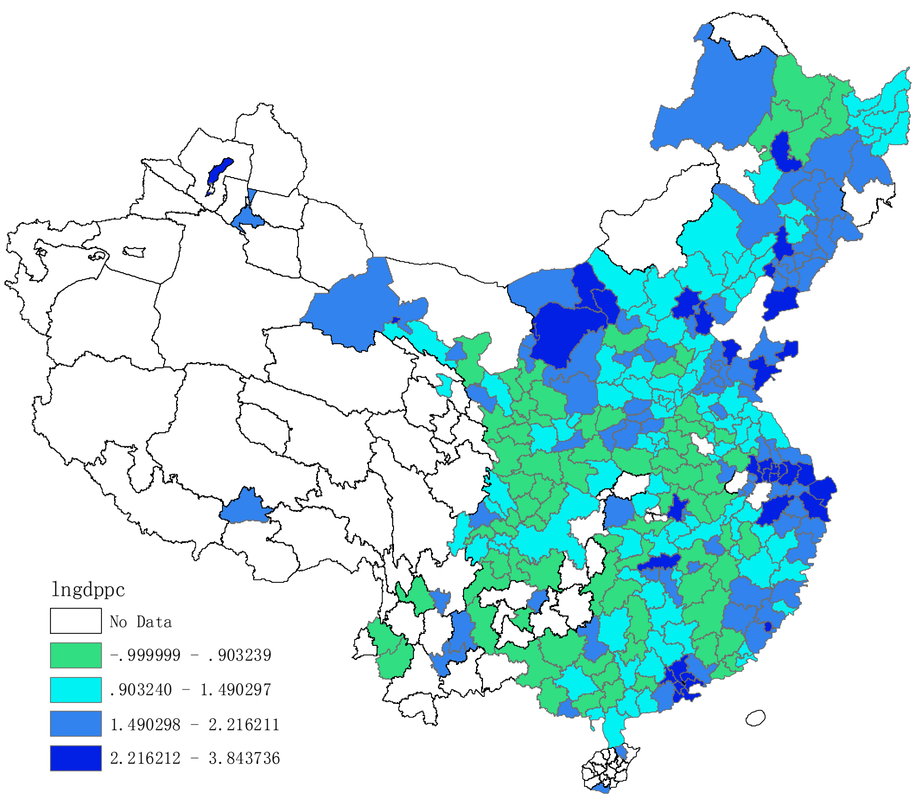


Figure B.3: Map of GDP per Capita in 2013

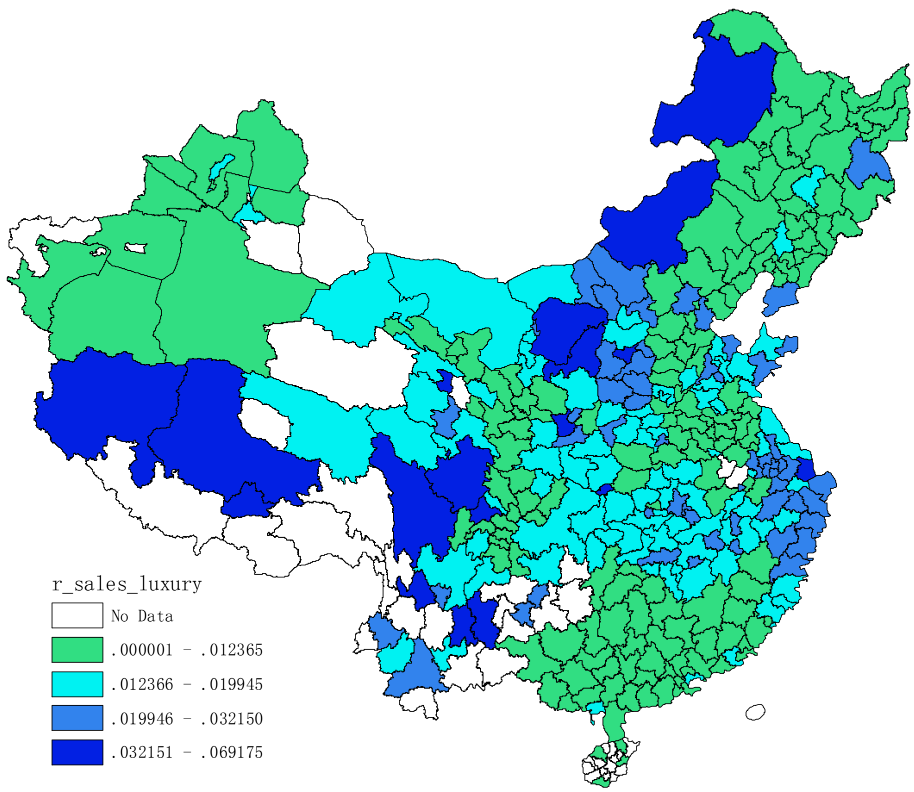


Figure B.4: Seasional Pattern: National Sales of Premium Cigarettes

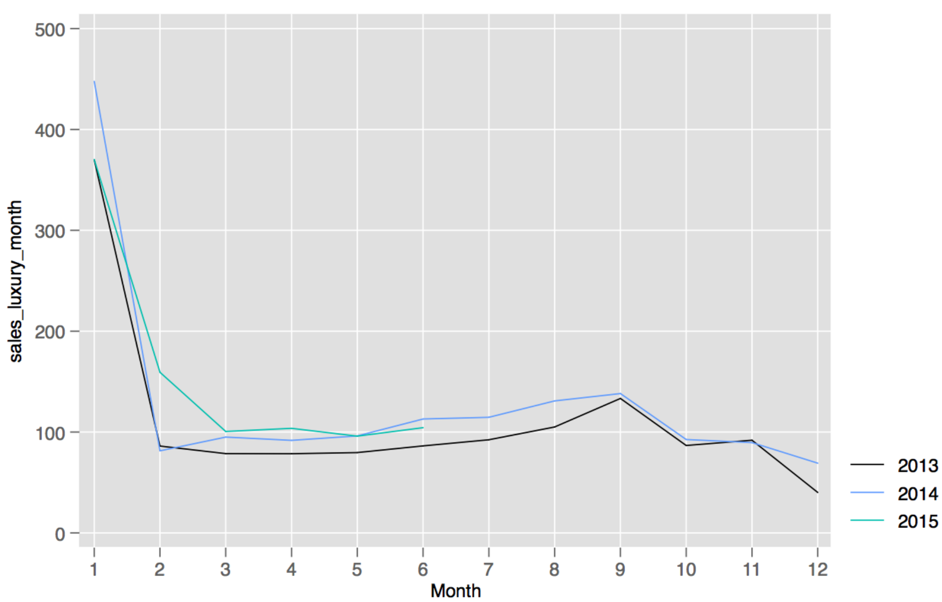


Table B.1: Data source of all Variables

Variable	Definition	Time Window	Data Source
r_sales_luxury	share of sales of luxury cigarette	2007–2014	Department of State Tobacco Monopoly Administration
sales_luxury	sales (unit) of luxury cigarette	2007–2014	Department of State Tobacco Monopoly Administration
r_sales_luxury (monthly)	share of sales of luxury cigarette	2007 Jan–2015 July	Department of State Tobacco Monopoly Administration
sales_luxury (monthly)	sal s of luxury cigarette	2007 Jan–2015 July	Department of State Tobacco Monopoly Administration
gdp	GDP (100 million RMB)	2007–2015	China City Statistical Yearbook (CCSY)
population	Population (in 10,000)	2007–2015	China City Statistical Yearbook (CCSY)
pop_density	population density (person / km^2)	2007–2015	China City Statistical Yearbook (CCSY)
r_service_gdp	share (%) of service industry in GDP	2007–2015	China City Statistical Yearbook (CCSY)
producer_luxury_city	dummy: luxury brand producer in the city	invariant (2007–2015)	(website)
exp_ttlpc	fiscal expenditure (10,000 RMB) per capita	2007	China Ministry of Finance Report
recruit_cddt_post	ratio of number of applicants to posted number of slots	2017	government website
r_v_foreign	share of foreign invested firm output in gross industry output	2007–2015	China City Statistical Yearbook (CCSY)
r_n_soe	share of people working in state owned firms	2007–2009	Urban Household Survey(UHS)
gift_eatdrinksmoke	gift (unit) from other people: eat, drink, or smoke	2007–2009	Urban Household Survey(UHS)
bribe	share of household used *guanxi* with government	2013	City Public Governance Survey (CPGS)
i_election	number of villages having elections, averaged based on village level	2007	China Family Panel Studies (CFPS)
gift_sending	number of gift given to friends, averaged based on household level	2007	China Family Panel Studies (CFPS)
gift_received	number of gift received, averaged based on household level	2007	China Family Panel Studies (CFPS)
v_english	rating of importance of English, averaged based on individual level	2007	China Family Panel Studies (CFPS)
i_travelout	number of people been to HMT*, averaged based on individual level	2007	China Family Panel Studies (CFPS)
inspection	number of inspections from CPC Central Commission for Discipline	2007 Jan–2015 July	government website
inspection_prov	number of inspections from CPC Provincial Commission for Discipline	2007 Jan–2015 July	government website
anticase_prov	number of declarations of provincial level official crimes	2007 Jan–2015 July	government website
anticase_city	number of declarations of city level official crimes	2007 Jan–2015 July	government website

Notes: *: Hong Kung, Macau, Taiwan

Table B.2: Descriptive statistics

Variable	Obs	Mean	Std. Dev.	Min	Max
r_sales_luxury	2,011	.0097616	.0071813	.0005058	.044766
sales_luxury	2,011	1556.127	2575.303	.85	35805.9
gdppc	2,011	3.943141	4.060653	.3547228	48.1692
r_service_gdp	2,011	36.2037	8.391318	11.8	77.95
pop_density	2,011	468.0424	328.6323	4.82	2648.11
producer_luxury_city	2,011	.8209846	.3834606	0	1
exp_ttlpc	1,869	1178.145	806.8132	135.6818	8776.027
recruit_cddt_post	1,995	55.50625	28.61516	12.12676	449.6667
r_v_foreign	2,011	.0970047	.106743	.000032	.5396062
number of provinces	28				
number of cities	264				
number of luxury brands	33				
number of luxury SKU	100				

Table B.3: Baseline specification

Variables	(1) lgt_r_sales_luxury	(2) ln_sales_luxury	(3) lgt_r_sales_luxury	(4) ln_sales_luxury
lngdppc	0.610***	0.680***	0.362***	0.366***
	(0.0237)	(0.0315)	(0.0214)	(0.0325)
r_service_gdp	0.00151	0.0280***	0.00569***	0.0335***
	(0.00183)	(0.00272)	(0.00150)	(0.00249)
pop_density	4.66e-05	0.000759***	0.000145***	0.000883***
	(4.46e-05)	(6.19e-05)	(4.01e-05)	(6.48e-05)
producer_luxury	-0.194***	0.0982*	-0.150***	0.154***
	(0.0373)	(0.0543)	(0.0334)	(0.0501)
Constant	-5.462***	4.517***	-6.057***	3.753***
	(0.0683)	(0.0961)	(0.0679)	(0.103)
Year dummy	N	N	Y	Y
Observations	2,011	2,011	2,011	2,011
R-squared	0.334	0.409	0.522	0.545

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table B.4: Institutional efficiency results

VARIABLES	(1) lgt_r_sales_luxury	(2) ln_sales_luxury	(3) lgt_r_sales_luxury	(4) ln_sales_luxury	(5) lgt_r_sales_luxury	(6) ln_sales_luxury
lnexp_ttlpc	0.105***	0.351***			0.105***	0.351***
	(0.0242)	(0.0359)			(0.0246)	(0.0360)
recruit_cddt_post	0.00356***	0.00515***			0.00366***	0.00525***
	(0.000490)	(0.000702)			(0.000483)	(0.000696)
r_v_foreign			-0.649***	-0.584**	-0.702***	-0.663***
			(0.138)	(0.237)	(0.141)	(0.234)
lngdppc	0.329***	0.207***	0.457***	0.464***	0.374***	0.249***
	(0.0223)	(0.0346)	(0.0232)	(0.0325)	(0.0251)	(0.0371)
r_service_gdp	0.00390***	0.0292***	0.00843***	0.0348***	0.00651***	0.0317***
	(0.00149)	(0.00210)	(0.00157)	(0.00242)	(0.00151)	(0.00218)
pop_density	0.000170***	0.000882***	0.000175***	0.000844***	0.000201***	0.000911***
	(4.17e-05)	(6.12e-05)	(4.21e-05)	(6.38e-05)	(4.19e-05)	(6.26e-05)
producer_luxury	-0.137***	0.223***	-0.117***	0.231***	-0.122***	0.237***
	(0.0340)	(0.0468)	(0.0348)	(0.0490)	(0.0336)	(0.0466)
Constant	-6.926***	1.264***	-6.188***	3.714***	-7.006***	1.188***
	(0.177)	(0.262)	(0.0686)	(0.100)	(0.177)	(0.263)
Year dummy	Y	Y	Y	Y	Y	Y
Observations	1,853	1,853	1,853	1,853	1,853	1,853
R-squared	0.560	0.599	0.547	0.567	0.566	0.601

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table B.5: UHIES survey results

VARIABLES	(1) lgt_r_sales_luxury	(2) ln_sales_luxury	(3) lgt_r_sales_luxury	(4) ln_sales_luxury
r_n_soe	0.662**	0.587	0.627**	0.506
	(0.289)	(0.432)	(0.295)	(0.443)
gift_eatdrinksmoke			0.00184**	0.00427***
			(0.000894)	(0.00132)
lngdppc	0.301***	0.440***	0.315***	0.472***
	(0.0594)	(0.0825)	(0.0611)	(0.0823)
r_service_gdp	0.0170***	0.0484***	0.0165***	0.0474***
	(0.00590)	(0.00909)	(0.00582)	(0.00880)
pop_density	0.000565***	0.00109***	0.000549***	0.00105***
	(9.60e-05)	(0.000167)	(9.46e-05)	(0.000152)
producer_luxury	0.113	0.297*	0.0858	0.233
	(0.107)	(0.166)	(0.108)	(0.169)
Constant	-6.954***	2.935***	-6.958***	2.926***
	(0.236)	(0.327)	(0.235)	(0.328)
Observations	202	202	202	202
R-squared	0.432	0.583	0.440	0.598
Number of cities	71	71	71	71
Number of provinces	7	7	7	7

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table B.6: City Public Governance Survey in 2013

VARIABLES	(1) lgt_r_sales_luxury	(2) ln_sales_luxury
bribe	0.320*	0.616***
	(0.176)	(0.231)
lngdppc	0.824***	1.175***
	(0.0780)	(0.112)
r_service_gdp	-0.00590	-0.00283
	(0.00366)	(0.00602)
pop_density	0.000144	0.000322**
	(9.94e-05)	(0.000157)
producer_luxury	-0.620***	0.191
	(0.0812)	(0.137)
Constant	-5.218***	5.569***
	(0.228)	(0.340)
Observations	231	231
R-squared	0.440	0.511
Number of cities	29	29

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table B.7: CDI inspections

Variables	(1) lgt_r_sales_luxury	(2) ln_sales_luxury	(3) lgt_r_sales_luxury	(4) ln_sales_luxury
L.inspection	0.0172	0.00745	0.0214	0.0111
	(0.0289)	(0.0298)	(0.0292)	(0.0301)
L2.inspection	-0.121***	-0.139***	-0.120***	-0.140***
	(0.0317)	(0.0391)	(0.0319)	(0.0392)
L3.inspection	-0.0948***	-0.0777**	-0.0991***	-0.0827***
	(0.0287)	(0.0306)	(0.0291)	(0.0311)
L4.inspection	-0.0536**	-0.0730***	-0.0559**	-0.0745***
	(0.0250)	(0.0262)	(0.0249)	(0.0260)
L5.inspection	-0.00873	0.138***	-0.00476	0.143***
	(0.0353)	(0.0377)	(0.0356)	(0.0380)
L.inspection_prov			0.0545**	0.0313
			(0.0254)	(0.0284)
L2.inspection_prov			-0.0505*	-0.0490
			(0.0300)	(0.0325)
L3.inspection_prov			-0.0308	-0.0623*
			(0.0337)	(0.0360)
L4.inspection_prov			-0.0305	-0.0203
			(0.0467)	(0.0485)
L5.inspection_prov			0.0425	0.0567*
			(0.0324)	(0.0343)
lngdppc	0.00131	0.0440	0.00263	0.0472
	(0.0502)	(0.0836)	(0.0516)	(0.0865)
r_service_gdp	0.0226***	0.0318***	0.0227***	0.0322***
	(0.00316)	(0.00367)	(0.00321)	(0.00374)
pop_density	4.42e-06	6.22e-05	6.32e-06	6.79e-05
	(0.000400)	(0.000473)	(0.000401)	(0.000475)
L.lgt_r_sales_luxury	0.121***		0.121***	
	(0.0142)		(0.0142)	
L.ln_sales_luxury		0.0566***		0.0561***
		(0.0142)		(0.0143)
Constant	-3.946***	4.463***	-3.952***	4.446***
	(0.242)	(0.302)	(0.243)	(0.305)
Month dummy	Y	Y	Y	Y
Observations	6,690	6,690	6,690	6,690
R-squared	0.327	0.511	0.328	0.512
Number of cities	274	274	274	274

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table B.8: Corruption cases

Variables	(1) lgt_r_sales_luxury	(2) ln_sales_luxury	(3) lgt_r_sales_luxury	(4) ln_sales_luxury
L.corruptcase_prov	-0.0242*	-0.0317*	-0.0238	-0.0312*
	(0.0146)	(0.0166)	(0.0146)	(0.0167)
L2.corruptcase_prov	-0.0441***	-0.0392**	-0.0441***	-0.0394***
	(0.0141)	(0.0152)	(0.0141)	(0.0152)
L3.corruptcase_prov	-0.0208	-0.0252	-0.0197	-0.0242
	(0.0163)	(0.0208)	(0.0164)	(0.0207)
L4.corruptcase_prov	0.0531***	0.0569***	0.0533***	0.0571***
	(0.0143)	(0.0151)	(0.0142)	(0.0151)
L5.corruptcase_prov	0.0565***	0.110***	0.0572***	0.111***
	(0.0160)	(0.0182)	(0.0161)	(0.0182)
L.corruptcase_city			-0.00771	-0.0249
			(0.0307)	(0.0341)
L2.corruptcase_city			0.0258	0.0221
			(0.0287)	(0.0313)
L3.corruptcase_city			-0.0541**	-0.0244
			(0.0262)	(0.0308)
L4.corruptcase_city			-0.0102	0.00558
			(0.0352)	(0.0367)
L5.corruptcase_city			-0.00960	-0.0481
			(0.0311)	(0.0346)
lngdppc	-0.00135	0.0432	-0.000272	0.0441
	(0.0523)	(0.0870)	(0.0531)	(0.0876)
r_service_gdp	0.0207***	0.0288***	0.0209***	0.0290***
	(0.00321)	(0.00368)	(0.00324)	(0.00373)
pop_density	4.18e-05	0.000127	5.21e-05	0.000139
	(0.000386)	(0.000462)	(0.000383)	(0.000461)
L.lgt_r_sales_luxury	0.121***		0.121***	
	(0.0144)		(0.0144)	
L.ln_sales_luxury		0.0539***		0.0541***
		(0.0143)		(0.0143)
Constant	-3.895***	4.552***	-3.907***	4.536***
	(0.235)	(0.289)	(0.236)	(0.290)
Month dummy	Y	Y	Y	Y
Observations	6,690	6,690	6,690	6,690
R-squared	0.327	0.513	0.328	0.513
Number of cities	274	274	274	274

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table B.9: Culture factors

VARIABLES	(1) lgt_r_sales_luxury	(2) ln_sales_luxury	(3) lgt_r_sales_luxury	(4) ln_sales_luxury
gift_sending	0.00273***	0.00192***		
	(0.000298)	(0.000404)		
gift_received	9.64e-05***	0.000245***		
	(1.38e-05)	(2.04e-05)		
i_election			-0.753***	-0.656***
			(0.0703)	(0.116)
v_english			-0.701***	-0.629***
			(0.0672)	(0.0934)
i_travel2gat			-1.160***	-1.235***
			(0.368)	(0.463)
lngdppc	0.357***	0.374***	0.375***	0.375***
	(0.0212)	(0.0323)	(0.0210)	(0.0330)
r_service_gdp	0.00513***	0.0328***	0.00795***	0.0378***
	(0.00158)	(0.00253)	(0.00157)	(0.00256)
pop_density	0.000147***	0.000884***	0.000138***	0.000869***
	(4.08e-05)	(6.30e-05)	(4.46e-05)	(6.46e-05)
producer_luxury	-0.148***	0.132**	-0.0141	0.259***
	(0.0344)	(0.0526)	(0.0336)	(0.0535)
Constant	-6.229***	3.493***	-3.698***	5.807***
	(0.0685)	(0.104)	(0.218)	(0.300)
Year dummy	Y	Y	Y	Y
Observations	1,974	1,974	1,974	1,974
R-squared	0.548	0.570	0.562	0.563

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table B.10: Overall evaluation

VARIABLES	(1) lgt_r_sales_luxury	(2) ln_sales_luxury
lnexp_ttlpc	0.105***	0.329***
	(0.0292)	(0.0422)
recruit_cddt_post	0.00233***	0.00423***
	(0.000551)	(0.000773)
r_v_foreign	-0.770***	-0.578**
	(0.143)	(0.234)
gift_sending	0.00249***	0.00150***
	(0.000315)	(0.000412)
gift_received	7.56e-05***	0.000222***
	(1.58e-05)	(2.15e-05)
i_election	-0.314***	-0.150
	(0.0880)	(0.132)
v_english	-0.482***	-0.429***
	(0.0799)	(0.112)
lngdppc	0.393***	0.285***
	(0.0253)	(0.0381)
r_service_gdp	0.00713***	0.0297***
	(0.00176)	(0.00264)
pop_density	0.000180***	0.000901***
	(4.56e-05)	(6.06e-05)
producer_luxury	-0.0739**	0.245***
	(0.0359)	(0.0527)
Constant	-5.564***	2.479***
	(0.304)	(0.389)
Year dummy	Y	Y
Observations	1,820	1,820
R-squared	0.601	0.625

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Appendix C

Appendix for Chapter 3

C.1 Proofs of Main Results

Proof of Result 4 (Decline in Search)

We prove this result by showing the conditions for search in each period and compare the fractions of customers in the two periods.

A customer searches in Period t only if $\mathbb{E}U_t^s \geq \mathbb{E}U_t^0$. These expectations are taken with respect to the customer's beliefs about $\Pr(H|m_1)$ and $\Pr(m_2 = 1|m_1)$. In Period 1, the customer believes $r_H = r_L = 1/2$, and the probability of discovery is $1/2\alpha_H$. In Period 2, a customer's posterior belief is updated based on $m_1 \in \{0, 1, \varnothing\}$. If she observes $m_1 = 0$, she updates her beliefs about the state to $\Pr(H|m_1 = 0) = \frac{1-\alpha_H}{2-\alpha_H}$ and $\Pr(L|m_1 = 0) = \frac{1}{2-\alpha_H}$, and the probability of discovery is $\frac{\alpha_H(1-\alpha_H)}{2-\alpha_H}$ in Period 2. If she observes $m_1 = 1$, she knows $\theta = H$, and the probability of discovery is 1 in Period 2. If she did not search in Period 1 ($m_1 = \varnothing$), she still believes $r_H = r_L = 1/2$, and the probability of discovery is $\frac{1}{2}\alpha_H + \alpha_H(1 - \alpha_H) = \frac{1}{2}\alpha_H(2 - \alpha_H)$.

By backward induction, we start from the decision in Period 2, which is contingent on the customer's observation m_1, so we need to separate these cases. If $m_1 = 0$, she believes $\Pr(H|m_1 = 0) = \frac{1-\alpha_H}{2-\alpha_H}$, and $\Pr(m_2 = 1|m_1 = 0) = \frac{\alpha_H(1-\alpha_H)}{2-\alpha_H}$. The customer's period utilities are$\mathbb{E}U_2^s = v_b - b - s + \frac{\alpha_H(1-\alpha_H)}{2-\alpha_H}(v_N - p)^+$, and $\mathbb{E}U_2^0 = v_b - b$; the solution to $\mathbb{E}U_2^s \geq \mathbb{E}U_2^0$ is given by $v_N \geq p + \frac{(2-\alpha_H)s}{\alpha_H(1-\alpha_H)}$. If $m_1 = 1$, the customer will not search again in Period 2, and her period utility is $\mathbb{E}U_2^0 = v_b - b + v_N - p$. If $m_1 = \varnothing$, the customer's period utilities are $\mathbb{E}U_2^s = v_b - b - s + \frac{1}{2}\alpha_H(2-\alpha_H)(v_N - p)^+$, and $\mathbb{E}U_2^0 = v_b - b$; the solution is $v_N \geq p + \frac{2s}{\alpha_H(1-\alpha_H)}$, conditional on this set including customers who did not search in Period 1.

In Period 1, the customer's utilities are given by Equations 3.1. We know from the derivations in Period 1 that the two $[\cdot]^+$ operators in these equations are non-negative when $v_N \geq p + \frac{(2-\alpha_H)s}{\alpha_H(1-\alpha_H)}$ and $v_N \geq p + \frac{2s}{\alpha_H(1-\alpha_H)}$, respectively. To solve $\mathbb{E}U_1^s \geq \mathbb{E}U_1^0$, we need to compare these two quantities; since $p + \frac{(2-\alpha_H)s}{\alpha_H(1-\alpha_H)} > p + \frac{2s}{\alpha_H(1-\alpha_H)}$ can be simplified to $(1-\alpha_H)^2 + 1 > 0$, this inequality holds true for any $\alpha_H \in [0, 1]$, and we can discuss the cases in each value segment separately.

If $v_N \geq p + \frac{(2-\alpha_H)s}{\alpha_H(1-\alpha_H)}$, both $[\cdot]^+$ operators return non-negative, and the solution to $\mathbb{E}U_1^s \geq \mathbb{E}U_1^0$ is $v_N \geq p + \frac{(2-\alpha_H)s}{\alpha_H}$. The final solution considering the value range is $v_N \geq \frac{(2-\alpha_H)s}{\alpha_H(1-\alpha_H)}$. If $v_N \in \left[p + \frac{2s}{\alpha_H(1-\alpha_H)}, p + \frac{(2-\alpha_H)s}{\alpha_H(1-\alpha_H)}\right)$, only the second $[\cdot]^+$ operator returns non-negative, and the solution to $\mathbb{E}U_1^s \geq \mathbb{E}U_1^0$ is $v_N \geq p$, so $p \in \left[p + \frac{2s}{\alpha_H(1-\alpha_H)}, p + \frac{(2-\alpha_H)s}{\alpha_H(1-\alpha_H)}\right)$. If $v_N \in \left[p, p + \frac{2s}{\alpha_H(1-\alpha_H)}\right)$, the solution is $v_N \geq p + \frac{s}{\alpha_H}$. It's easy to show that $p + \frac{s}{\alpha_H} < p + \frac{2s}{\alpha_H(1-\alpha_H)}$, so the final solution is $v_N \in \left[p, p + \frac{s}{\alpha_H}\right)$.

To summarize, the searchers and non-searchers in Period 1 are $v_N \in \left[p + \frac{s}{\alpha_H}, \bar{v}\right]$ and $v_N \in \left[0, p + \frac{s}{\alpha_H}\right)$, respectively.

We need to then refine the Period 2 search decision by incorporating Period 1 search conditions. The non-searchers in Period 1, $v_N \in \left[0, p + \frac{s}{\alpha_H}\right)$, contradicts the condition for searching again in Period 2 ($v_N \geq p + \frac{2s}{\alpha_H(1-\alpha_H)}$), so the non-searchers in Period 1 will not search in Period 2.

To summarize, the searchers in Period 2 are $v_N \in \left[p + \frac{(2-\alpha_H)s}{\alpha_H(1-\alpha_H)}, \bar{v}\right]$ when $m_1 = 0$ (since $p + \frac{(2-\alpha_H)s}{\alpha_H(1-\alpha_H)} > p + \frac{s}{\alpha_H}$), and $v_N \in \varnothing$ when $m_1 \in \{1, \varnothing\}$.

Comparing the conditions for Periods 1 and 2, we conclude that the total search declines in Period 2 regardless of the search outcome in Period 1.

■

Proof of Result 5 (Customer's Preference with Differentiated Products)

When the customer observes $m_t = 10$, the condition for $N_I \succ B$ is given by $v_b + v + \tau v_N - (b+p) \geq v_b - b$, which gives $v_N \geq \frac{1}{\tau}(p - v)$. When $v > p$, this means everyone prefers N_I. When the customer observes $m_t = 01$, the condition for $N_E \succ B$ is given by $v_b + v + \tau(1 - v_N) - (b+p) \geq v_b - b$, which gives $v_N \leq 1 - \frac{1}{\tau}(p - v)$. When $v > p$, this means everyone prefers N_E.

When the customer observes $m_t = 11$, she can choose one product from $\{B, N_I, N_E\}$, and she makes decision by solving the below problem:

$$\max\{v_b - b,\, v_b + v + \tau v_N - (b+p),\, v_b + v + \tau(1 - v_N) - (b+p)\}$$

All customers $v_N \in U[0,1]$ will choose between N_I and N_E when presented this three-product choice problem; $v_N \in [0, 1/2]$ customers will choose to buy N_E, and $v_N \in (1/2, 1]$ customers will choose to buy N_I.

One can easily show with first order conditions that in complete information version of these games and when firm can price their new product, the optimal price is v when there is one firm (I or E) offering the new product in the market and τ when both firms offer new products, and this justifies the assumption that $p \leq v$ in the customer decision-making model.

■

Proof of Result 6 (Decline in Search with Product Competition)

In Period 1, a customer holds her prior belief on the state, given by $r_H = r_L = 1/2$. She thinks the probabilities of discovering N_I and N_E in Period 1 are thus $\frac{1}{2}$ and $\frac{1}{2}\alpha_H$.

The customer's beliefs in Period 2 are: If $m_1 = 11$, the customer knows $h_2 = 11$, so $\Pr(\text{Discover } N_E \text{ in Period 2}) = 1$. If $m_1 = 10$, then $h_2 = 11, 10$ with probabilities α_H and $1 - \alpha_H$, so she believes $\Pr(\text{Discover } N_E \text{ in Period 2}) = \alpha_H$. If $m_1 = 00$, the customer knows $h_2 = 00$. If $m_1 = \varnothing$, the customer thinks $r_H = 1/2$; she discovers N_I and N_E in Period 2 with probabilities $\frac{1}{2}$ and $\frac{1}{2}\alpha_H(2 - \alpha_H)$.

Search conditions in Period 2. A customer's search decision is contingent on m_1. When $m_1 = 00$, then $\theta = L$ and $m_2 = h_2 = 00$, so she will not search, and get $\mathbb{E}U_2^0 = v_b - b$. Everyone buys B. When $m_1 = 11$, then $\theta = H$ and $m_2 = h_2 = 11$, so she will not search, and get $\mathbb{E}U_2^0 = v_b - b + v + \tau v_N - p + \tau(1 - 2v_N)^+$. So $v_N \in [0, 1/2]$ customers will choose to buy N_E, and $v_N \in (1/2, 1]$ customers will choose to buy N_I.

When $m_1 = 10$, then $\Pr(\text{Discover } N_E \text{ in Period 2}) = \alpha_H$. Her search is only meaningful when she prefers N_E and discover it in Period 2. Her period utilities are $\mathbb{E}U_2^s = v_b - b + v + \tau v_N - p - s + \alpha_H \cdot \tau(1 - 2v_N)^+$ and $\mathbb{E}U_2^0 = v_b - b + v + \tau v_N - p$. We next find the conditions in which customers choose to search. To solve $\mathbb{E}U_2^s \geq \mathbb{E}U_2^0$, we realize the $(\cdot)^+$ holds non-negative when $v_N \leq 1/2$. When $v_N > 1/2$, it follows $\mathbb{E}U_2^s < \mathbb{E}U_2^0$, and these customers will not search. When $v_N \leq 1/2$, the search condition is $v_N \leq \hat{v}_1$, where

$$\hat{v}_1 \equiv \frac{1}{2}\left(1 - \frac{s}{\alpha_H \tau}\right).$$

We know $\hat{v}_1 < \frac{1}{2}$ always holds. From Assumption 1, $s - \alpha_H \tau < 0$, we have $\hat{v}_1 > 0$. The customers who will search in Period 2 are $v_N \in [0, \hat{v}_1]$.

When $m_1 = \varnothing$, the customer thinks $r_H = 1/2$. Her period utilities are $\mathbb{E}U_2^0 = v_b - b$ and $\mathbb{E}U_2^s = v_b - b - s + \frac{1}{2}(v + \tau v_N - p) + \frac{1}{2}\alpha_H(2 - \alpha_H) \cdot \tau(1 - 2v_N)^+$ To solve $\mathbb{E}U_2^s \geq \mathbb{E}U_2^0$, we realize the $(\cdot)^+$ holds non-negative when $v_N \leq 1/2$. When $v_N > 1/2$, the solution to $\mathbb{E}U_2^s \geq \mathbb{E}U_2^0$ is $v_N \geq \hat{v}_2$, where

$$\hat{v}_2 \equiv \frac{1}{\tau}(p - v + 2s),$$

and from Assumption 1, $\hat{v}_2 > \frac{1}{2}$.

When $v_N \leq 1/2$, the search condition under Assumption 1 is $v_N \in \varnothing$. To derive this, first see the solution to $\mathbb{E}U_2^s \geq \mathbb{E}U_2^0$ is

$$\left[2(1-\alpha_H)^2 - 1\right] v_N \geq \frac{1}{\tau}\left[p - v + 2s + \tau\left[(1-\alpha_H)^2 - 1\right]\right] = \hat{v}_2 + (1-\alpha_H)^2 - 1,$$

where $2(1-\alpha_H)^2 - 1 > 0$ when $\alpha_H < 1 - \frac{1}{\sqrt{2}}$. If $\alpha_H < 1 - \frac{1}{\sqrt{2}}$, $v_N \geq \frac{1}{2(1-\alpha_H)^2-1}[\hat{v}_2 + (1-\alpha_H)^2 - 1]$; the right-hand side of the inequality equals $\frac{\hat{v}_2+(1-\alpha_H)^2-1}{2(1-\alpha_H)^2-1} > \frac{\frac{1}{2}+(1-\alpha_H)^2-1}{2(1-\alpha_H)^2-1} = \frac{1}{2}$, which contradicts $v_N \leq 1/2$, so nobody searches. If $\alpha_H = 1 - \frac{1}{\sqrt{2}}$, the inequality becomes $\hat{v}_2 - \frac{1}{2} \leq 0$, and it contradicts $\hat{v}_2 > 1/2$; nobody searches.

If $\alpha_H > 1 - \frac{1}{\sqrt{2}}$, $v_N \leq \frac{\hat{v}_2+(1-\alpha_H)^2-1}{2(1-\alpha_H)^2-1}$; suppose $x \equiv (1-\alpha_H)^2$, we know from $\alpha_H > 1 - \frac{1}{\sqrt{2}}$ that $x < \frac{1}{2}$, and the right-hand side of the inequality equals $\frac{\hat{v}_2+x-1}{2x-1}$. If $\hat{v}_2 + x < 1$, then $v_N \leq \frac{\hat{v}_2+x-1}{2x-1} > 0$, meaning searchers are $v_N \in [0, \frac{\hat{v}_2+x-1}{2x-1}]$, conditional on these customers being non-searchers in Period 1. if $\hat{v}_2 + x \geq 1$, then $v_N \leq \frac{\hat{v}_2+x-1}{2x-1} \leq 0$, meaning nobody searches. From Assumption 1, $\alpha_H \leq 1 - \sqrt{1-\hat{v}_2}$ when $\hat{v}_2 \in (\frac{1}{2}, 1)$, it implies $\hat{v}_2 + x \geq 1$ whenever $\hat{v}_2 \in (\frac{1}{2}, 1)$. In addition, when $\hat{v}_2 \geq 1$, $\hat{v}_2 + x \geq 1$ always holds for any $\alpha_H > 1 - \frac{1}{\sqrt{2}}$. This means $\hat{v}_2 + x \geq 1$ and nobody searches.

To summarize, when $m_1 = \varnothing$ and under Assumption 1, Period 2 searchers are $v_N \in [\hat{v}_2, 1]$; these customers prefer N_I. This search condition is pending refinements after having search conditions in Period 1.

Search conditions in Period 1. For a customer, she discovers N_I and N_E with probabilities $\frac{1}{2}$ and $\frac{1}{2}\alpha_H$. Looking forward to Period 2, if she does not search in Period 1, $\Pr(\text{Discover } N_I) = \frac{1}{2}$ and $\Pr(\text{Discover } N_E) = \frac{1}{2}\alpha_H(2-\alpha_H)$. Her period utilities are thus given by Equations 3.2. Next, we solve the inequality $\mathbb{E}U_1^s \geq \mathbb{E}U_1^0$ to find customers who search in Period 1.

When $v_N > 1/2$, the utility becomes $\mathbb{E}U_1^s = 2(v_b - b) - s + (v + \tau v_N - p)$ and $\mathbb{E}U_1^0 = 2(v_b - b) + \left[-s + \frac{1}{2}(v + \tau v_N - p)\right]^+$. The boundary condition for making the $[\cdot]^+$ operator non-negative is $v_N \geq \hat{v}_2 > \frac{1}{2}$, from our discussion of Period 2 search decisions. If $\frac{1}{2} < v_N < \hat{v}_2$, the solution to $\mathbb{E}U_1^s \geq \mathbb{E}U_1^0$ is given by $v_N \geq \hat{v}_3$, where

$$\hat{v}_3 \equiv \frac{1}{\tau}(p - v + s)$$

From Assumption 1, $\hat{v}_3 > \frac{1}{2}$; since $\frac{1}{2} < \hat{v}_3 < \hat{v}_2$, the final solution is $v_N \in [\hat{v}_3, \hat{v}_2)$. If $v_N \geq \hat{v}_2$, the solution is $v_N \geq \frac{1}{\tau}(p - v) < 0$, and the final solution is $v_N \geq \hat{v}_2$.

Combine the above analysis, we know the searchers with $v_N > 1/2$ are $[\hat{v}_3, 1]$, with preference $N_I \succ N_E \succ B$. The non-searchers are $(\frac{1}{2}, \hat{v}_3)$.

Next, we refine the search decision in Period 2 for $v_N > 1/2$ considering Period 1 searchers being $[\hat{v}_3, 1]$. In Period 2, these searchers will not search again. For the non-searchers, $(\frac{1}{2}, \hat{v}_3)$, they will search if $v_N > \hat{v}_2 > \hat{v}_3$, which means none of them will search, and $v_N \in \varnothing$. Notably, even without the two regularity conditions on $\hat{v}_2$ and $\hat{v}_3$, $\hat{v}_3 < \hat{v}_2$ always holds true, and the result remains that nobody searches in Period 2, which is a decrease from the Period 1 search volume.

When $v_N \leq 1/2$, the $[\cdot]^+$ operator in $\mathbb{E}U_1^s$ is non-negative when $v_N \leq \hat{v}_1$, where $\hat{v}_1 \equiv \frac{1}{2}(1 - \frac{s}{\alpha_H \tau})$. Based on our discussion about customer decisions when $v_N \leq 1/2$ and $m_1 = \varnothing$, the $[\cdot]^+$ operator in $\mathbb{E}U_1^s$ is always negative, which means that the $v_N \leq 1/2$ customers will not search in Period 2 if they do not search in Period 1. We look at the two cases separately.

- If $v_N > \hat{v}_1$, this implies that the customer will *not* search Period 2 after observing $m_1 = 10$. Solve $\mathbb{E}U_1^s \geq \mathbb{E}U_1^0$, and we get $(1 - 2\alpha_H)v_N \geq \hat{v}_3 - \alpha_H$. Since $1 - 2\alpha_H > 0$ is equivalent to $\alpha_H < \frac{1}{2}$, we can check each case it implies. If $\alpha_H = \frac{1}{2}$, then $0 \geq \hat{v}_3 - \frac{1}{2}$, which contradicts $\hat{v}_3 > \frac{1}{2}$, so nobody searches in this case. If $\alpha_H < \frac{1}{2}$, then $v_N \geq \frac{\hat{v}_3 - \alpha_H}{1 - 2\alpha_H} > \frac{\frac{1}{2} - \alpha_H}{1 - 2\alpha_H} = \frac{1}{2}$, which contradicts $v_N \leq 1/2$, so nobody searches in this case.

  If $v_N > \hat{v}_1$ and $\alpha_H > \frac{1}{2}$, then $v_N \leq \frac{\hat{v}_3 - \alpha_H}{1 - 2\alpha_H}$. It follows immediately from $\hat{v}_3 > \frac{1}{2}$ that $\frac{\hat{v}_3 - \alpha_H}{1 - 2\alpha_H} < \frac{1}{2}$. In order for this search condition to include some searchers, we need $\frac{\hat{v}_3 - \alpha_H}{1 - 2\alpha_H} > \hat{v}_1$. Since $\alpha_H > \frac{1}{2}$ and $\hat{v}_1 > \frac{1}{2}$, this inequality becomes $\alpha_H > \frac{\hat{v}_1 - \hat{v}_3}{2\hat{v}_1 - 1}$, and searchers exist only if $\frac{\hat{v}_1 - \hat{v}_3}{2\hat{v}_1 - 1} < 1$; this last inequality can be simplified to $\hat{v}_1 + \hat{v}_3 < 1$.

To summarize, when $v_N \in (\hat{v}_1, \frac{1}{2}]$, the searchers are $\left(\bar{v}_1, \frac{\hat{v}_3 - \alpha_H}{1-2\alpha_H}\right]$ if $\hat{v}_1 + \hat{v}_3 < 1$ and $\alpha_H > \frac{1}{2}$; the searchers are are $v_N \in \varnothing$ otherwise.

- If $v_N \leq \hat{v}_1$, this implies that the customer will search Period 2 after observing $m_1 = 10$. Solve $\mathbb{E}U_1^s \geq \mathbb{E}U_1^0$, and we get

$$y(\alpha_H) \cdot v_N \geq \hat{v}_4 + \frac{1}{2}(y(\alpha_H) - 1)$$

$$y(\alpha_H) \equiv \alpha_H(\alpha_H - 3) + 1$$

$$\hat{v}_4 \equiv \frac{1}{\tau}\left[p - v + \frac{1}{2}(3 - \alpha_H)s\right]$$

  The condition for $y(\alpha_H) > 0$ is $\alpha_H < \frac{1}{2}(3 - \sqrt{5})$, which is smaller than $\frac{1}{2}$. We can then look at the segments it creates separately.

  First, it's easy to see that $\frac{1}{2} < \hat{v}_3 < \hat{v}_4 < \hat{v}_2$.

  If $y(\alpha_H) = 0$, the inequality becomes $0 \geq \hat{v}_4 - \frac{1}{2}$; this contradicts $\hat{v}_4 > \frac{1}{2}$, which implies searchers don't exist. If $y(\alpha_H) > 0$, the inequality becomes $v_N \geq \frac{\hat{v}_4}{y} + \frac{1}{2}(1 - \frac{1}{y}) > \frac{1}{2}$; this contradicts $v_N \leq 1/2$, again implying searchers don't exist.

  If $y(\alpha_H) < 0$, the inequality becomes $v_N \leq \frac{\hat{v}_4}{y} + \frac{1}{2}(1 - \frac{1}{y})$. To ensure that searchers exist, combining $v_N \in [0, \hat{v}_1]$, we need $\frac{\hat{v}_4}{y} + \frac{1}{2}(1 - \frac{1}{y}) > 0$, meaning $2\hat{v}_4(\alpha_H) + y(\alpha_H) < 1$. Depending on the monotonicity of $2\hat{v}_4(\alpha_H) + y(\alpha_H)$ (undetermined), this implicit inequality can then give a value range for α_H, implied by Assumption 1.

From the above analysis, we know Period 1 searchers for $v_N \leq 1/2$ exist only when: (1) $\alpha_H > \frac{1}{2}(3 - \sqrt{5})$ and $2\hat{v}_4(\alpha_H) + y(\alpha_H) < 1$, and these searchers are $v_N \in \left[0, \min\left\{\hat{v}_1,\ \frac{\hat{v}_4}{y} + \frac{1}{2}(1 - \frac{1}{y})\right\}\right]$; (2) $\alpha_H > \frac{1}{2}$ and $\hat{v}_1 + \hat{v}_3 < 1$, and these searchers are $\left(\bar{v}_1, \frac{\hat{v}_3 - \alpha_H}{1-2\alpha_H}\right]$.

Next, we refine the search decision in Period 2 based on the segmentation of searchers and non-searchers for $v_N \leq 1/2$. In Period 2, from earlier analysis, we know the non-searchers will not search again. For the searchers, they will search if $m_1 = 10$ and $v_N \in [0, \hat{v}_1]$; this means all Period 1 searchers in $\left[0, \min\left\{\hat{v}_1,\ \frac{\hat{v}_4}{y} + \frac{1}{2}(1 - \frac{1}{y})\right\}\right]$ search again in Period 2 when

$m_1 = 10$.

To summarize the search dynamics, the total search declines in Period 2 compared to Period 1.

■

Proof of Result 7 (Truth-telling Product Launch)

We prove this result by verifying that firm I playing $m_t^* = h_t$ constitutes an equilibrium strategy in a separating equilibrium. We first show customers' optimal decisions under this strategy, and then show firm I does not have profitable off-equilibrium path deviations under these customer decisions. In Period 1, customers believe $r_H = \frac{1}{2}$, and the probability of discovery is $\frac{1}{2}$. In Period 2, a customer's belief is updated based on $m_1 \in \{00, 10, 11, 01, \varnothing\}$. If $m_1 = 00$, she knows $\theta = L$, and she will observe $m_2 = 11$ with probability 1 in Period 2. If $m_1 \in \{10, 11, 01\}$, she knows $\theta = H$, and she will at least see one N product in Period 2. If she did not search in Period 1, she still believes $r_H = 1/2$, and the probability of discovery is $1/2$.

Search decisions. Starting from customer's search decision in Period 2 conditional on m_1. If $m_1 = 00$, customers will not search in Period 2, and $\mathbb{E}U_2^0 = v_b - b$. If $m_1 \in \{10, 11, 01\}$, customers will not search again in Period 2, and $\mathbb{E}U_2^0 = v_b - b + v_N - p$. If $m_1 = \varnothing$, customers' utilities are $\mathbb{E}U_2^s = v_b - b - s + \frac{1}{2}(v_N - p)^+$ and $\mathbb{E}U_2^0 = v_b - b$. The search condition is $v_N \geq p + 2s$, conditional on this set including Period 1 non-searchers.

For Period 1 problem described in Equation 3.3, if we solve $\mathbb{E}U_1^s \geq \mathbb{E}U_1^0$, it gives $v_N \in [p + s, \bar{v}]$. This solution is due to derivations as follows. If $v_N \geq p + 2s$, the term of $\mathbb{E}U_1^0$ in the $[\cdot]^+$ operator returns non-negative, and the search condition is $v_N \geq p$. Combined with the range, the solution is $v_N \geq p + 2s$. If $v_N < p + 2s$, the $[\cdot]^+$ operator returns zero, and the search condition is $v_N \geq p + s$, so $p \in [p + s, p + 2s)$.

To refine the search condition in Period 2, since Period 1 non-searchers are $v_N \in [0, p + s)$, it contradicts the condition to search in Period 2 after $m_1 = \varnothing$ ($v_N \geq p + 2s$), so Period 1 non-searchers will not search in Period 2. In addition, Period 1 searchers will also not search in Period 2.

Equilibrium conditions. We can then derive a condition in which it's optimal for firm I not to deviate from playing m_t^*, given by $K \leq \hat{K}$, where

$$\hat{K} \equiv \min\left\{\frac{1}{2}(b+p), p\right\} \cdot \left[1 - \frac{1}{\bar{v}}(p+s)\right].$$

To see this, we first derive the demand functions D_b, D_N^I and D_N^E, represented by customer value segments, for each strategic path that involves a possible off-equilibrium path deviation. When $h_1 = m_1^* = 11$, these demand functions are $[0, p+s)$, $\frac{1}{2}[p+s, \bar{v}]$ and $\frac{1}{2}[p+s, \bar{v}]$. The likely deviation ($m_1' = 01$) leads to demand: $[0, p+s)$, $\varnothing$, $[p+s, \bar{v}]$.

When $h_1 = m_1^* = 10$, these demand functions are $[0, p+s)$, $[p+s, \bar{v}]$ and $\varnothing$. The likely deviation ($m_1' = 01$) leads to everyone buying $B : [0, \bar{v}]$. In Period 2, after $h_1 = m_1^* = 11$, $h_2 = m_2^* = 11$, then demand functions are: $[0, p+s)$, $\frac{1}{2}[p+s, \bar{v}]$ and $\frac{1}{2}[p+s, \bar{v}]$. If the deviation is $m_2' = 01$, then demand: $[0, \bar{v}]$, $\varnothing$, $\varnothing$. If the deviation is $m_1'01$, then demand: $[0, \bar{v}]$, $\varnothing$, $\varnothing$.

After $h_1 = m_1^* = 10$, $h_2 = m_2^* = 11$, then demand functions are: $[0, p+s)$, $[p+s, \bar{v}]$ and $\varnothing$. If the deviation is $m_2' = 01$, then demand: $[0, \bar{v}]$, $\varnothing$, $\varnothing$. If the deviation is $m_1' = 00$, then demand: $[0, \bar{v}]$, $\varnothing$, $\varnothing$. After $h_1 = m_1^* = 10$, $h_2 = m_2^* = 10$, then demand functions are: $[0, p+s)$, $[p+s, \bar{v}]$ and 0. If the deviation is $m_2' = 00$, then demand: $[0, \bar{v}]$, $\varnothing$, $\varnothing$. If the deviation is $m_1' = 00$, then demand: $[0, \bar{v}]$, $\varnothing$, $\varnothing$.

Equilibrium conditions: profit. We can then derive the profit functions under different strategies, the comparison of which gives us the equilibrium condition as above.

- For Period 1 deviations, we notice customers do not search in Period 2 and always repeat their Period 1 decisions. When $h_1 = m_1^* = 11$ and $m_1' = 01$, the equilibrium and deviation profit functions are $T\Pi_1^* = 2\frac{b}{\bar{v}}(p+s) + (b+p)\left[1 - \frac{1}{\bar{v}}(p+s)\right] - 2K$ and $\Pi_2' = \frac{b}{\bar{v}}(p+s)$, so firm I will not deviate if $K \leq \frac{1}{2}(b+p)\left[1 - \frac{1}{\bar{v}}(p+s)\right]$.

When $h_1 = m_1^* = 10$ and $m_1' = 00$; the equilibrium profit and the deviation profit are $T\Pi_2^* = 2b + 2p\left[1 - \frac{1}{\bar{v}}(p+s)\right] - 2K$ and $\Pi_2' = 2b$, so firm I will not deviate if $K \leq p\left[1 - \frac{1}{\bar{v}}(p+s)\right]$.

- For Period 2 deviations, when $h_1 = m_1^* = 11$ and $h_2 = m_2^* = 11$, we consider $m_2' = 01$; the equilibrium profit and the deviation profit are $\Pi_2^* = \frac{b}{\bar{v}}(p+s) + \frac{1}{2}(b+p)\left[1 - \frac{1}{\bar{v}}(p+s)\right] - K$ and $\Pi_2' = \frac{b}{\bar{v}}(p+s)$, so firm I will not deviate if $K \leq \frac{1}{2}(b+p)\left[1 - \frac{1}{\bar{v}}(p+s)\right]$.

  When $h_1 = m_1^* = 10$ and $h_2 = m_2^* = 11$, we consider $m_2' = 01$; the equilibrium profit and the deviation profit are $\Pi_2^* = b + p\left[1 - \frac{1}{\bar{v}}(p+s)\right] - K$ and $\Pi_2' = b$, so firm I will not deviate if $K \leq p\left[1 - \frac{1}{\bar{v}}(p+s)\right]$. When $h_1 = m_1^* = 10$ and $h_2 = m_2^* = 10$, $m_2' = 00$, the profit functions and derived condition are the same.

■

Proof of Result 8 (Strategic Withholding)

We prove this result by verifying that firm I playing the strategy detailed in Table 3.1 constitutes an equilibrium strategy in a semi-separating equilibrium. We first show customers' optimal decisions under this strategy, and then show firm I does not have profitable off-equilibrium path deviations under these customer decisions.

Customer's belief updates. In Period 1, the customer's beliefs are $r_H = r_L = 1/2$, and $\Pr(m_1 = 11) = \frac{1}{2}\alpha_H$, $\Pr(m_1 = 00) = r_L \Pr(m_1 = 00|L) + r_H \Pr(m_1 = 00|H) = 1/2 + 1/2 \cdot \frac{1}{2}\alpha_H = \frac{1}{2}(2 - \alpha_H)$. The probability of discovery is $\frac{1}{2}\alpha_H$.

In Period 2, the customer's beliefs conditional on m_1 are: If $m_1 = 11$, $\Pr(H|m_1 = 11) = 1$, and $m_2^* = h_2 = 11$ with probability 1. If $m_1 = 10$, $\Pr(H|m_1 = 10) = 1$, and she will observe $m_2^* = 01, 10$ ($h_2 = 11, 10$) with probabilities $\alpha_H, 1 - \alpha_H$, respectively; the probability of discovery is 1. If $m_1 = 01$, then $\Pr(H|m_1 = 01) = 1$, and $m_2^* = h_2 = 11$; the probability of discovery is 1. If $m_1 = \varnothing$, then $r_H = r_L = 1/2$, and $\Pr(m_2 = 11) = \frac{1}{2}\alpha_H$, $\Pr(m_2 = 00) =$

$1/2$, $\Pr(m_2 = 10) = \frac{1}{2}(1-\alpha_H)^2$, $\Pr(m_2 = 01) = \frac{1}{2}\alpha_H(1-\alpha_H)$; the probability of discovery is $\frac{1}{2}$.

Period 2 customer's beliefs after observing $m_1 = 00$ are $\Pr(H|m_1 = 00) = \frac{1-\alpha_H}{2-\alpha_H}$. The customer observes m_2 in Period 2 with probabilities $\Pr(m_2 = 00|m_1 = 00) = \Pr(h_1 = 00|m_1 = 00)\Pr(m_2 = 00|h_1 = 00) = \frac{1}{2-\alpha_H}$, $\Pr(m_2 = 10|m_1 = 00) = \Pr(h_1 = 10|m_1 = 00)\Pr(m_2 = 10|h_1 = 10) = \frac{(1-\alpha_H)^2}{2-\alpha_H}$, $\Pr(m_2 = 01|m_1 = 00) = \Pr(h_1 = 10|m_1 = 00)\Pr(m_2 = 01|h_1 = 10) = \frac{\alpha_H(1-\alpha_H)}{2-\alpha_H}$, $\Pr(m_2 = 11|m_1 = 00) = 0$, and the probability of discovery is $\frac{1-\alpha_H}{2-\alpha_H}$.

Search decisions in Period 2. A customer's search decision is contingent on m_1. If $m_1 = 11$, the customer believes $m_2^* = h_2 = 11$, and she will not search in Period 2 with period utility given by $\mathbb{E}U_2^0 = v_b - b + (v_N - p)^+$. If $m_1 = 10$, the probability of discovery is 1, and the probabilities of observing $m_2^* = 01, 10$ are α_H and $1-\alpha_H$; if she does not search, she can buy the discovered N_I with probability $1-\alpha_H$ ($m_2^* = 10$). Her period utilities thus are $\mathbb{E}U_2^s = v_b - b - s + (v_N - p)^+$ and $\mathbb{E}U_2^0 = v_b - b + (1-\alpha_H)(v_N - p)^+$. The search condition is $v_N \geq p + \frac{s}{\alpha_H}$. If $m_1 = 01$, then $\Pr(h_2 = 11) = 1$, and $m_2 = x1$, $x \in \{0, 1\}$, so she can always buy N_E even off the equilibrium path. So she will not search with period utility $\mathbb{E}U_2^0 = v_b - b + (v_N - p)^+$.

If $m_1 = 00$, the probability of discovery is $\frac{1-\alpha_H}{2-\alpha_H}$, and the customer's period utilities are $\mathbb{E}U_2^s = v_b - b - s + \frac{1-\alpha_H}{2-\alpha_H}(v_N - p)^+$ and $\mathbb{E}U_2^0 = v_b - b$. The search condition is $v_N \geq p + \frac{(2-\alpha_H)s}{1-\alpha_H}$.

If $m_1 = \varnothing$, the probability of discovery is $1/2$, and her period utilities are $\mathbb{E}U_2^s = v_b - b - s + \frac{1}{2}(v_N - p)^+$ and $\mathbb{E}U_2^0 = v_b - b$. The search condition is $v_N \geq p + 2s$. It follows:

$$p + \frac{s}{\alpha_H} \leq p + 2s \leq p + \frac{(2-\alpha_H)s}{1-\alpha_H},\ \alpha_H \in [1/2, 1]. \tag{C.1}$$

Search decisions in Period 1. In Period 1, the consumer's period utilities are given by Equations 3.6. We solve for search conditions by considering each segment created by the $[\cdot]^+$ operators.

The two option value terms (last terms) in above equations gives $v_N \geq p + \frac{(2-\alpha_H)s}{1-\alpha_H}$

and $v_N \geq p + 2s$, corresponding to the two search conditions in Period 2 when $m_1 =$ 00 and $\varnothing$. Since the semi-separating equilibrium could hold only for $\alpha_H \in [1/2, 1]$, and because of Equation C.1, we discuss all cases as follows. When $v_N \geq p + \frac{(2-\alpha_H)s}{1-\alpha_H}$, the search condition is $v_N \geq p + \frac{(2-\alpha_H)s}{\alpha_H}$. Considering the range, it becomes $v_N \in \left[p + \frac{(2-\alpha_H)s}{1-\alpha_H}, \bar{v}\right]$. When $v_N \in \left[p + 2s, p + \frac{(2-\alpha_H)s}{1-\alpha_H}\right)$, the solution to $\mathbb{E}U_2^s \geq \mathbb{E}U_2^0$ is $v_N \geq p$, which means the search condition is $v_N \in \left[p + 2s, p + \frac{(2-\alpha_H)s}{1-\alpha_H}\right)$. When $v_N \in [p, p + 2s)$, the solution is $v_N \geq p + \frac{s}{\alpha_H}$. Considering the range, the search condition is $v_N \in \left[p + \frac{s}{\alpha_H}, p + 2s\right]$. To summarize, the searchers are: $v_N \in \left[p + \frac{s}{\alpha_H}, \bar{v}\right]$.

To refine the search condition in Period 2, since Period 1 non-searchers are $v_N \in [0, p + \frac{s}{\alpha_H})$, it contradicts the condition to search in Period 2 after $m_1 = \varnothing$ ($v_N \geq p + 2s$), so Period 1 non-searchers will not search in Period 2. In addition, from Equation C.1, we also conclude that Period 1 searchers will search in Period 2 only if $m_1 = 10$, or $v_N \geq p + \frac{(2-\alpha_H)s}{1-\alpha_H}$ after $m_1 = 00$.

Equilibrium conditions: demand. Firm I's demand functions can be derived from the above discussion. We denote these demand functions (ranges) as D_b, D_N^I, D_N^E, and we use the customer's value ranges to characterize them for easy of analysis. We then drive these demand functions for cases where firm I has a possible off-equilibrium path deviation (when $\theta = H$). In Period 1, (D_b, D_N^I, D_N^E) are:

- $h_1 = 11$, $m_1^* = 11$, half of those who searched will by N from firm I, and (D_b, D_N^I, D_N^E) are $\left[0, p + \frac{s}{\alpha_H}\right)$, $\frac{1}{2} \cdot \left[p + \frac{s}{\alpha_H}, \bar{v}\right]$, $\frac{1}{2} \cdot \left[p + \frac{s}{\alpha_H}, \bar{v}\right]$. For the likely deviation, $m_1' = 01$, those who searched buy E : $\left[0, p + \frac{s}{\alpha_H}\right)$, $\varnothing$, $\left[p + \frac{s}{\alpha_H}, \bar{v}\right]$.

- $h_1 = 10$, $m_1^* = 00$, everyone buys b, $[0, \bar{v}]$, $\varnothing$, $\varnothing$. For the likely deviation, $m_1' = 10$, those who searched buy I: $\left[0, p + \frac{s}{\alpha_H}\right)$, $\left[p + \frac{s}{\alpha_H}, \bar{v}\right]$, $\varnothing$.

In Period 2, to these demand functions, first notice that the Period 1 non-searchers in will not search in Period 2. The Period 2 demand functions (D_b, D_N^I, D_N^E) are given by:

- $h_1 = m_1^* = 11$; $h_2 = m_2^* = 11$, no Period 1 searchers will search again, and keep Period

1 choice: $\left[0, p+\frac{s}{\alpha_H}\right), \frac{1}{2}\cdot\left[p+\frac{s}{\alpha_H}, \bar{v}\right], \frac{1}{2}\cdot\left[p+\frac{s}{\alpha_H}, \bar{v}\right]$. For the likely Period 1 deviation, $m_1' = 01$; $h_2 = m_2^* = 11$, Period 1 searchers will keep Period 1 choices and buy N_E: $\left[0, p+\frac{s}{\alpha_H}\right)$, $\varnothing$, $\left[p+\frac{s}{\alpha_H}, \bar{v}\right]$. For the likely Period 2 deviation, $h_2 = 11$, $m_2' = 01$, no Period 1 searchers will search again, and they all buy N_E since N_I is not offered: $\left[0, p+\frac{s}{\alpha_H}\right)$, $\varnothing$, $\left[p+\frac{s}{\alpha_H}, \bar{v}\right]$.

- $h_1 = 10$, $m_1^* = 00$; $h_2 = 11$, $m_2^* = 01$, all Period 2 searchers buy N_E : $\left[0, p+\frac{(2-\alpha_H)s}{1-\alpha_H}\right)$, $\varnothing$, $\left[p+\frac{(2-\alpha_H)s}{1-\alpha_H}, \bar{v}\right]$. For the likely Period 1 deviation, $m_1' = 10$; $h_2 = 11$, $m_2^* = 01$, all Period 2 searchers buy N_E since N_I is not offered: $\left[0, p+\frac{s}{\alpha_H}\right)$, $\varnothing$, $\left[p+\frac{s}{\alpha_H}, \bar{v}\right]$; notably, more people are informed of the existence of innovations than on the equilibrium path. For the likely Period 2 deviation, $h_2 = 11$, $m_2' = 11$, half of Period 2 searchers buy N_I: $\left[0, p+\frac{(2-\alpha_H)s}{1-\alpha_H}\right), \frac{1}{2}\cdot\left[p+\frac{(2-\alpha_H)s}{1-\alpha_H}, \bar{v}\right], \frac{1}{2}\cdot\left[p+\frac{(2-\alpha_H)s}{1-\alpha_H}, \bar{v}\right]$.

- $h_1 = 10$, $m_1^* = 00$; $h_2 = 10$, $m_2^* = 10$, all Period 2 searchers buy N_I: $\left[0, p+\frac{(2-\alpha_H)s}{1-\alpha_H}\right)$, $\left[p+\frac{(2-\alpha_H)s}{1-\alpha_H}, \bar{v}\right]$, $\varnothing$. For the likely Period 1 deviation, $m_1' = 10$, all Period 2 searchers buy N_I $\left[0, p+\frac{s}{\alpha_H}\right)$, $\left[p+\frac{s}{\alpha_H}, \bar{v}\right]$, $\varnothing$. For the likely Period 2 deviation, $h_2 = 10$, $m_2' = 00$, everyone can only buy B: $[0, \bar{v}], \varnothing, \varnothing$.

Equilibrium conditions: profit. Based on Equations 3.4, 3.5 and 3.7, we can derive firm I's profits. Then we can check all possible deviations to derive conditions under which the equilibrium can be sustained.

- When $h_1 = m_1^* = 11$ and $h_2 = m_2^* = 11$, firm I deviates to play $m_2' = 01$. The profits under the equilibrium and deviation strategies are $\Pi_2^* = b\cdot\frac{1}{\bar{v}}\left(p+\frac{s}{\alpha_H}\right)+\frac{1}{2}(b+p)\frac{1}{\bar{v}}\left[\bar{v}-p-\frac{s}{\alpha_H}\right]-K$ and $\Pi_2' = b\cdot\frac{1}{\bar{v}}\left(p+\frac{s}{\alpha_H}\right)$. The solution to $\Pi_2^* \geq \Pi_2'$ is $\alpha_H \geq \frac{1}{G_1}$, where
$$G_1 \equiv \frac{1}{s}\left(\bar{v}-p-\frac{2K\bar{v}}{b+p}\right).$$
To ensure $\alpha_H \in [1/2, 1]$, we have $G_1 > 1$.

- When $h_1 = 10$, $m_1^* = 00$ and $h_2 = 11, m_2^* = 01$, firm I deviates to play $m_2' = 11$. The profits under the equilibrium and deviation strategies are $\Pi_2^* = b \cdot \frac{1}{\bar{v}}\left[p + \frac{(2-\alpha_H)s}{1-\alpha_H}\right]$ and $\Pi_2' = b \cdot \frac{1}{\bar{v}}\left[p + \frac{(2-\alpha_H)s}{1-\alpha_H}\right] + \frac{1}{2}(b+p)\frac{1}{\bar{v}}\left[\bar{v} - p - \frac{(2-\alpha_H)s}{1-\alpha_H}\right] - K$. The solution to $\Pi_2^* \geq \Pi_2'$ is $\alpha_H \geq \frac{G_1-2}{G_1-1}$. To ensure $\alpha_H \in [1/2, 1]$, we again have $G_1 > 1$.

- When $h_1 = 10$, $m_1^* = 00$ and $h_2 = 10, m_2^* = 10$, firm I deviates to play $m_2' = 00$. The profits under these strategies are $\Pi_2^* = b + p\frac{1}{\bar{v}}\left[\bar{v} - p - \frac{(2-\alpha_H)s}{1-\alpha_H}\right] - K$ and $\Pi_2' = b$. The solution to $\Pi_2^* \geq \Pi_2'$ is $\alpha_H \leq \frac{G_2-2}{G_2-1}$, where

$$G_2 \equiv \frac{1}{s}\left(\bar{v} - p - \frac{K\bar{v}}{p}\right),$$

  and we need to have $G_2 > 3$ to ensure $\alpha_H \in [1/2, 1]$.

- When $h_1 = 10$, $m_1^* = 00$; firm I deviates to play $m_1' = 10$. The profits under these strategies are $T\Pi_1^* = b + \alpha_H b \cdot \frac{1}{\bar{v}}\left[p + \frac{(2-\alpha_H)s}{1-\alpha_H}\right] + (1-\alpha_H)\left[b + p \cdot \frac{1}{\bar{v}}\left(\bar{v} - p - \frac{(2-\alpha_H)s}{1-\alpha_H}\right) - K\right]$, and $T\Pi_1' = p \cdot \frac{1}{\bar{v}}\left[\bar{v} - p - \frac{s}{\alpha_H}\right] + \alpha_H b \cdot \frac{1}{\bar{v}}\left(p + \frac{s}{\alpha_H}\right) + (1-\alpha_H)\left[b + p \cdot \frac{1}{\bar{v}}\left(\bar{v} - p - \frac{s}{\alpha_H}\right) - K\right] + b - K$. Solve the inequality $T\Pi_1^* \geq T\Pi_1'$, it gives $\frac{1}{\alpha_H} + \left(\frac{2-\alpha_H}{1-\alpha_H} - \frac{1}{\alpha_H}\right)((q+1)\alpha_H - 1) \geq G_2$, where

$$q \equiv \frac{b}{p},\ G_2 \equiv \frac{1}{s}\left(\bar{v} - p - \frac{K\bar{v}}{p}\right),$$

  where we require $G_2 > 3$ from a former discussion. The solution to the inequality then gives that for $G_2 > 3$, there exists $\hat{q} > 3$, such that for any $\alpha_H > \frac{1}{2}$ and $q > \hat{q}$, firm I will not deviate. As G_2 increases, the minimum $\hat{q}$ also increases.

- When $h_1 = 11$, $m_1^* = 11$; firm I deviates to play $m_1' = 01$. The profits under these strategies are $T\Pi_1^* = 2b \cdot \frac{1}{\bar{v}}\left(p + \frac{s}{\alpha_H}\right) + (b+p)\frac{1}{\bar{v}}\left[\bar{v} - p - \frac{s}{\alpha_H}\right] - 2K$ and $T\Pi_1' = 2b \cdot \frac{1}{\bar{v}}\left(p + \frac{s}{\alpha_H}\right)$. Notice that this is the same problem as when $h_1 = m_1^* = 11$ and $h_2 = m_2^* = 11$, firm I deviates to play $m_2' = 01$. So the equilibrium condition is to have $\alpha_H \geq \frac{1}{G_1}$ and $G_1 > 1$, where $G_1 \equiv \frac{1}{s}\left(\bar{v} - p - \frac{2K\bar{v}}{b+p}\right)$.

To summarize, the existence of this withholding equilibrium requires α_H satisfy

$$\alpha_H \in \left[\max\left\{\frac{1}{2}, \frac{1}{G_1}, \frac{G_1-2}{G_1-1}\right\}, \frac{G_2-2}{G_2-1}\right],$$

where $G_1 > 1$ and $G_2 > 3$, and for a given G_3, we can find a $\hat{q}$ such that $q > \hat{q} > 3$.

■

References

[1] L. Li, W. Chu, J. Langford, and R. E. Schapire, "A contextual-bandit approach to personalized news article recommendation," in *Proceedings of the 19th international conference on World wide web*, ser. WWW '10, Raleigh, North Carolina, USA: Association for Computing Machinery, Apr. 2010, pp. 661–670.

[2] P. I. Frazier, W. B. Powell, and S. Dayanik, "A knowledge-gradient policy for sequential information collection," *SIAM Journal on Control and Optimization*, vol. 47, no. 5, pp. 2410–2439, Jan. 2008.

[3] S. E. Chick and K. Inoue, "New two-stage and sequential procedures for selecting the best simulated system," *Operations research*, vol. 49, no. 5, pp. 732–743, Oct. 2001.

[4] P. Frazier, W. Powell, and S. Dayanik, "The knowledge-gradient policy for correlated normal beliefs," *INFORMS journal on computing*, vol. 21, no. 4, pp. 599–613, Nov. 2009.

[5] W. B. Powell and I. O. Ryzhov, *Optimal Learning*, en. John Wiley & Sons, Apr. 2012.

[6] X. Liu, "Dynamic coupon targeting using batch deep reinforcement learning: An application to livestream shopping," *Marketing Science*, vol. 42, no. 4, pp. 637–658, Jul. 2023.

[7] O. Rafieian and H. Yoganarasimhan, "Targeting and privacy in mobile advertising," en, *Marketing science*, vol. 40, no. 2, pp. 193–218, Mar. 2021.

[8] D. Simester, A. Timoshenko, and S. I. Zoumpoulis, "Targeting prospective customers: Robustness of machine-learning methods to typical data challenges," *Management science*, vol. 66, no. 6, pp. 2495–2522, Jun. 2020.

[9] D. Simester, A. Timoshenko, and S. I. Zoumpoulis, "Efficiently evaluating targeting policies: Improving on champion vs. challenger experiments," *Management science*, vol. 66, no. 8, pp. 3412–3424, Aug. 2020.

[10] J.-P. Dubé and S. Misra, "Personalized pricing and consumer welfare," *The journal of political economy*, vol. 131, no. 1, pp. 131–189, Jan. 2023.

[11] S. Fujimoto, D. Meger, and D. Precup, "Off-policy deep reinforcement learning without exploration," in *Proceedings of the 36th International Conference on Machine Learning*, K. Chaudhuri and R. Salakhutdinov, Eds., ser. Proceedings of Machine Learning Research, vol. 97, PMLR, 2019, pp. 2052–2062.

[12] N. Si, F. Zhang, Z. Zhou, and J. Blanchet, "Distributionally robust batch contextual bandits," *Management science*, vol. 69, no. 10, pp. 5772–5793, Oct. 2023.

[13] D. Dzyabura and J. R. Hauser, "Active machine learning for consideration heuristics," *Marketing Science*, vol. 30, no. 5, pp. 801–819, Sep. 2011.

[14] J. C. Gittins, "Bandit processes and dynamic allocation indices," en, *Journal of the Royal Statistical Society. Series B, Statistical methodology*, vol. 41, no. 2, pp. 148–164, Jan. 1979.

[15] A. Krause and C. S. Ong, "Contextual gaussian process bandit optimization," in *Advances in Neural Information Processing Systems 24*, J. Shawe-Taylor, R. S. Zemel, P. L. Bartlett, F. Pereira, and K. Q. Weinberger, Eds., Curran Associates, Inc., 2011, pp. 2447–2455.

[16] T. L. Lai and H. Robbins, "Asymptotically efficient adaptive allocation rules," en, *Advances in applied mathematics*, vol. 6, no. 1, pp. 4–22, Mar. 1985.

[17] S. Agrawal and N. Goyal, "Near-optimal regret bounds for thompson sampling," *Journal of the ACM*, vol. 64, no. 5, pp. 1–24, Sep. 2017.

[18] E. M. Schwartz, E. T. Bradlow, and P. S. Fader, "Customer acquisition via display advertising using multi-armed bandit experiments," *Marketing Science*, vol. 36, no. 4, pp. 500–522, Jul. 2017.

[19] I. Weaver and V. Kumar, "Nonparametric bandits leveraging informational externalities to learn the demand curve," en, *SSRN Electronic Journal*, 2022.

[20] J. R. Hauser, G. L. Urban, G. Liberali, and M. Braun, "Website morphing," *Marketing Science*, vol. 28, no. 2, pp. 202–223, Mar. 2009.

[21] S. Lin, J. Zhang, and J. R. Hauser, "Learning from experience, simply," *Marketing Science*, vol. 34, no. 1, pp. 1–19, Jan. 2015.

[22] K. Misra, E. M. Schwartz, and J. Abernethy, "Dynamic online pricing with incomplete information using multiarmed bandit experiments," *Marketing Science*, vol. 38, no. 2, pp. 226–252, Mar. 2019.

[23] S. S. Tehrani and A. T. Ching, "A heuristic approach to explore: The value of perfect information," *Management science*, Jul. 2023.

[24] A. Alban, S. E. Chick, and S. I. Zoumpoulis, "Learning personalized treatment strategies with predictive and prognostic covariates in adaptive clinical trials," *Available at SSRN 4160045*, 2022.

[25] L. Ding, L. J. Hong, H. Shen, and X. Zhang, "Technical note—knowledge gradient for selection with covariates: Consistency and computation," en, *Naval Research Logistics*, vol. 69, no. 3, pp. 496–507, Apr. 2022.

[26] J. Wu and P. Frazier, "Practical two-step lookahead bayesian optimization," in *Advances in Neural Information Processing Systems*, vol. 32, 2019, pp. 9813–9823.

[27] G. L. Urban and J. R. Hauser, ““listening in” to find and explore new combinations of customer needs,” *Journal of marketing*, vol. 68, no. 2, pp. 72–87, Apr. 2004.

[28] O. Toubia, D. I. Simester, J. R. Hauser, and E. Dahan, “Fast polyhedral adaptive conjoint estimation,” *Marketing Science*, vol. 22, no. 3, pp. 273–303, Aug. 2003.

[29] O. Rafieian, “Optimizing user engagement through adaptive ad sequencing,” *Marketing Science*, vol. 42, no. 5, pp. 910–933, Sep. 2023.

[30] O. Toubia, E. Johnson, T. Evgeniou, and P. Delquié, “Dynamic experiments for estimating preferences: An adaptive method of eliciting time and risk parameters,” *Management science*, vol. 59, no. 3, pp. 613–640, Mar. 2013.

[31] D. I. Simester, P. Sun, and J. N. Tsitsiklis, “Dynamic catalog mailing policies,” *Management science*, vol. 52, no. 5, pp. 683–696, May 2006.

[32] F. Gönül and M. Z. Shi, “Optimal mailing of catalogs: A new methodology using estimable structural dynamic programming models,” *Management science*, vol. 44, no. 9, pp. 1249–1262, Sep. 1998.

[33] T. Desautels, A. Krause, and J. W. Burdick, “Parallelizing exploration-exploitation tradeoffs in gaussian process bandit optimization,” *Journal of Machine Learning Research*, vol. 15, pp. 3873–3923, 2014.

[34] Z. Gao, Y. Han, Z. Ren, and Z. Zhou, “Batched multi-armed bandits problem,” in *Advances in Neural Information Processing Systems*, vol. 32, 2019, pp. 503–513.

[35] A. Karbasi, V. Mirrokni, and M. Shadravan, “Parallelizing thompson sampling,” *Advances in neural information processing systems*, vol. 34, pp. 10 535–10 548, 2021.

[36] V. Perchet, P. Rigollet, S. Chassang, and E. Snowberg, “Batched bandit problems,” en, *The Annals of Statistics*, vol. 44, no. 2, pp. 660–681, Apr. 2016.

[37] Z. Wang, C. Li, S. Jegelka, and P. Kohli, “Batched high-dimensional bayesian optimization via structural kernel learning,” in *Proceedings of the 34th International Conference on Machine Learning - Volume 70*, ser. ICML’17, Sydney, NSW, Australia: JMLR.org, Aug. 2017, pp. 3656–3664.

[38] J. Wang, S. C. Clark, E. Liu, and P. I. Frazier, “Parallel bayesian global optimization of expensive functions,” *Operations research*, vol. 68, no. 6, pp. 1850–1865, Nov. 2020.

[39] J. Wu and P. Frazier, “The parallel knowledge gradient method for batch bayesian optimization,” in *Advances in Neural Information Processing Systems 29*, D. D. Lee, M. Sugiyama, U. V. Luxburg, I. Guyon, and R. Garnett, Eds., Curran Associates, Inc., 2016, pp. 3126–3134.

[40] V. Hadad, D. A. Hirshberg, R. Zhan, S. Wager, and S. Athey, “Confidence intervals for policy evaluation in adaptive experiments,” en, *Proceedings of the National Academy of Sciences of the United States of America*, vol. 118, no. 15, Apr. 2021.

[41] M. Tabord-Meehan, “Stratification trees for adaptive randomisation in randomised controlled trials,” *Review of Economic Studies*, vol. 90, no. 5, pp. 2646–2673, 2023.

[42] Z. Ren, Z. Zhou, and J. R. Kalagnanam, “Batched learning in generalized linear contextual bandits with general decision sets,” *IEEE Control Systems Letters*, vol. 6, pp. 37–42, 2022.

[43] M. Joo, M. L. Thompson, and G. M. Allenby, “Optimal product design by sequential experiments in high dimensions,” *Management science*, vol. 65, no. 7, pp. 3235–3254, Jul. 2019.

[44] A. N. Elmachtoub and P. Grigas, “Smart “predict, then optimize”,” *Management science*, vol. 68, no. 1, pp. 9–26, Jan. 2022.

[45] E. M. Feit and R. Berman, “Test & roll: Profit-maximizing a/b tests,” *Marketing Science*, vol. 38, no. 6, pp. 1038–1058, Nov. 2019.

[46] D. Simester, A. Timoshenko, and S. Zoumpoulis, "A sample size calculation for training and certifying targeting policies," en, *SSRN Electronic Journal*, 2022.

[47] H. Yoganarasimhan, E. Barzegary, and A. Pani, "Design and evaluation of optimal free trials," en, *Management science*, vol. 69, no. 6, pp. 3220–3240, Jun. 2023.

[48] C. E. Rasmussen and C. K. I. Williams, *Gaussian processes for machine learning* (Adaptive Computation and Machine Learning Series). London, England: MIT Press, Jun. 2019, vol. 2.

[49] C. Robert, *The Bayesian Choice: From Decision-Theoretic Foundations to Computational Implementation*, en. Springer Science & Business Media, Aug. 2007.

[50] O. Saarela, L. R. Belzile, and D. A. Stephens, "A bayesian view of doubly robust causal inference," en, *Biometrika*, vol. 103, no. 3, pp. 667–681, Jul. 2016.

[51] J. R. Hauser and O. Toubia, "The impact of utility balance and endogeneity in conjoint analysis," *Marketing Science*, vol. 24, no. 3, pp. 498–507, Aug. 2005.

[52] Q. Liu, T. Otter, and G. M. Allenby, "Investigating endogeneity bias in marketing," *Marketing Science*, vol. 26, no. 5, pp. 642–650, Sep. 2007.

[53] C. Olivier, S. Bernhard, and Z. Alexander, "A discussion of semi-supervised learning and transduction," in *Semi-Supervised Learning*, London, England: The MIT Press, Sep. 2006, pp. 473–478.

[54] X. Zhu, "Semi-supervised learning tutorial," in *International Conference on Machine Learning (ICML)*, 2007, pp. 1–135.

[55] R. S. Sutton and A. G. Barto, *Reinforcement Learning, second edition: An Introduction*, en. MIT Press, Nov. 2018.

[56] M. Hu, Z. C. Rich, D. Luo, and S. Xiao, "Cigarette sharing and gifting in rural china: A focus group study," *Nicotine & Tobacco Research*, vol. 14, no. 3, pp. 361–367, 2012.

[57] Z. C. Rich and S. Xiao, "Tobacco as a social currency: Cigarette gifting and sharing in china," *Nicotine & Tobacco Research*, vol. 14, no. 3, pp. 258–263, 2011.

[58] J. L. Townsend, "Cigarette tax, economic welfare and social class patterns of smoking," *Applied Economics*, vol. 19, no. 3, pp. 355–365, 1987.

[59] S. T. Yen, "Zero observations and gender differences in cigarette consumption," *Applied Economics*, vol. 37, no. 16, pp. 1839–1849, 2005.

[60] H. Cai, H. Fang, and L. C. Xu, "Eat, drink, firms, government: An investigation of corruption from the entertainment and travel costs of chinese firms," *The Journal of Law and Economics*, vol. 54, no. 1, pp. 55–78, 2011.

[61] S. H. Park and Y. Luo, "Guanxi and organizational dynamics: Organizational networking in chinese firms," *Strategic management journal*, vol. 22, no. 5, pp. 455–477, 2001.

[62] P. Bardhan, "Corruption and development: A review of issues," *Journal of economic literature*, vol. 35, no. 3, pp. 1320–1346, 1997.

[63] G. Hofstede, *Culture's consequences: Comparing values, behaviors, institutions and organizations across nations*. Sage publications, 2001.

[64] H. Takada and D. Jain, "Cross-national analysis of diffusion of consumer durable goods in pacific rim countries," *Journal of Marketing*, vol. 55, no. 2, pp. 48–54, 1991.

[65] G. J. Tellis, J. C. Prabhu, and R. K. Chandy, "Radical innovation across nations: The preeminence of corporate culture," *Journal of marketing*, vol. 73, no. 1, pp. 3–23, 2009.

[66] S. Dwyer, H. Mesak, and M. Hsu, "An exploratory examination of the influence of national culture on cross-national product diffusion," *Journal of International Marketing*, vol. 13, no. 2, pp. 1–27, 2005.

[67] J. L. Hamilton, “The demand for cigarettes: Advertising, the health scare, and the cigarette advertising ban,” *The Review of Economics and Statistics*, pp. 401–411, 1972.

[68] S. L. Holak and S. K. Reddy, “Effects of a television and radio advertising ban: A study of the cigarette industry,” *The Journal of Marketing*, pp. 219–227, 1986.

[69] T. E. Keeler, T.-w. Hu, M. Ong, and H.-Y. Sung, “The us national tobacco settlement: The effects of advertising and price changes on cigarette consumption,” *Applied Economics*, vol. 36, no. 15, pp. 1623–1629, 2004.

[70] A. E. Tuchman, “Advertising and demand for addictive goods: The effects of e-cigarette advertising,” working paper, Tech. Rep., 2016.

[71] K. E. Warner, “Cigarette smoking in the 1970’s: The impact of the antismoking campaign on consumption,” *Science*, vol. 211, no. 4483, pp. 729–731, 1981.

[72] R. Borland, S. Chapman, N. Owen, and D. Hill, “Effects of workplace smoking bans on cigarette consumption.,” *American Journal of Public Health*, vol. 80, no. 2, pp. 178–180, 1990.

[73] A. Bonfrer, P. K. Chintagunta, J. H. Roberts, and D. Corkindale, “Assessing the sales impact of plain packaging regulation for cigarettes: Evidence from australia,” *Marketing Science*, 2019.

[74] S. James and T. Lahti, “Taxes, cigarette consumption, and smoking intensity,” *The American Economic Review*, vol. 96, no. 4, pp. 1013–1028, 2006.

[75] J. Adda and F. Cornaglia, “The effect of bans and taxes on passive smoking,” *American Economic Journal: Applied Economics*, vol. 2, no. 1, pp. 1–32, 2010.

[76] M. S. Amato, R. G. Boyle, and B. Brock, “Higher price, fewer packs: Evaluating a tobacco tax increase with cigarette sales data,” *American journal of public health*, vol. 105, no. 3, e5–e8, 2015.

[77] H. Liu, J. A. Rizzo, Q. Sun, and F. Wu, "How do smokers respond to cigarette taxes? evidence from china's cigarette industry," *Health economics*, vol. 24, no. 10, pp. 1314–1330, 2015.

[78] R. W. Belk, "Gift-giving behavior/bebr no. 449," *Faculty working papers; no. 449-450*, 1977.

[79] A. Joy, "Gift giving in hong kong and the continuum of social ties," *Journal of Consumer Research*, vol. 28, no. 2, pp. 239–256, 2001.

[80] Z. C. Rich, M. Hu, and S. Xiao, "Gifting and sharing cigarettes in a rural chinese village: A cross-sectional study," *Tobacco control*, vol. 23, no. 6, pp. 496–500, 2014.

[81] D. Kaufmann, "Can corruption adversely affect public finances in industrialized countries?" *Brookings Institution, April*, vol. 19, 2010.

[82] L. K. Cheng and Y. K. Kwan, "What are the determinants of the location of foreign direct investment? the chinese experience," *Journal of international economics*, vol. 51, no. 2, pp. 379–400, 2000.

[83] O. Belkhodja, M. Mohiuddin, and E. Karuranga, "The determinants of fdi location choice in china: A discrete-choice analysis," *Applied Economics*, vol. 49, no. 13, pp. 1241–1254, 2017.

[84] B. Smarzynska Javorcik, "Does foreign direct investment increase the productivity of domestic firms? in search of spillovers through backward linkages," *American economic review*, vol. 94, no. 3, pp. 605–627, 2004.

[85] Z. Liu, "Foreign direct investment and technology spillovers: Theory and evidence," *Journal of Development Economics*, vol. 85, no. 1-2, pp. 176–193, 2008.

[86] H. Cai, Y. Chen, and L.-A. Zhou, "Income and consumption inequality in urban china: 1992–2003," *Economic Development and Cultural Change*, vol. 58, no. 3, pp. 385–413, 2010.

[87] L. L. Tsai and Y. Xu, "Outspoken insiders: Political connections and citizen participation in authoritarian china," *Political Behavior*, vol. 40, no. 3, pp. 629–657, 2018.

[88] B. J. Bronnenberg, J.-P. H. Dubé, and M. Gentzkow, "The evolution of brand preferences: Evidence from consumer migration," *The American Economic Review*, vol. 102, no. 6, pp. 2472–2508, 2012.

[89] N. Qian and J. Wen, "The impact of xi jinping's anti-corruption campaign on luxury imports in china," 2015.

[90] X. Lan and W. Li, "Swiss watch cycles: Evidence of corruption during leadership transition in china," *Journal of Comparative Economics*, vol. 46, no. 4, pp. 1234–1252, 2018.

[91] R. Fisman, "Estimating the value of political connections," *The American economic review*, vol. 91, no. 4, pp. 1095–1102, 2001.

[92] B. A. Olken, "Monitoring corruption: Evidence from a field experiment in indonesia," *Journal of political Economy*, vol. 115, no. 2, pp. 200–249, 2007.

[93] Y. Gorodnichenko and K. S. Peter, "Public sector pay and corruption: Measuring bribery from micro data," *Journal of Public economics*, vol. 91, no. 5, pp. 963–991, 2007.

[94] H. Cai, J. V. Henderson, and Q. Zhang, "China's land market auctions: Evidence of corruption?" *The Rand journal of economics*, vol. 44, no. 3, pp. 488–521, 2013.

[95] H. Fang, Q. Gu, and L.-A. Zhou, "The gradients of power: Evidence from the chinese housing market," National Bureau of Economic Research, Tech. Rep., 2014.

[96] H. Ding, H. Fang, S. Lin, and K. Shi, "Equilibrium consequences of corruption on firms: Evidence from china's anti-corruption campaign," 2017.

[97] D. Ma and M. R. Pang, "The rise and fall of electoral democracy: A social evolutionary approach to direct election experiments in local china," *Journal of Chinese Political Science*, vol. 22, no. 4, pp. 601–624, 2017.

[98] C. M. Christensen, *The Innovator's Dilemma: When New Technologies Cause Great Firms to Fail*, en. Harvard Business Review Press, Oct. 2013.

[99] G. M. Grossman and E. Helpman, "Quality ladders in the theory of growth," *The Review of economic studies*, vol. 58, no. 1, pp. 43–61, Jan. 1991.

[100] P. Aghion and P. Howitt, "A model of growth through creative destruction," *Econometrica: journal of the Econometric Society*, vol. 60, no. 2, pp. 323–351, 1992.

[101] P. Aghion, U. Akcigit, and P. Howitt, "The schumpeterian growth paradigm," *Annual review of economics*, vol. 7, no. Volume 7, 2015, pp. 557–575, Aug. 2015.

[102] S. (Ansari and P. Krop, "Incumbent performance in the face of a radical innovation: Towards a framework for incumbent challenger dynamics," *Research policy*, vol. 41, no. 8, pp. 1357–1374, Oct. 2012.

[103] H. Chun, J.-W. Kim, R. Morck, and B. Yeung, "Creative destruction and firm-specific performance heterogeneity," *Journal of financial economics*, vol. 89, no. 1, pp. 109–135, Jul. 2008.

[104] T. Bresnahan, S. Greenstein, and R. Henderson, "Schumpeterian competition and diseconomies of scope; illustrations from the histories of microsoft and ibm," *ERN: Technology (Topic)*, Jan. 2011.

[105] B. Wernerfelt, "When does the underdog win?" *Strategy Science*, vol. 7, no. 1, pp. 1–9, Mar. 2022.

[106] B. F. Jones, "The burden of knowledge and the "death of the renaissance man": Is innovation getting harder?" en, *The Review of economic studies*, vol. 76, no. 1, pp. 283–317, Jan. 2009.

[107] N. Bloom, C. I. Jones, J. Van Reenen, and M. Webb, "Are ideas getting harder to find?" *The American economic review*, vol. 110, no. 4, pp. 1104–1144, Apr. 2020.

[108] A. Bell, R. Chetty, X. Jaravel, N. Petkova, and J. Van Reenen, "Who becomes an inventor in america? the importance of exposure to innovation," en, *The quarterly journal of economics*, vol. 134, no. 2, pp. 647–713, May 2019.

[109] T. Fitzgerald, B. Balsmeier, L. Fleming, and G. Manso, "Innovation search strategy and predictable returns," *Management science*, vol. 67, no. 2, pp. 1109–1137, Feb. 2021.

[110] J. Miklós-Thal, M. Raith, and M. Selove, "What are we really good at? product strategy with uncertain capabilities," *Marketing Science*, vol. 37, no. 2, pp. 294–309, Mar. 2018.

[111] M. L. Weitzman, "Optimal search for the best alternative," *Econometrica (pre-1986)*, vol. 47, no. 3, p. 641, 1979.

[112] T. T. Ke, Z.-J. M. Shen, and J. M. Villas-Boas, "Search for information on multiple products," *Management Science*, vol. 62, no. 12, pp. 3576–3603, 2016.

[113] P. A. Diamond, "Search, sticky prices, and inflation," *The Review of Economic Studies*, vol. 60, no. 1, pp. 53–68, 1993.

[114] M. Meurer and D. O. Stahl, "Informative advertising and product match," *International Journal of Industrial Organization*, vol. 12, no. 1, pp. 1–19, Mar. 1994.

[115] R. P. Greminger, "Optimal search and discovery," *Management science*, vol. 68, no. 5, pp. 3904–3924, May 2022.

[116] C. R. Clark, "Advertising restrictions and competition in the children's breakfast cereal industry," *The Journal of Law and Economics*, vol. 50, no. 4, pp. 757–780, Nov. 2007.

[117] A. Barroso and G. Llobet, “Advertising and consumer awareness of new, differentiated products,” *JMR, Journal of marketing research*, vol. 49, no. 6, pp. 773–792, Dec. 2012.

[118] D. Dzyabura and J. R. Hauser, “Recommending products when consumers learn their preference weights,” *Marketing Science*, vol. 38, no. 3, pp. 417–441, May 2019.

[119] J. R. Hauser and B. Wernerfelt, “An evaluation cost model of consideration sets,” en, *The Journal of consumer research*, vol. 16, no. 4, pp. 393–408, Mar. 1990.

[120] J. H. Roberts and J. M. Lattin, “Development and testing of a model of consideration set composition,” *JMR, Journal of marketing research*, vol. 28, no. 4, pp. 429–440, Nov. 1991.

[121] N. Mehta, S. Rajiv, and K. Srinivasan, “Price uncertainty and consumer search: A structural model of consideration set formation,” *Marketing Science*, vol. 22, no. 1, pp. 58–84, Feb. 2003.

[122] J. M. Villas-Boas, “Communication strategies and product line design,” *Marketing Science*, vol. 23, no. 3, pp. 304–316, Aug. 2004.

[123] L. Liu and A. Dukes, “Consideration set formation with multiproduct firms: The case of within-firm and across-firm evaluation costs,” *Management science*, vol. 59, no. 8, pp. 1871–1886, Aug. 2013.

[124] A. Caplin, M. Dean, and J. Leahy, “Rational inattention, optimal consideration sets, and stochastic choice,” *The Review of economic studies*, vol. 86, no. 3, pp. 1061–1094, May 2019.

[125] V. Mahajan, E. Muller, and F. M. Bass, “New product diffusion models in marketing: A review and directions for research,” *Journal of marketing*, vol. 54, no. 1, pp. 1–26, Jan. 1990.

[126] V. Mahajan and E. Muller, “Advertising pulsing policies for generating awareness for new products,” *Marketing Science*, vol. 5, no. 2, pp. 89–106, May 1986.

[127] V. Mahajan, E. Muller, and S. Sharma, “An empirical comparison of awareness forecasting models of new product introduction,” *Marketing Science*, vol. 3, no. 3, pp. 179–197, Aug. 1984.

[128] R. Peres, E. Muller, and V. Mahajan, “Innovation diffusion and new product growth models: A critical review and research directions,” *International Journal of Research in Marketing*, vol. 27, no. 2, pp. 91–106, Jun. 2010.

[129] B. N. Anand and R. Shachar, “Advertising, the matchmaker,” en, *The Rand journal of economics*, vol. 42, no. 2, pp. 205–245, Jun. 2011.

[130] D. Mayzlin and J. Shin, “Uninformative advertising as an invitation to search,” *Marketing Science*, vol. 30, no. 4, pp. 666–685, Jul. 2011.

[131] P. M. Gardete, “Cheap-talk advertising and misrepresentation in vertically differentiated markets,” *Marketing Science*, vol. 32, no. 4, pp. 609–621, Jul. 2013.

[132] J. Miklós-Thal and J. Zhang, “(de)marketing to manage consumer quality inferences,” *JMR, Journal of marketing research*, vol. 50, no. 1, pp. 55–69, Feb. 2013.

[133] D. Kuksov, R. Shachar, and K. Wang, “Advertising and consumers’ communications,” *Marketing Science*, vol. 32, no. 2, pp. 294–309, Mar. 2013.

[134] E. Kamenica and M. Gentzkow, “Bayesian persuasion,” *The American economic review*, vol. 101, no. 6, pp. 2590–2615, Oct. 2011.

[135] T. T. Ke, S. Lin, and M. Y. Lu, “Information design of online platforms,” *Working paper*, 2022.

[136] J. I. Bulow, J. D. Geanakoplos, and P. D. Klemperer, “Multimarket oligopoly: Strategic substitutes and complements,” *The journal of political economy*, vol. 93, no. 3, pp. 488–511, 1985.

[137] S. C. Salop, “Strategic entry deterrence,” *The American economic review*, vol. 69, no. 2, pp. 335–338, 1979.

[138] H. C. Hoppe and I. H. Lee, “Entry deterrence and innovation in durable-goods monopoly,” *European economic review*, vol. 47, no. 6, pp. 1011–1036, Dec. 2003.

[139] R. Schmalensee, “Advertising and entry deterrence: An exploratory model,” *The journal of political economy*, vol. 91, no. 4, pp. 636–653, Aug. 1983.

[140] U. Doraszelski and S. Markovich, “Advertising dynamics and competitive advantage,” en, *The Rand journal of economics*, vol. 38, no. 3, pp. 557–592, Jan. 2008.

[141] G. Ellison and S. F. Ellison, “Strategic entry deterrence and the behavior of pharmaceutical incumbents prior to patent expiration,” *American Economic Journal: Microeconomics*, vol. 3, no. 1, pp. 1–36, Feb. 2011.

[142] M. Armstrong and J. Zhou, “Search deterrence,” *The Review of economic studies*, vol. 83, no. 1 (294), pp. 26–57, 2016.

[143] P. Klemperer, “Entry deterrence in markets with consumer switching costs,” en, *The Economic Journal*, vol. 97, no. Supplement, pp. 99–117, Dec. 1987.

[144] V. Grossmann, “Advertising, in-house r&d, and growth,” *Oxford economic papers*, vol. 60, no. 1, pp. 168–191, Jan. 2008.

[145] O. Turut and E. Ofek, “Innovation strategy and entry deterrence,” en, *Journal of Economics & Management Strategy*, vol. 21, no. 3, pp. 583–631, Sep. 2012.

[146] R. J. Gilbert and D. M. G. Newbery, “Preemptive patenting and the persistence of monopoly,” *The American economic review*, vol. 72, no. 3, pp. 514–526, 1982.

[147] C. Cunningham, F. Ederer, and S. Ma, “Killer acquisitions,” *The journal of political economy*, vol. 129, no. 3, pp. 649–702, Mar. 2021.

[148] S. Shavell and T. van Ypersele, “Rewards versus intellectual property rights,” *The Journal of Law and Economics*, vol. 44, no. 2, pp. 525–547, Oct. 2001.

[149] B. Sampat and H. L. Williams, “How do patents affect follow-on innovation? evidence from the human genome,” en, *The American economic review*, vol. 109, no. 1, pp. 203–236, 2019.

[150] A. Galasso and H. Luo, “Product liability litigation and innovation: Evidence from medical devices,” National Bureau of Economic Research, Tech. Rep. w32215, Mar. 2024.

[151] M. A. Heller and R. S. Eisenberg, “Can patents deter innovation? the anticommons in biomedical research,” en, *Science*, vol. 280, no. 5364, pp. 698–701, May 1998.

[152] E. Budish, B. N. Roin, and H. Williams, “Do firms underinvest in long-term research? evidence from cancer clinical trials,” en, *The American economic review*, vol. 105, no. 7, pp. 2044–2085, Jul. 2015.

[153] J. Krieger, D. Li, and D. Papanikolaou, “Missing novelty in drug development,” en, *The review of financial studies*, vol. 35, no. 2, pp. 636–679, Mar. 2021.