3-2022

Consumer search and marketing actions in retailing

Yi PENG
Singapore Management University, yi.peng.2017@pbs.smu.edu.sg

Cover Page (PhD Dissertation)

CONSUMER SEARCH AND MARKETING ACTIONS IN RETAILING

YI PENG

SINGAPORE MANAGEMENT UNIVERSITY

2022

Title Page (PhD Dissertation)

Consumer Search and Marketing Actions in Retailing

Yi Peng

Submitted to Lee Kong Chian School of Business
in partial fulfillment of the requirements for the
Degree of Doctor of Philosophy in Business (Marketing)

Dissertation Committee:

Sandeep R. Chandukala (Supervisor / Chair)
Associate Professor of Marketing
Singapore Management University

Kapil R. Tuli
Lee Kong Chian Professor of Marketing
Singapore Management University

Hulya Karaman
Assistant Professor of Marketing
Singapore Management University

Girish Mallapragada
Associate Professor of Marketing
Indiana University

Singapore Management University
2022

Declaration Page (PhD Dissertation)

I hereby declare that this PhD dissertation is my original work
and it has been written by me in its entirety.
I have duly acknowledged all the sources of information
which have been used in this dissertation.

This PhD dissertation has also not been submitted for any degree
in any university previously.

YI PENG

Yi Peng

06 April 2022

Abstract

This dissertation seeks to gain insight into the critical roles of consumers and marketers in a retail context using a variety of unique and rich data sources (e.g., tracking data, retail scanner data, ad intel data and publicly available data). The main aim of the two essays is to focus on unique aspects of retail analytics.

The first essay examines how consumers conduct haptic search to make purchase decisions using a unique dataset collected by the state-of-the-art sensing technology. This research contributes to the literature by defining key attributes of the shoppers' speed, consideration set, and shopping path at the shelf space and investigating the effects of consumer haptic search on price paid across food and non-food categories. This paper further provides managerial implications regarding in-store category management and shelf layout.

The second essay investigates the spillover effects of recreational cannabis legalization (RCL) on related categories (i.e., alcohol, tobacco, candy, and salty snacks) using secondary data (Nielsen retailer scanner data and ad intel data). This study employs synthetic control method to show that RCL resulted in an increase in per capita dollar sales and per capita unit sales of alcohol, salty snacks, and candy, while this was not observed for tobacco sales. To rule out alternative explanations, this work identifies a "null category" (i.e., batteries) and demonstrates that RCL did not lead to changes in pricing or advertising. The findings are likely to help policymakers in understanding unintended consequences and potential problems associated with RCL such as excessive drinking and junk food consumption resulting in increasing health care expenses.

Keywords: retail analytics; haptic search; RCL

Table of Contents

Acknowledgement

First of all, I am extremely grateful to my supervisor Prof. Chandukala for his helpful guidance and tremendous support. I feel so fortunate to have such a wonderful advisor who is knowledgeable and caring. I highly appreciate all the research resources he has offered and the huge amount of time he has spent on our discussions.

Second, I would like to express my sincere thanks to Prof. Tuli for his treasured support at every stage of my PhD journey. I am deeply grateful for his guidance on research skills and career development.

Third, I would like to thank Prof. Karaman and Prof. Mallapragada for their insightful comments and suggestions which significantly improved my thesis. Additionally, I would like to offer my special thanks to all the faculty, admin staff and PhD students from the School of Business.

Finally, I would like to express my sincere gratitude to my parents and my husband for their continuous support.

Chapter 1 Introduction

The rapid growth of digital businesses and technology has shaped the retail sector and turned it into an increasingly data rich environment (PwC 2022). Given the complicated issues faced by the industry (e.g., competition from digitalization and product launches, highly informed consumers), retailers should boost performance through leveraging and harnessing data which they own or have access to (Deloitte 2013).

Retail data exists in a variety of forms. Traditionally, retailers have been relying on retail scanner data and consumer panel data to conduct analyses. Retail scanner data mainly records point-of-sale data, pricing, promotion, display, feature, product descriptions and so on, whereas consumer panel data tracks information regarding shopping trips and household demographics. Third-party companies (e.g., ACNielsen and IRI) collect retail scanner data from distinct retailers without disclosing key information (i.e., the names of firms and retail stores) and obtain consumer panel data through the use of scanners (Cox 2011). The data not only provides researchers with a great deal of information to analyze the retail sector, but also opens up opportunities to gain a better understanding of the market and competitors for retailers (Cox 2011). Despite the substantial advantages, this data captures the purchase outcomes rather than the search processes of shoppers.

Recent development of tracking technology gives rise to new forms of retail data. For instance, eye tracking captures visual attention via eye fixations (Chandon et al. 2009). Radio Frequency Identification (RFID) technology usually tracks the movement of a product, shopping basket or cart (Seiler and Pinna 2017). Video tracking records all the consumers' in-store activities such as searching at the shelf space, interacting with shopping assistants and talking on the phone (Zhang et al. 2014). However, the above-mentioned technologies may have certain limitations. Specifically, eye tracking is likely to be intrusive, RFID involves potential measurement errors when a shopping cart is left behind by a shopper during the shopping trip and video tracking tends to incur consumer privacy concerns. Therefore, the latest sensing technology allows retailers to observe how consumers physically interact with the shelf space without being intrusive and suffering from related issues. As a shopper steps into a category, the sensors installed on top of the shelf track when and how an item is touched, picked up or returned in the category.

Adopting the state-of-the-art sensing technology to collect a unique dataset from a local grocery retailer, the first essay[1] of this dissertation examines how consumers conduct haptic search to make purchase decisions. This research contributes to the literature by defining key attributes of the shoppers' speed, consideration set, and shopping path at the shelf space (i.e., search speed, search product dispersion, search price dispersion, promotional search propensity, and peripheral search propensity). This study also investigates the effects of consumer haptic search on price paid across food and non-food categories. Gauri and Grewal (2021) emphasize the importance of establishing collaborative relationships between academics and practitioners in retailing. This paper, therefore, provides practitioners with managerial implications regarding in-store category management and shelf layout.

[1] The essay is the joint work with Sandeep R. Chandukala (Associate Professor of Marketing, Lee Kong Chian School of Business, Singapore Management University), Kapil R. Tuli (Lee Kong Chian Professor of Marketing, Lee Kong Chian School of Business, Singapore Management University) and Øyvind Christensen (Founder & Chief Executive Officer, Flow Insights).

Given the rich information contained in the traditional retail data, the second essay supplements Nielsen retailer scanner data and ad intel data with additional publicly available data (e.g., US Bureau of Economic Analysis, Federal Reserve Economic data, and Federal Bureau of Investigation data) to investigate the spillover effects of recreational cannabis legalization (RCL) on related categories (i.e., alcohol, tobacco, candy, and salty snacks). To build causal inference, this essay[2] adopts synthetic control method which relaxes the parallel trends assumption that is a key assumption for difference-in-differences analysis. It is considered as “arguably the most important innovation in the policy evaluation literature in the last 15 years” (Athey and Imbens 2017, p.9) and “a new and relatively cleaner identification strategy” in the marketing literature (Lu et al. 2021, p.3). The findings suggest that RCL resulted in an increase in per capita dollar sales and per capita unit sales of alcohol, salty snacks, and candy, while this was not observed for tobacco sales. To rule out alternative explanations, this work identifies a "null category" (i.e., batteries) and demonstrates that RCL did not lead to changes in pricing or advertising. The results are likely to help policymakers in understanding unintended consequences and potential problems associated with RCL such as excessive drinking and junk food consumption resulting in increasing health care expenses.

To sum up, this dissertation seeks to gain insight into the critical roles of consumers and marketers in a retail context using a variety of unique and rich data sources (e.g., tracking data, retail scanner data, ad intel data and publicly available data). The main aim of the two essays is to focus on unique aspects of

[2] The essay is the joint work with Sandeep R. Chandukala (Associate Professor of Marketing, Lee Kong Chian School of Business, Singapore Management University) and Kapil R. Tuli (Lee Kong Chian Professor of Marketing, Lee Kong Chian School of Business, Singapore Management University).

retail analytics. The two essays will be discussed in detail in the following chapters.

Chapter 2 Capturing Shopper Haptic Search at Shelf Space Utilizing Sensing Technology

2.1 Introduction

A recent report in the Wall Street Journal outlines several efforts by retailers to have a better understanding of shopper's in-store search behavior as it is likely to be a key determinant of their purchase behavior (Gasparro and Kang 2020). Even large traditional brick-and-mortar stores like Walmart are installing cameras and sensors to understand more than just inventory and product quality checks[3]. With the use of computer vision technology there is a renewed interest in stores like Amazon Go in understanding shopper in-store search behavior. Reflecting the importance of in-store shopper search behavior, a large body of literature has studied consumer's shopping paths using a variety of technologies such as RFID tags and eye tracking devices (e.g., Larson et al. 2005, Seiler and Pinna 2017, Seiler and Yao 2017, Hui, Bradlow and Fader 2009). Complementing the focus on shopping paths, studies also examine consumers' behavior at the shelf space (Cobb and Hoyer 1985, Dickson and Sawyer 1990, Hoyer 1984). Indeed, in a recent study, Chen et al. (2021) use eye tracking technology to examine consumers' attention to products placed at different locations on the shelf.

Few studies, however, explore consumers' haptic search behavior at the shelf space. Haptic search is, "active seeking and pickup of information by the hand" (Peck and Childers 2003) and is widely viewed as a critical aspect of consumers' in-store search behavior. Haptic search increases consumers'

[3] https://www.siasearch.io/blog/computer-vision-startups-retail-industry

perceived ownership, enhances a persuasive attitude towards products, and facilitates impulse buying (see Krishna and Schwarz 2014). Indeed, higher number of product touches indicates greater shopper engagement and even higher probability of unplanned product purchases (Hui et al. 2013). Accordingly, this study examines the haptic search behavior of consumers at the shelf space and seeks to make two key contributions.

First, using a unique state-of-the-art non-intrusive sensing technology, we present perhaps the first in-depth examination of shoppers' haptic search behavior at the shelf space by identifying and defining five key attributes of the shoppers' speed, consideration set, and shopping path at the shelf space. Reflecting on prior work on the speed of shopping path in the store (see Zhang et al. 2014), we examine shoppers' speed of haptic search at the shelf space. Drawing on extant research that underscores the importance of shoppers' consideration set (e.g., Mitra and Lynch 1995), we propose that the diversity of products and prices touched by shoppers at the shelf space reflect their consideration set and can provide insights into their purchasing behavior. Finally, consistent with the importance of shopping path (e.g., Larson et al. 2005), we study the shopper's propensity to start their haptic search at the shelf space with promoted products and at the periphery of the aisle, as opposed to the center of the aisle (e.g., Chandon et al. 2000).

To study the five proposed constructs, we collect non-intrusive in-store data over a 2-week period in a Southeast Asian retail store for 3 product categories covering 460 SKUs of 75 brands and measure the haptic search efforts of 1968 shoppers. Results show that the proposed constructs have both statistically and economically significant effects on the average price paid by the shopper, a key consideration for the retailers (Glandon 2018). In sharp contrast to the received view that encourages retailers to speed up the in-store shopping process (e.g., Van den Bergh et al. 2016), we find that higher search speed at the shelf space lowers the average price paid by .4 cents per unit. Bringing to fore the nuances of shopper consideration sets, we find opposing effects of product and price search dispersion. Whereas one standard deviation increase in product search dispersion lowers the price paid by 1.4 cents per unit, one standard deviation increase in price search dispersion increases the price paid by .9 cents per unit. Underscoring the importance of considering the shopper path at the shelf space, we also find that one standard deviation increase in the promotion and peripheral search propensity lowers the price paid by the shopper by 2 cents and 1 cent per unit, respectively.

Second, we contribute to the literature by investigating how the effects of haptic search at the shelf space on price paid differ *across* food and non-food categories. While higher search speed decreases price paid in both categories, the effect is stronger in the food category. Moreover, the magnitude of the negative effect of search product dispersion on price paid is larger in the non-food category. In addition, we only observe significant impacts of search price dispersion, peripheral search propensity and promotional search propensity on price paid in the non-food category.

The remainder of this article is organized as follows: First, we provide background on haptic search and the associated literature on offline search. Next, we introduce each of the five dimensions (i.e., search speed, search product dispersion, search price dispersion, promotional search propensity, and peripheral search propensity) of haptic search and motivate the moderating role of category. We follow it up with our data collection strategy using sensing technology. We then provide specific examples of our variable operationalization, discuss the proposed empirical methodology, our strategy to address endogeneity, and present our results. Finally, we discuss the theoretical and managerial implications of our study and conclude with limitations and potential avenues for future research.

2.2 Haptic Search in Offline Stores

2.2.1 In-store Search Behavior

Given the importance of in-store search behavior of consumers for managers, a growing body of work leverages a variety of technologies to examine multiple dimensions of shopper in-store search behavior (see Roggeveen et al. 2020). One stream of this research underscores the importance of the specific path followed by the shoppers and the time they spend during a store visit (Larson et al. 2005). For example, Hui, Bradlow and Fader (2009) find that shoppers are likely to be more purposeful and less exploratory as they spend more time in the store (also see Reutskaja et al. 2011, Seiler and Yao 2017, Zhang et al. 2014). More recently, Seiler and Pinna (2017) find that shoppers who spend more time in the store, are more likely to purchase lower priced items.

Another stream of studies leverages eye tracking and video capture equipment to examine shopper's attention to different product categories (e.g., Hui, Bradlow and Fader 2009, Hui, Fader, and Bradlow 2009). More recently, Chen et al. (2021) leverage ambulatory eye tracking equipment to examine shopper's lateral and vertical biases related to the physical location of SKUs on the shelf space. Importantly, Chen et al. (2021) underscore the need for research to examine shoppers' search behavior at the point of purchase, i.e., the shelf space, as such efforts will provide retailers more insights into the specific purchases made by the shoppers.

2.2.2 Haptic Search Behavior

A key aspect of shoppers' in-store search behavior is haptic search, that is, their tendency to touch the products (see Peck and Wiggins 2006, Peck and Shu 2009). Indeed, Hui et al. (2013) consider the number of product touches by a shopper as an engagement indicator with the product and find that the number of product touches is significantly higher for purchases of unplanned products. Taken together, prior research highlights the importance of considering shopper behavior at the shelf space and haptic search in general. Synthesizing these two ideas, we propose that shoppers' haptic search behavior at the shelf space is likely to provide insights into their purchase behavior. We draw on prior work that emphasizes the importance of shoppers' speed (e.g., Zhang et al. 2014), consideration set (e.g., Van Nierop et al. 2010), and the specific path followed by the shopper (e.g., Larson et al. 2005) to study five haptic search attributes at the shelf space.

First, building on literature on walking speed of shoppers in a store (Van Den Bergh et al. 2016), we focus on the speed of touching products in a category to capture consumer haptic search effort at the shelf space. Given that speed is likely to reduce consumers' recall and recognition of products on a shelf, it is a critical factor that indicates the intensity of information processing or the urgency of their needs (Zhang et al. 2014). *Second*, we propose that haptic search efforts of shoppers reflect their consideration set and therefore offer insight into their purchases. We study both the products and the price points during the shoppers' haptic search at the shelf space. We examine search product dispersion, i.e., the degree to which a shopper explores multiple types of products. The variety of products included in the shopper's haptic search at the shelf space directly reflects the consideration set size, a key determinant of the shopper's price elasticity (Mitra and Lynch 1995). Similarly, shoppers' search price dispersion reflects the variety of prices considered by a shopper, a key factor in their decision-making process as they try to gain more price information by investing time and effort on price search (e.g., Mazumdar and Jun 1993, Mehta et al. 2003).

Third, complementing studies on shopping path across the retail store, we examine the effects of shopping path at the shelf space. Reflecting importance of the shelf-center for consumers (e.g., Chandon et al. 2009), we study the degree to which a consumer searches from shelf edges rather than the shelf center in a category, that is, peripheral search propensity. In addition, we explore the degree to which a consumer is likely to prioritize exploring a promoted product, that is, promotional search propensity. Despite the literature on the effectiveness of sales promotions (e.g., Chandon et al. 2000), the effect of searching a promoted product first in a specific category during a shopping trip is unknown.

2.3 Dimensions of Haptic Search

We now define the five dimensions of haptic search i.e., search speed, search product dispersion, search price dispersion, peripheral search propensity, and promotional search propensity. For each of the dimensions of haptic search we also provide cursory justification for their impact on price paid.

2.3.1 Search Speed

Search speed refers to the number of distinct products considered per minute by the consumer in a specific category during a shopping trip. Search speed allows us to capture two important dimensions of consumer product search effort, i.e., the number of products considered by the consumer and the time spent in considering these products (Malhotra 1982, Larson et al. 2005). If we only consider the time spent in a store by a consumer, we do not know whether the time was spent examining products or only for browsing activity. Similarly, if we only consider the number of products considered by a consumer, we do not know how long it took the consumer to do so. We believe that a consumer that examines five products in a minute is expending more intense product search effort as compared to a consumer that considers only one product in a minute.

As consumers search more unique products in a given amount of time, they tend to spend less time on each product considered during a shopping trip. This indicates that they are less active in seeking out information about products, demonstrating a low level of involvement for each product considered (Zaichkowsky 1985). As low product involvement implies that consumers are less concerned with product benefits, a decrease in product involvement is likely to lower price acceptability level and increase price consciousness of the consumers (Lichtenstein et al. 1988). Given that price conscious consumers are unwilling to pay for additional product benefits, they tend to pay less for each product considered, resulting in lower average price paid.

2.3.2 Search Product Dispersion

Search product dispersion refers to the degree to which a consumer explores multiple types of products in a specific category during a shopping trip. Search product dispersion, therefore, reflects the diversity of the consumers consideration set. The more diverse the consumers' consideration set, the higher the probability that the consumer is not wedded to a specific choice and therefore the higher the price elasticity of the shopper (see for example, Mitra and Lynch 1995). In addition, when exploring extensive options, individuals are likely to feel more enjoyable due to more possibilities available but experience more frustrations associated with the difficulty to process substantial amount of product information (Iyengar and Lepper 2000).

As a result, consumers are likely to rely more on an affective system and less on an analytical system of decision-making (Strack and Deutsch 2004). Therefore, consumers are less likely to assess product attributes such as price to infer product quality (Shafir et al. 1993). Indeed, the fact that shoppers' haptic search effort encompasses more diverse set of products can indicate their perception of greater choice, and therefore higher quality of offerings (see e.g., Mogilner et al. 2008). They are less likely to hold price-quality belief that higher prices indicate better quality and tend to become more price sensitive, leading to lower price paid (Cronley et al. 2005).

2.3.3 Search Price Dispersion

Search price dispersion refers to the degree of variance of the price points explored by a consumer in a category during a shopping trip. As such, search price dispersion captures the variability of all the price points in a category searched by shoppers and reflects the price search effort during a shopping trip. Consumers tend to utilize the price information about competing products (i.e., external reference prices) to evaluate the price of focal product at the point-of-purchase (Mayhew and Winer 1992, Kumar et al. 1998). While consumers who have the same external reference prices may search within the same price range, they are likely to differ in the amount of variation around their price standards. For instance, two shoppers who search across the same price range (e.g., $1, $2, $3 vs. $1, $3) are centered around the same average price point (both have mean of $2) but may have different search price dispersions (standard deviations are 1 vs. 1.41).

The average of the prices searched by a consumer during a shopping trip serves as a reference point in terms of opportunity costs and determines the level of pain of paying and hence indicates her budget (Soster et al. 2014). The reference point represents prospective losses when the budget shrinks and signals prospective gains in the case of budget expansion (Carlson et al. 2015). Higher search price dispersion, therefore, shows that the consumer has a broader consideration set and is more likely to deviate from the external reference prices formed during a shopping trip i.e., her budget (Jacobson and Obermiller 1990), and therefore are less likely to stick to her budget constraints. Indeed, high price dispersion means shoppers have a high price latitude and are more likely to accept higher prices (see Sorce and Widrick 1991). Consumers who are less likely to follow a budget tend to experience less pain of payment, resulting in higher price paid.

2.3.4 Peripheral Search Propensity

Peripheral search propensity refers to the degree to which a consumer searches from shelf edges rather than the shelf center in a specific category during a shopping trip. The more likely the first product a consumer searches from the edges of a shelf as opposed to the center of the shelf, the higher the peripheral search propensity. Consumers believe that products displayed on the center of a shelf are likely to be of not only higher quality (Raghubir and Valenzuela 2006), but also more popular (Valenzuela and Raghubir 2009). Indeed, studies of retail shelf space do find that brands located in the center of a shelf gain more attention and therefore are more likely to be purchased (Chandon et al. 2007, 2009).

High peripheral search propensity of a consumer, therefore, suggests that she is less likely to rely on shelf locations to make inferences about the popularity and quality of the product. As such, consumers with high peripheral search propensity are willing to put in more effort in inferring product quality and are less likely to rely on indicators such as price and shelf locations that are less effortful (Jacobson and Obermiller 1990). This suggests that consumers with higher peripheral search propensity are also less likely to hold the belief that higher price signals better quality. As consumers with a weak price-quality belief are less willing to pay more for a product in return for good quality, they tend to be more price sensitive and pay lower prices (Cronley et al. 2005).

2.3.5 Promotional Search Propensity

Promotional search propensity refers to the degree to which a consumer is likely to prioritize exploring a promoted product in a specific category during a shopping trip. The more likely that the first product touched by a consumer during a shopping trip is a promoted product, the higher the promotional search propensity. Therefore, promotional search propensity reflects the extent to which consumers are sensitive to price promotions.

The higher the promotional search propensity, the more likely that a consumer is active in locating in-store promotions and is sensitive to in-store deals (Schneider and Currim 1991). Hence, the consumer is likely to lower her external reference price formed by the first product she searches and hence pays less (Mayhew and Winer 1992). As a result, higher promotional search propensity is likely to result in lower average price paid.

2.3.6 Moderating Role of Category

Given that food consumers are concerned with their health and safety, eager to know more about the product information such as the ingredients and willingness to pay higher prices, consumers, in general, tend to focus more on product quality in food category compared to non-food category (PwC 2015, Olayanju 2019). Therefore, we examine the moderation effect of category (i.e., food vs. non-food) on the relationship between haptic search and price paid. Purchase frequency is typically higher in food category due to higher perishability of food items. As food shoppers visit the category more frequently, they are more likely to be exposed to price information, resulting in price information rehearsal and storage in long-term memory (Bettman 1979, Winer 1986). In addition, high purchase frequency is associated with the need for allocating consumer budgets constantly and repetitively (Monroe 2003, Nagle and Holden 1987). As such, high purchase frequency is positively associated with price knowledge (Estelami and Maeyer 2004). Consumer shopping process in the food category, therefore, is accompanied by higher prior price knowledge. Higher purchase frequency also means that consumer long-term knowledge about the products and product options are likely to be higher in the food category (see Kalyanaram and Little 1994).

Higher long-term memory of product knowledge and price points implies that even if search speed is high, consumer decision making is more directed and therefore is less likely to be an indicator of lower consumer involvement in the purchasing process. As such, the negative effect of search speed on price paid is weaker in food category. Similarly, higher search product dispersion is less likely to result in more affective decision making in food category. In addition, the propensity of consumers to deviate from their budget constraints is likely to be lower even when they explore different price points during their purchase process in the food category. This is because while shopping in the food category, consumers can draw on their long-term memory of product knowledge and price points even when they are exploring a variety of different products (see Jacobson and Obermiller 1990). As such, the negative effects of search product and price dispersion on price paid by the consumers is likely to be weaker for purchases in the food category.

Greater prior knowledge of product and prices in the food category also means that extrinsic cues such as product location on the shelf space are likely to be less salient for consumers. This is because high knowledge of product and price means that consumers are more capable of evaluating product attributes and rely on intrinsic cues such as, brand name, country of origin, and packaging (Rao and Sieben 1992). As such, the negative effect of peripheral search propensity on price paid is likely to be lower for purchases in the food category. In addition, given that consumers have richer prior knowledge of product and prices about food items, they are more likely to form internal reference prices based on memory and rely less on external reference prices based on stimulus during the search process (see Mazumdar et al. 2005). Hence, higher promotional search propensity (i.e., the more likely that a consumer searches for a promoted item first during a shopping trip) will result in a smaller decrease in external reference prices, leading to a smaller decrease in price paid in the food category. As such, the negative effect of promotional search propensity on price paid is likely to be weaker in the food category.

2.4 Sample and Data Collection

2.4.1 Tracking Technology

The unique tracking technology adopted in the current study collects consumer in-store search information in front of the shelves. As a shopper enters a category, the sensors installed on top of the shelf record consumer sequential physical engagement e.g., timestamp of all activities including when a shopper enters and leaves a category as well as when and which products are touched, picked up and returned in the category (see Figure 1). Table 1 summarizes the various tracking methodologies commonly used in extant research and compares our sensing technology over other existing tracking approaches. For instance, eye-tracking technology enables researchers to capture visual attention via eye fixations (Chandon et al. 2009) while being intrusive. The current sensing technology captures accurate information by tracking the movement of a shopper and shopper’s hands rather than a basket or cart, as is usually done using RFID technology, and thus addresses potential measurement errors, especially when a cart is left behind by a shopper during the shopping trip (Seiler and Pinna 2017). Additionally, our data collection approach does not suffer from consumer privacy concerns as it does not store any private information, which could be a potential issue with video tracking. Finally, our sensing technology builds on the anonymized and fragmented event-based (AFE) tracking technology (Kakatkar and Spann 2019) by capturing all the shopper activities in a category.

Table 1:
Comparisons of Tracking Technologies Used in the Literature

Characteristics	Eye Tracking	RFID	Video Tracking	AFE Tracking	**Sensing**
Literature	Chandon et al. (2009)	Seiler and Pinna (2017)	Zhang et al. (2014)	Kakatkar and Spann (2019)	**Current study**
Non-Intrusiveness	N	Y	Y	Y	**Y**
No-Privacy-Concern	N	Y	N	Y	**Y**
Information Accuracy	Y	N	Y	Y	**Y**
Continuous Shopping	Y	Y	Y	N	**Y**

Figure 1:
Illustration of the Sensing Technology Adopted in the Current Work

2.4.2 Data Collection

We collected data from a brick-and-mortar retail store that belongs to a large supermarket chain located in Southeast Asia. Apart from the consumer-level data (obtained using sensing technology), we also obtained product (e.g., price, promotion, and shelf location) and brand information (e.g., brand name and type) from the store data. We obtained data on three product categories (i.e., cereal, household cleaning, and pasta) because they represent food and non-food categories in the store. As can be seen in Table 2, the cereal category has 201 unique SKUs and 31 brands, the household cleaning category has 163 SKUs and 30 brands, and the pasta category has 96 SKUs and 14 brands. Having food (cereal and pasta) and non-food (household cleaning) categories helps us in investigating differences arising due to varying shopper involvement. Figure 2 (a), (b) and (c) provide the shelf layouts for the cereal, household cleaning and pasta categories, respectively.

Table 2:
Summary of the Cereal, Household Cleaning, and Pasta categories

Characteristics	Cereal	Household Cleaning	Pasta
Number of unique SKUs	201.000	163.000	96.000
Number of unique brands	31.000	30.000	14.000
Average unit price	.022	.065	.012
Number of shoppers	1,116.000	349.000	503.000
Average shopper age	32.195	36.447	32.207
Proportion of basket/cart users	.453	.493	.553
Proportion of group shoppers	.341	.226	.256

Notes: Average unit price refers to the average unit price of all the items in a category. Basket/cart users refer to consumers shopping with baskets or carts. Group shoppers are individuals shopping with others.

Figure 2:
The Layout of Each Category

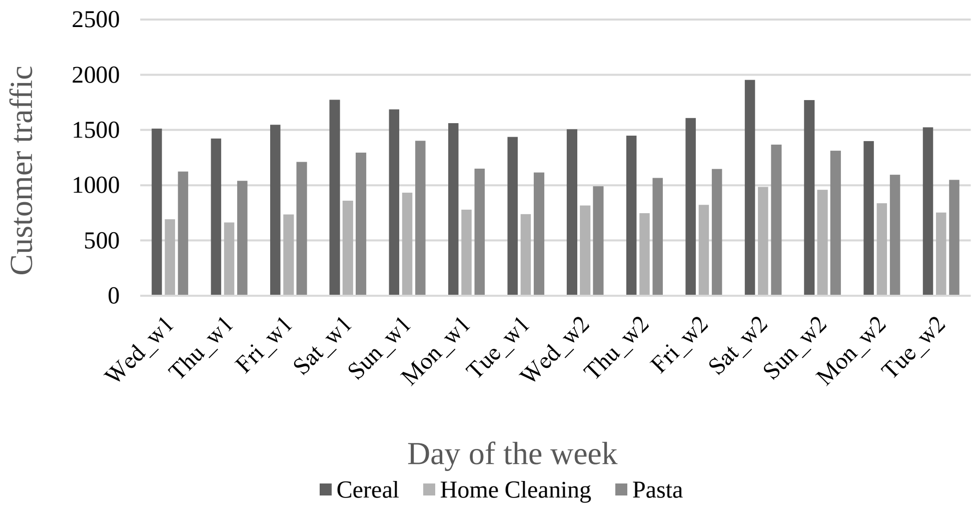
(a) Cereal Category

(b) Household Cleaning Category

(c) Pasta Category

Our data is collected for two weeks from September 19th 2018 to October 2nd 2018. During the two weeks, 1,968 shoppers purchased at least one item across the three categories[4]. Figure 3 shows that, as expected, across all three categories we observe higher customer traffic during the weekend.

Figure 3:
Customer Traffic per Day in Cereal, Home Cleaning, and Pasta (Sept. 19 – Oct. 2, 2018)



Notes: Customer traffic refers to the number of consumers who visited each of the three categories in the store on a specific day. Wed_w1, Thu_w1, Fri_w1, Sat_1, Sun_w1, Mon_w1 and Tue_w1 refer to Wednesday, Thursday, Friday, Saturday, Sunday, Monday and Tuesday of the first week, respectively (i.e., Sept. 19 – Sept. 25, 2018). Wed_w2, Thu_w2, Fri_w2, Sat_2, Sun_w2, Mon_w2 and Tue_w2 refer to Wednesday, Thursday, Friday, Saturday, Sunday, Monday and Tuesday of the second week, respectively (i.e., Sept. 26 – Oct 2, 2018).

[4] While we do not have transaction data that can be linked to individual shoppers, we assume that a purchase was made when a consumer picked up a product in a category but did not put it back on the shelf before exiting the category. The high correlations between estimated sales and actual sales in each category (i.e., around 0.8) indicates that our assumption is reliable.

2.4.3 Data Description

We find that 1,116 shoppers in the cereal category, 349 shoppers in the household cleaning category and 503 shoppers in the pasta category purchased at least one product (Table 2). Although the sensing technology did not record consumer profiles, we still gained some non-private information about the shoppers such as approximate age, shopping with or without a basket/cart, and shopping alone versus in a group. We observe that 45.3%, 49.3%, and 55.3% of shoppers used baskets or carts across the three categories. Additionally, 34.1%, 22.6%, and 25.6% of customers shopped with other people in the three categories.

2.5 Variables and Measures

2.5.1 Independent Variables

Search speed. It is measured as the ratio of the number of unique SKUs to the amount of time spent searching (deliberation time) by the shopper. We illustrate this measure in Table 3(A) using a stylized example of three shoppers in a given category. The second column in Table 3(A) provides the sequence of products (SKUs) searched by each shopper with all shoppers starting their search with product ‘a’ followed by product ‘b’. Shopper 1’s search concludes with choosing product ‘b’ after considering product ‘b’ for the second time. However, shopper 2’s search concludes with choosing product ‘b’ (after considering it only once), while shopper 3’s search concludes with choosing product ‘c’ after considering product ‘b’. The amount of time each shopper spends, referred to as deliberation time, searching for the products is provided in column 3 of Table 3(A). Deliberation time is the total duration of time after a shopper physically touches the first product in a category until exiting the category. Deliberation time, therefore, includes the total time expended by a shopper in considering the various products in a category during the physical search process.

Table 3 (A):
Examples of Search Speed, Product Dispersion, Price Dispersion, and Price Paid

Shopper	Search Sequence (Unit Price)	Deliberation (Minute)	Search Speed	Search Product Dispersion	Search Price Dispersion	Average Price Paid
1	*a* ($1) –> *b* ($2) –> *b* ($2)	12.000	.167	.637	.577	1.667
2	*a* ($1) –> *b* ($2)	10.000	.200	.693	.707	1.500
3	*a* ($1) –> *b* ($2) –> *c* ($5)	12.000	.250	1.099	2.082	2.667

The fourth column in Table 3(A) provides search speed as the ratio of number of unique SKUs searched to the deliberation time. Based on Table 3(A) we find that shopper 1 searches product ‘a’ once, followed by searching product ‘b’ twice in 12 minutes. The search speed of shopper 1, therefore will be,

2 / 12 = 0.167 products/minute

Shopper 2 searches product ‘a’ first and then product ‘b’ only once. As such, shopper 2 spends 10 minutes deliberating about two unique SKUs and hence the search speed will be,

2 / 10 = 0.2 products/minute

Finally, shopper 3 searches three unique products ‘a’, ‘b’, and ‘c’ once each in 12 minutes, hence, the number of unique SKUs searched per minute is illustrated below,

3 / 12 = 0.25 products/minute

Thus, our measure accounts for the number of unique products searched and the time spent on searching these products and therefore, provides an accurate measure of search speed.

Search product dispersion. It is measured as the Shannon Entropy (Shannon and Weaver 1949, entropy in short hereafter). The formula for entropy is given as

$$H = -\sum_i p_i \, ln(p_i), \tag{1}$$

where p_i is the probability of a shopper searching for product i (i.e., number of times product ‘i’ was searched by a given shopper over total of all products searched by the same shopper). Intuitively, the higher the entropy, the higher the uncertainty of consumer preferences.

As shown in Table 3(A), shopper 1 searches two unique products (i.e., product ‘a’ and ‘b’) in a category and purchases the last product searched i.e., product ‘b’. Shopper 1 searches ‘a’ first, followed by searching ‘b’ twice. The probabilities of searching ‘a’ and ‘b’ are 1/3 and 2/3, respectively. As such, the entropy of the first shopper is computed as,

$$- [(1/3) \times \ln(1/3) + (2/3) \times \ln(2/3)] = 0.637$$

Similarly, the entropy of the second shopper is,

$$- [(1/2) \times \ln(1/2) + (1/2) \times \ln(1/2)] = 0.693$$

The entropy of the third shopper who searches ‘a’, ‘b’, and ‘c’ consecutively and each only once is computed as,

$$- [(1/3) \times \ln(1/3) + (1/3) \times \ln(1/3) + (1/3) \times \ln(1/3)] = 1.099$$

Thus, our measure of search product dispersion accounts for the number of times a given product is included in a shopper’s consideration during the entire search process.

Search price dispersion. It is measured as the standard deviation of all the prices searched by a shopper in a category. Table 3(A) shows that shopper 1 searches product ‘a’ priced at $1 first and then searches product ‘b’ priced at $2 twice. The standard deviation of all the prices searched is 0.577. Shopper 2 searches ‘a’ and ‘b’ once, respectively. Therefore, the standard deviation of prices searched is 0.707. Shopper 3 searches a third product (i.e., product ‘c’) priced at $5, and the standard deviation is 2.082. Our measure of search price dispersion, therefore, captures the extent of variation of prices searched by a shopper during their entire search process in the category.

Peripheral Search Propensity. It is measured as the Euclidean distance between the location (coordinates) of the first product searched by a shopper and the category shelf center. Table 3(B) shows that shopper 1 starts searching product ‘a’ and then searches ‘b’ twice. The first product searched is product ‘a’ and its shelf location has the coordinates of 100 cm and 150 cm. 100 cm is the horizontal distance from the left end of the category to the location of the product, while 150 cm is the vertical distance from the bottom of the shelf to the product. Similarly, the category shelf center has the coordinates of 300 cm and 100 cm along the horizontal and vertical dimensions, respectively. Thus, the distance between product ‘a’ and shelf center is computed as,

$$\sqrt{[(100 - 300)^2 + (150 - 100)^2]} = 206.155 \text{ cm}$$

Similarly, shopper 2 searches product ‘c’ first, which has coordinates of 400 cm and 100 cm. The coordinates of the center of the category are 300 cm and 100 cm as described before. Therefore, the distance to shelf center from product c for shopper 2 is,

$$\sqrt{[(400 - 300)^2 + (100 - 100)^2]} = 100 \text{ cm}$$

Therefore, our measure of peripheral search propensity captures the relative location of the shopper’s starting search point from the central location of the shelf. The larger the peripheral search propensity value the farther away the first product searched is from the shelf center.

Table 3 (B):
Examples of Peripheral Search Propensity

Shopper	Search Sequence	First Product Searched and its Location (cm)	Location of Shelf Center (cm)	Peripheral Search Propensity (cm)
1	*a* –> *b* –> *b*	*a* (100, 150)	(300,100)	206.155
2	*c* –> *c* –> *d*	*c* (400, 100)	(300,100)	100

Notes: The length in the shelf location is the distance to the left side of the category shelf. The height in the shelf location is the distance to the bottom of the category shelf.

Promotional search propensity. It is measured as one minus the ratio of the number of times a shopper searches any product prior to searching the first promoted product to the total search frequency. For instance, as indicated in Table 3(C), shopper 1 searches product ‘a’ first, which is on promotion. Promotional search propensity, therefore, is calculated as, one minus the ratio of the number of products searched before the first promoted product to the total number of products searched, for shopper 1 is:

$$1 - 0 / (1+2) = 1$$

Shopper 2 searches promoted product ‘a’ after searching non-promoted product ‘b’ once. As such, promotional search propensity is computed as,

$$1 - 1/ (1+2+1) = 0.67$$

On the contrary, shopper 5 does not search any promoted item and hence the promotional search propensity measure will be

$$1 - 3/ (1+1+1) = 0$$

To sum up, promotional search propensity is a continuous variable ranging from 0 to 1. The larger the value of this measure, the earlier in the search sequence a promoted product is searched. Thus, this measure captures the initial presence of a promotional product in a shopper’s search sequence.

Table 3 (C):
Examples of Promotional Search Propensity

Shopper	Search Sequence	First Promoted Product Searched	Promotional Search Propensity
1	*a* –> *b* –> *c*	*a*	1 – 0/3 = 1
2	*b* –> *a* –> *c*	*a*	1 – 1/3 = 0.67
3	*b* –> c –> *a*	*a*	1 – 2/3 = 0.33
4	*b* –> *a* –> *a*	*a*	1 – 1/3 = 0.67
5	*b* –> c –> *b*	*None*	1 – 3/3 = 0

2.5.2 Dependent Variable and Control Variables

Average price paid. We use the average unit price paid (Seiler and Pinna 2017) by a consumer as our dependent variable. It is measured as the ratio of basket size (price per unit) to purchase quantity. If a consumer, at the end of the search, picks up only one product and leaves the category then the average price paid is the price of this product. However, if the consumer picks up multiple products we then compute the average price paid as the ratio of price per unit to purchase quantity. For expositional purposes we stick to the same examples shown in Table 3(A) and assume that consumers purchase all the items searched. Shopper 1 purchases one product ‘a’ and two products ‘b’. The average price paid is calculated as,

$$(1*1 + 2*2) / (1+2) = \$1.667$$

Shopper 2 purchases one product ‘a’ and one product ‘b’. The average price paid will be,

$$(1*1 + 1*2) / (1+1) = \$1.500$$

Shopper 3 purchases one product ‘a’, one product ‘b’ and one product ‘c’. The average price paid by the shopper is computed as,

$$(1*1 + 1*2 + 1*5) / (1+1+1) = \$2.667$$

Control Variables. We include several control variables to account for shopper search behavior, shopper characteristics, category, time, and day effects. Specifically, we control for group shopping as a dummy variable that equals 1 when a shoppers is with others and 0 when the shopper shops alone. Shopping with a basket/cart is also a dummy variable (1 = shopping with a basket/cart, 0 = shopping without a basket or a cart). Lastly, we also control for the age of the shoppers, time of the day, and day of the week effects. We use food category dummy as a moderator in our study as discussed previously. All the continuous variables have been winsorized at the 99% level. Table 4 outlines the summary statistics and correlations among these variables. As can be seen from Table 4 we do not observe any high correlations.

Table 4:
Summary Descriptive and Correlations between Key Variables

No.	Variables	Mean	SD	Min	Max	1	2	3	4	5	6	7	8	9	10
1	Average price paid	.020	.026	.003	.178	1.000									
2	Search speed	5.013	5.204	.632	30.000	-.025	1.000								
3	Search product dispersion	.582	.639	.000	2.322	-.016	***-.269***	1.000							
4	Search price dispersion	.004	.012	.000	.082	***.343***	***-.126***	***.384***	1.000						
5	Peripheral search propensity	187.261	116.600	19.329	442.097	***.080***	-.014	-.010	***.049***	1.000					
6	Promotional search propensity	.288	.423	.000	1.000	***-.109***	***-.085***	***.191***	.027	-.009	1.000				
7	Age	32.952	10.531	20.000	70.000	-.010	-.027	.021	.024	-.034	.017	1.000			
8	Group shopping	.299	.458	.000	1.000	.011	***-.159***	***.194***	***.092***	***.047***	***.053***	***.051***	1.00		
9	Shopping with a basket/cart	.485	.500	.000	1.000	-.027	***-.064***	.018	.032	***-.092***	.002	***.081***	***.048***	1.000	
10	Food category	.823	.382	.000	1.000	***-.313***	***-.044***	***.065***	***-.234***	***.116***	***.054***	***-.154***	***.074***	-.007	1.000

Notes: All the continuous independent variables are winsorized at 99% significance level. The outcome variable is also winsorized at 99% significance level (Kogan et al. 2017). All correlations in bold and italicized are significant at 90% level of significance. Average price paid refers to average unit price paid per customer.

2.6 Model Specification and Results

2.6.1 Model Specification

This research aims to investigate the impact of search speed, search product dispersion, search price dispersion, peripheral search propensity, and promotional search propensity on the average price paid. Our proposed model is given by,

$$\begin{aligned}\text{Price\_Paid}_i = {} & \theta_0 + \theta_1\times\text{Search\_Speed}_i + \theta_2\times\text{Product\_Dispersion}_i \\ & + \theta_3\times\text{Price\_Dispersion}_i + \theta_4\times\text{Peripheral\_Search}_i \\ & + \theta_5\times\text{Promotional\_Search}_i + \theta_6\times\text{Age}_i + \theta_7\times\text{Group}_i \\ & + \theta_8\times\text{Basket\_Cart}_i + \theta_9\times\text{Food\_Category}_i \\ & + \textstyle\sum_{k=10}^{15} \theta_k\times\text{Day\_Week}_{ik} \\ & + \textstyle\sum_{m=16}^{18} \theta_m\times\text{Time\_Day}_{im} + \omega_i \qquad (2)\end{aligned}$$

Moreover, we seek to understand if the effects differ for food and non-food categories. We further estimate the moderation effects using the model shown below,

$$\begin{aligned}\text{Price\_Paid}_i = {} & \beta_0 + \beta_1\times\text{Search\_Speed}_i + \beta_2\times\text{Product\_Dispersion}_i \\ & + \beta_3\times\text{Price\_Dispersion}_i + \beta_4\times\text{Peripheral\_Search}_i \\ & + \beta_5\times\text{Promotional\_Search}_i + \beta_6\times\text{Food\_Category}_i \\ & + \beta_7\times\text{Food\_Search\_Speed}_i + \beta_8\times\text{Food\_Product\_Dispersion}_i \\ & + \beta_9\times\text{Food\_Price\_Dispersion}_i + \beta_{10}\times\text{Food\_Peripheral\_Search}_i \\ & + \beta_{11}\times\text{Food\_Promotional\_Search}_i + \beta_{12}\times\text{Age}_i + \beta_{13}\times\text{Group}_i \\ & + \beta_{14}\times\text{Basket\_Cart}_i \\ & + \textstyle\sum_{k=15}^{20} \beta_k\times\text{Day\_Week}_{ik} \\ & + \textstyle\sum_{m=21}^{23} \beta_m\times\text{Time\_Day}_{im} + \varphi_i \qquad (3)\end{aligned}$$

Where, $Price\_Paid_i$ is the average price paid per trip by shopper *i*. $Search\_Speed_i$, $Product\_Dispersion_i$, $Price\_Dispersion_i$, $Peripheral\_Search_i$, and $Promotional\_Search_i$ are the focal independent variables that capture the search activity for shopper *i*. $Food\_category_i$ is a dummy variable that takes a value of 1 if shopper *i* makes a purchase in the food category else a value of 0. In addition, $Food\_Search\_Speed_i$, $Food\_Product\_Dispersion_i$, $Food\_Price\_Dispersion_i$, $Food\_Peripheral\_Search_i$, and $Food\_Promotional\_Search_i$ are the interaction terms of Food_Category and the five focal constructs. As noted, we control for consumer characteristics such as age (Age_i), group shopping ($Group_i$), and shopping with a basket/cart ($Basket\_Cart_i$). We also take into account day of the week and time of the day.

2.6.2 Addressing Potential Endogeneity

Assessing the impact of consumer search behaviour through observational data requires addressing potential endogeneity concerns. Specifically, consumers could be strategic in their search behaviour, i.e., some consumers might be faster in their search because they are younger or familiar with the category layout or have time constraints due to other commitments. While some variables are observable, and hence can be controlled for, there could be unobservable product characteristics (e.g., quality or durability) that influence the focal search variables (search price dispersion and promotional search propensity) and also the consumer decision of price paid. Similarly, retailers are strategic in their behaviour and some of the characteristics of the shelf layout like packaging and placement of certain brands are not random and could impact consumer search variables like search speed, search product dispersion, and peripheral search propensity and can potentially be correlated with the error term. Thus, all five focal search variables in our study are potentially endogenous.

In addition, potentially correlated omitted variables and measurement error can be sources of endogeneity. For example, consumer in-store search behaviour could be influenced by omitted variables like advertising activities which are unobserved in our case. Furthermore, consumer search behaviour could be different when they shop alone as opposed to shopping in a group and this could in turn impact the price paid (a consumer might pay less when shopping alone and hence could search more). It is therefore imperative to account for these diverse sources of endogeneity.

We follow a twofold approach to address the endogeneity concerns raised above. First, we control for relevant covariates. Specifically, we account for customer and shopping characteristics by incorporating consumer age, usage of cart/basket, and whether the consumer is shopping alone or in a group. Additionally, differences in consumer search behaviour during different days of the week and different times of the day can be accounted for using variables for days of the week and time of the day. These variables address only some of the issues discussed above related to strategic consumer behaviour and omitted variables.

Second, to account for other sources of endogeneity, we use an instrument-free approach that involves Gaussian copulas. Gaussian copulas allow us to model the joint distribution of the endogenous regressors and the structural error term and make inferences on model parameters by maximizing the resulting likelihood function (Park and Gupta 2012). Regardless of the underlying reason (omitted variables, strategic firm or consumer behavior), the main issue of endogeneity arises due to the correlation between the independent (endogenous) variable and the error term. Gaussian copula addresses the correlation between the normal error term and the non-normal endogenous variable by introducing an additional variable (copula) that directly accounts for this correlation (Park and Gupta 2012). Papies et al. (2017) demonstrate, through simulation studies, that conditional on non-normal endogenous regressor and normal structural error, Gaussian copula method addresses the endogeneity bias and is as efficient as an instrumental variable approach.

The implementation of the copula approach is similar to the control function approach and, thus, has been widely adopted in prior literature to correct for endogeneity (e.g., Bombaij and Dekimpe 2020, Atefi et al. 2018). Specifically, copula approach has been used to account for endogeneity due to retailer's strategic decisions (Bombaij and Dekimpe 2020), consumer's strategic decisions (Schweidel and Knox 2013), omitted variables (Lim et al. 2018), unobserved variables (Atefi et al. 2018, Burmester et al. 2015), and endogeneity of marketing mix variables (Datta et al. 2017, Datta et al. 2015). An advantage of using the copula approach, especially when incorporating moderators in our analysis, is that it requires only one correction term to address both the endogenous regressor and its interaction term (Papies et al. 2017).

A critical prerequisite of this approach is that the endogenous regressor should not follow either Bernoulli or normal distribution. Given that our endogenous regressors are continuous variables, they do not follow a Bernoulli distribution. Additionally, we conduct the Shapiro-Wilk test and find that search speed ($W = .686$, p-value $< .000$), search product dispersion ($W = .832$, p-value $< .000$), search price dispersion ($W = .396$, p-value $< .000$), peripheral search propensity (W = .933, p-value < .000) and promotional search propensity (W = .640, p-value < .000) are not normally distributed. As such, we can use copulas in our empirical setting which involves including additional regressors that control for each of the specific endogenous variables. Therefore, our final proposed model is:

$$\begin{aligned}
\text{Price_Paid}_i = {} & \alpha_0 + \alpha_1 \times \text{Search_Speed}_i + \alpha_2 \times \text{Product_Dispersion}_i \\
& + \alpha_3 \times \text{Price_Dispersion}_i + \alpha_4 \times \text{Peripheral_Search}_i \\
& + \alpha_5 \times \text{Promotional_Search}_i + \alpha_6 \times \text{Food_Category}_i \\
& + \alpha_7 \times \text{Food_Search_Speed}_i + \alpha_8 \times \text{Food_Product_Dispersion}_i \\
& + \alpha_9 \times \text{Food_Price_Dispersion}_i + \alpha_{10} \times \text{Food_Peripheral_Search}_i \\
& + \alpha_{11} \times \text{Food_Promotional_Search}_i + \alpha_{12} \times \text{Age}_i + \alpha_{13} \times \text{Group}_i \\
& + \alpha_{14} \times \text{Basket_Cart}_i + \alpha_{15} \times \textit{Search_Speed_Copula}_i \\
& + \alpha_{16} \times \textit{Product_Dispersion_Copula}_i \\
& + \alpha_{17} \times \textit{Price_Dispersion_Copula}_i \\
& + \alpha_{18} \times \textit{Peripheral_Search_Copula}_i \\
& + \alpha_{19} \times \textit{Promotional_Search_Copula}_i \\
& + \sum_{k=20}^{25} \alpha_k \times \text{Day_Week}_{ik} + \sum_{m=26}^{28} \alpha_m \times \text{Time_Day}_{im} + \upsilon_i \quad (4)
\end{aligned}$$

Following Papies et al. (2017), we include five copula terms for the five endogenous independent variables in the model. Where, *Search_Speed_Copula*$_i$, *Product_Dispersion_Copula*$_i$, *Price_Dispersion_Copula*$_i$, *Peripheral_Search_Copula*$_i$, and *Promotional_Search_Copula*$_i$ are the copula terms for the respective endogenous variables. Finally, following Park and Gupta (2012) we use bootstrapped standard errors.

2.6.3 Results

The results for the estimation of no interactions or copula terms (Equation 2) and no copula terms (Equation 3) are shown in columns (1) and (2) of Table 5, respectively. Column (3) of Table 5 summarizes the results for our focal model in Eq. (4). We find that search speed and search product dispersion have significantly negative effects on the average price paid. The empirical results show that one standard deviation increase in search speed leads to a .4 cents (α = -.004; converting price from dollars to cents) decrease in the average price paid. Also, a standard deviation increase in search product dispersion results in a 1.4 cent decrease in the average price paid (α = -.014). On the contrary, search price dispersion has a significant positive impact (α = .009) on the average price paid. Specifically, a consumer with a standard deviation increase in search price dispersion tends to pay .9 cent more on average. In addition, a standard deviation increase in peripheral search propensity gives rise to a 2 cent decrease in the average price paid (α = -.020). The impact of promotional search propensity on price paid is significantly negative (α = -.010), thus, a standard deviation increase in promotional search propensity gives rise to a 1 cent decrease in the average price paid. Overall, the results in Table 5 provide strong empirical evidence for the main effects.

In terms of the interaction effects, the negative impact of search speed on the average price paid does not differ between food and non-food categories. The effect of search product dispersion on price paid is less negative in the food category (α = .009). Also, the positive impact of search price dispersion on price paid is weaker in the food category (α = -.004). In addition, the negative effect of peripheral search propensity on price paid is weaker in the food category (α

= .022). The negative effect of promotional search propensity on price paid is mitigated in food category (α = .009). In summary, the main effects are largely mitigated in the food category.

Table 5 also reveals some interesting findings on consumer demographics and time effects. For instance, older people tend to buy cheaper products (α = -.001). Older people may have higher purchasing power and product knowledge than younger people. Therefore, they are less likely to rely on price to make quality inferences, leading to lower price paid. We also observe day of week and time of day effects in the results. Shoppers are likely to pay more on Monday, Wednesday, and Thursday. In addition, people tend to pay less in the morning due to the fact that consumers make informed choices when they are probably more alert. An alternative explanation is that if people go shopping in the morning, they are less likely to be at work or in a rush. Thus, they are more likely to process the product information thoroughly, resulting in lower price paid.

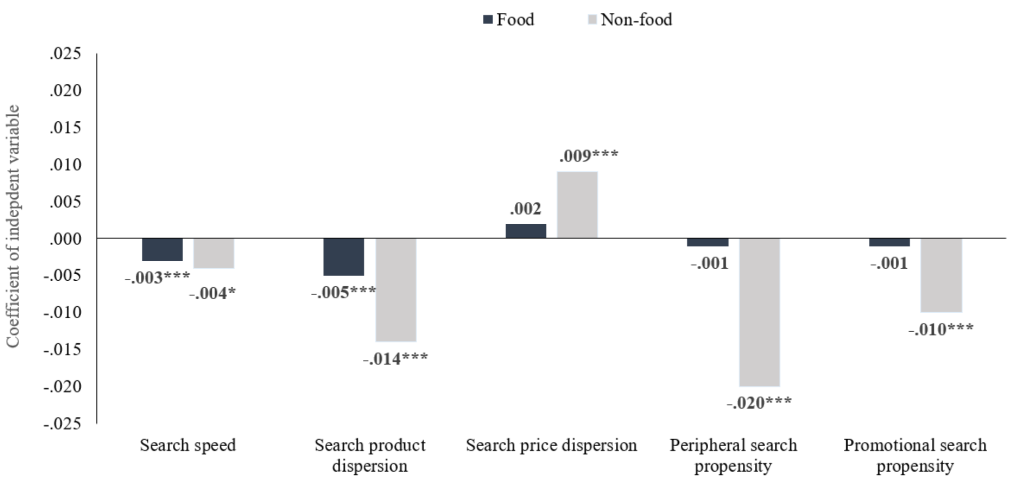
Table 5:
OLS Regressions and Regressions with Gaussian Copulas

	Average Price Paid					
	(1)		(2)		(3)	
	Coeff. (S.E.)	*Sig.*	*Coeff. (S.E.)*	*Sig.*	*Coeff. (S.E.)*	*Sig.*
Hypothesized Variables						
Search speed	-.001 (.000)	*	-.002 (.002)		-.004 (.002)	*
Search product dispersion	-.003 (.001)	***	-.010 (.003)	***	-.014 (.003)	***
Search price dispersion	.008 (.001)	***	.011 (.002)	***	.009 (.003)	***
Peripheral search propensity	.002 (.000)	***	-.015 (.004)	***	-.020 (.004)	***
Promotional search propensity	-.002 (.000)	***	-.010 (.002)	***	-.010 (.002)	***
Food category	-.017 (.003)	***	-.054 (.008)	***	-.053 (.008)	***
Food*Search speed			.001 (.002)		.001 (.002)	
Food*Search product dispersion			.009 (.003)	***	.009 (.003)	***
Food*Search price dispersion			-.004 (.001)	***	-.004 (.001)	***
Food* Peripheral search propensity			.022 (.005)	***	.022 (.005)	***
Food* Promotional search propensity			.009 (.002)	***	.009 (.002)	***
Control Variables						
Age	-.001 (.001)	**	-.001 (.001)	**	-.001 (.001)	**
Group shopping	.002 (.001)		.001 (.001)		.001 (.001)	
Shopping with a basket/cart	-.001 (.001)		-.001 (.001)		-.001 (.001)	
Monday	.005 (.002)	**	.005 (.002)	***	.005 (.002)	***
Tuesday	.001 (.002)		.001 (.002)		.002 (.002)	
Wednesday	.004 (.002)	**	.004 (.002)	**	.004 (.002)	**
Thursday	.004 (.002)	**	.004 (.002)	*	.004 (.002)	**
Friday	.003 (.002)		.003 (.002)		.003 (.002)	
Saturday	.003 (.002)		.002 (.002)		.003 (.002)	*
Q1	-.002 (.002)		-.002 (.002)		-.003 (.001)	*
Q2	.000 (.002)		.001 (.001)		.001 (.002)	
Q3	-.001 (.001)		-.002 (.001)		-.002 (.001)	
Copulas						
Search speed (Copula)					.002 (.001)	**
Search product dispersion (Copula)					.005 (.002)	***
Search price dispersion (Copula)					.004 (.005)	
Peripheral search propensity (Copula)					.004 (.001)	***
Promotional search propensity (Copula)					-.000 (.001)	
Model Fit						
F Statistic	12.620		18.710		14.892	
Adjusted R-squared	.207		.274		.165	

Notes: Coeff = Coefficient; SE = Standard Error; Sig = Statistical Significance; Food category = Cereal and Pasta; Q1= 9am-1pm, Q2= 1pm-4pm, Q3= 4pm-7pm; * = p < .10; ** = p < .05; *** = p < .01. Model (1) shows the OLS regression including independent and control variables. Model (2) indicates the OLS regression of interaction effects. Model (3) adds copula terms to the independent variables based on Model (2). All the independent variables and control variables except the ones that are dummy variables are standardized. For model (1) and (2), we use robust standard errors. For model (3), following Park and Gupta (2012), we use bootstrap standard errors with 1000 iterations.

Figure 4 probes the interaction effects and demonstrates the simple main effects of the five constructs in the food and non-food categories. It is clear that higher search speed leads to lower price paid in both food and non-food categories. However, the effect is much more significant in food category. While an increase in search product dispersion results in a decrease in price paid in both categories, the magnitude of the negative effect is much larger in the non-food category. Moreover, the positive effects of search price dispersion and the negative effects of peripheral search propensity and promotional search propensity on price paid are significant only in the non-food category. In summary, we find that non-food category dominates the overall effect.

Figure 4:
Probing the Interaction Effects



Notes: * = p < .10; ** = p < .05; *** = p < .01. All the independent variables are standardized.

2.6.4 Sensitivity Analysis

We conduct several robustness checks using alternative control variables and including significant copula terms only (see Table 6). First, following prior literature, we re-estimate Eq. (4) by only including significant copula terms, as the significance of a copula term illustrates the endogeneity issue for the focal variable (Mathys et al. 2016, Lim et al. 2018, Park and Gupta 2012). Second, we use bootstrap standard errors of 5,000 iterations instead of 1,000 iterations to conduct the analysis. Finally, we replace time of day controls with a weekday dummy variable. Specifically, it equals 1 when a consumer shops on weekdays, otherwise it is 0. Table 6 indicates that our results are highly robust and consistent with Table 5 across all these alternative specifications.

Table 6:
Sensitivity Analysis

	Average Price Paid					
	(1)		(2)		(3)	
	Coeff. (S.E.)	*Sig.*	*Coeff. (S.E.)*	*Sig.*	*Coeff. (S.E.)*	*Sig.*
Hypothesized Variables						
Search speed	-.004 (.002)	*	-.004 (.002)	*	-.004 (.002)	*
Search product dispersion	-.013 (.003)	***	-.014 (.003)	***	-.014 (.003)	***
Search price dispersion	.011 (.002)	***	.009 (.003)	***	.009 (.003)	***
Peripheral search propensity	-.020 (.004)	***	-.020 (.004)	***	-.019 (.004)	***
Promotional search propensity	-.010 (.002)	***	-.010 (.002)	***	-.010 (.002)	***
Food category	-.053 (.008)	***	-.053 (.008)	***	-.053 (.008)	***
Food*Search speed	.001 (.002)		.001 (.002)		.001 (.002)	
Food*Search product dispersion	.009 (.003)	***	.009 (.003)	***	.009 (.003)	***
Food*Search price dispersion	-.004 (.001)	***	-.004 (.001)	***	-.004 (.001)	***
Food* Peripheral search propensity	.022 (.005)	***	.022 (.005)	***	.022 (.005)	***
Food* Promotional search propensity	.009 (.002)	***	.009 (.002)	***	.009 (.002)	***
Control Variables						
Age	-.001 (.000)	**	-.001 (.001)	**	-.001 (.000)	**
Group shopping	.001 (.001)		.001 (.001)		.001 (.001)	
Shopping with a basket/cart	-.001 (.001)		-.001 (.001)		-.001 (.001)	
Weekday					.002 (.001)	**
Monday	.005 (.002)	***	.005 (.002)	***		
Tuesday	.002 (.002)		.002 (.002)			
Wednesday	.004 (.002)	**	.004 (.002)	**		
Thursday	.004 (.002)	**	.004 (.002)	**		
Friday	.003 (.002)		.003 (.002)	*		
Saturday	.003 (.002)	*	.003 (.002)	*		
Q1	-.003 (.002)	*	-.003 (.001)	*	-.003 (.001)	*
Q2	.001 (.002)		.001 (.001)		.001 (.002)	
Q3	-.002 (.001)		-.002 (.001)		-.002 (.001)	
Copulas						
Search speed (Copula)	.002 (.001)	**	.002 (.001)	**	.002 (.001)	**
Search product dispersion (Copula)	.005 (.002)	***	.005 (.002)	***	.005 (.002)	**
Search price dispersion (Copula)			.004 (.005)		.004 (.005)	
Peripheral search propensity (Copula)	.004 (.001)	***	.004 (.001)	***	.004 (.001)	***
Promotional search propensity (Copula)			-.000 (.001)		-.000 (.001)	
Model Fit						
F Statistic	21.606		15.139		18.603	
Adjusted R-squared	.214		.172		.171	

Notes: Coeff = Coefficient; SE = Standard Error; Sig = Statistical Significance; Food category = Cereal and Pasta; Q1= 9am-1pm, Q2= 1pm-4pm, Q3= 4pm-7pm; * = p < .10; ** = p < .05; *** = p < .01. Model (1) includes significant copula terms only. Model (2) conducts analysis using use bootstrap standard errors with 5000 iterations. Model (3) replaces day of week control variables with an alternative control variable (i.e., weekday). Weekday equals to one when a shopper visited the store during weekday, otherwise it is zero. All the independent variables and control variables except the ones that are dummy variables are standardized.

2.7 Discussion

Haptic search plays a pivotal role in consumers' in-store search behavior. Thus, offline retailers are keen on understanding the implications of consumers' haptic search with an aim to incorporate changes within a store. The present study seeks to examine and offer guidance on how multiple dimensions of consumer haptic search (i.e., search speed, search product dispersion, search price dispersion, peripheral search propensity, and promotional search propensity) affect the average price paid. These dimensions reflect the speed, consideration set, and shopping path of a shopper navigating through the category in a store. Furthermore, we provide a nuanced understanding of the impact of these dimensions on price paid by investigating the moderating role of category (food versus non-food).

This paper also contributes to the tracking literature by adopting a state-of-the-art sensing technology to record consumer in-store haptic search behaviour. The data collection approach has several merits over existing tracking technology. First, unlike eye-tracking technology which tracks the movement of eyeballs and is often used to capture attention, our sensing technology tracks the movement of hands and captures consumer search effort in the shelf space. More importantly, our methodology is non-intrusive and can be applied in real shopping environments. Second, relative to RFID technology that tracks the movement of shopping baskets/carts, our sensing technology captures accurate movement of hands and addresses the issues associated with moving while leaving the baskets/carts behind. Third, compared to video tracking technology, our approach protects consumer privacy by not capturing personal information of shoppers. Overall, our study advances several important contributions to the marketing literature.

2.7.1 Theoretical Contributions

Prior research on the role of haptic search on price paid has been debatable. While some studies suggest that consumer haptic search lowers the price paid, others find that greater haptic search efforts lead to an increase in purchase likelihood and higher average selling prices (e.g., Ratchford and Srinivasan 1993, Seiler and Pinna 2017, Hui, Bradlow and Fader 2009, Ngwe et al. 2019). Using five key attributes of shoppers' haptic search at the shelf space, which reflect speed, consideration set, and shopping path, we show differential and significant effects of these attributes on the average price paid by the shopper.

Our proposed constructs capture distinct aspects of shopper haptic search. Specifically, search speed involves two important dimensions of shopper haptic search effort, i.e., the number of products considered by the consumer and the time spent in considering these products (Malhotra 1982, Larson et al. 2005). Despite the interest in shopping path in prior work, insufficient attention has been paid to shopper speed (Larson et al. 2005, Van Den Bergh et al. 2016). Moreover, given that consumers tend to "waste" 80% of the time in a store, we focus on the speed of product search that captures search effort more accurately than walking speed examined in prior work (Sorensen 2009, Van Den Bergh et al. 2016, Zhang et al. 2014). In sharp contrast to prior work that emphasizes the need for retailers to increase shopper speed, we find that higher shopper speed at the shelf space lowers price paid per unit (Seiler and Pinna 2017, Zhang et al. 2014). This study, therefore, brings to fore the difference in shopper path at the overall store level as compared to the path followed by the shopper at the shelf space. Moreover, we account for the diversity of a shopper's consideration set during the haptic search process by proposing two constructs, search product dispersion and search price dispersion. The more diverse the consumers' consideration set, the higher the probability that the consumer is not bound to a specific choice and therefore the higher the price elasticity of the shopper (Mitra and Lynch 1995). As such, we contribute to the consumer consideration set literature via proposing two unique constructs that capture the diversity of a consumer's consideration set during haptic search (also see Mazumdar and Jun 1993, Sorce and Widrick 1991).

In addition, peripheral search propensity and promotional search propensity examine the effect of consumer's shopping path on purchase behavior. Extant research has shown that products displayed on the center of a shelf are likely to be of higher quality (Raghubir and Valenzuela 2006), more popular (Valenzuela and Raghubir 2009), gain more attention, and therefore more likely to be purchased (Chandon et al. 2007, 2009). Thus, we find that peripheral search propensity of a consumer influences their price paid. Also, the earlier a shopper is exposed to promotions during their search, the more likely that a consumer is active in locating in-store promotions and is sensitive to in-store deals (Schneider and Currim 1991). We, therefore, find that promotional search propensity influences the average price paid by a consumer. To the best of our knowledge, this paper is the first to study and demonstrate the impact of various dimensions of haptic search in a brick-and-mortar setting.

Additionally, we find that the type of category (food versus non-food), in conjunction with haptic search plays a vital role. Our findings indicate that while higher search speed lowers price paid in both food and non-food categories, the effect is more significant in the food category. The magnitude of the negative impact of search product dispersion on price paid is much larger in the non-food category than food category and the positive impact of search price dispersion on price paid is significant only in the non-food category. Moreover, we find that an increase in peripheral search propensity leads to a smaller decrease in average price paid in the food category. Finally, the negative effect of promotional search propensity on price paid is mitigated in food category. Thus, our research delves deeper into consumer's haptic search behavior by investigating the boundary conditions for the five proposed constructs of consumer search.

2.7.2 Managerial Implications

Conventional retail stores face the challenge of competing with e-commerce and more than 45 US retail chains have gone bankrupt in the last three years (McKinsey 2020). Furthermore, retailers have been struggling with maintaining profitability due to rising cost pressures from higher cost of property, staff, fuel, commodity prices and the investment in digitization to remain competitive (Deloitte 2017). Our findings provide several managerial implications for brick-and-mortar retailers. Specifically, our study indicates the critical role of lowering consumer search speed in driving profitability. Indeed, a small group of retailers (e.g., Origins, a beauty brand) strive to "slow down" customers in their stores (Eckhardt and Husemann 2018). We find that an additional item searched in a second results in a decrease in the price paid per unit by .4 cents. To capture the diversity of consumer consideration set embedded in product search and price search efforts, we examine the impacts of search product dispersion and search price dispersion. Results suggest that whereas a unit increase in search product dispersion reduces the average price paid by 1.4 cents, a unit increase in search price dispersion gives rise to an increase in the average price paid by .9 cents. Also, an increase in peripheral search propensity leads to a 2 cent decrease in the average price paid and an increase in promotional search propensity leads to a 1 cent decrease in the average price paid.

Taken together, our study indicates that consumer consideration set has mixed effects on the price paid by consumers in a category during a shopping trip. On one hand, search product dispersion lowers the average price paid. On the other hand, search price dispersion increases the price paid. Category management aiming at optimizing sales and margins has remained almost unchanged for decades and changes concerning shopper-centric approaches are in need (Cloud 2015). Contrary to the conventional view that larger category assortment is beneficial, our findings demonstrate that offering a smaller shelf allows shoppers to expend less effort to identify their desired choices. However, when it comes to the optimization of prices across unique products, it is essential to enlarge the price variations across distinct SKUs in a category. Overall, a smaller assortment with larger price variations is likely to boost category profits for brick-and-mortar retailers. A prudent approach that retailers can undertake is to explore a greater balance in the number of SKU offerings and investigate the average price paid within the store, while closely monitoring the impact on category demand. In addition, our findings about shopper shopping path, in a category, influencing price paid indicates that shoppers starting their search farther away from the center of the category or with a promoted product are more likely to pay a lower price. Thus, in-store sales representative should direct shopper attention more toward the center of the category and non-promoted products while assisting shoppers during their search process.

Our study also provides interesting implications for category managers of food and non-food categories. Specifically, in terms of shelf layout, non-food category managers need to recognize the importance of shelf center to engage shoppers and possibly not highlight promoted items to grab shopper attention. Additionally, category managers can aim to increase the price range of product offerings, reduce assortment sizes, or include product recommendations in non-food category. We also believe that our measures and the associated implications can be very useful for managers in other categories like apparel category, where haptic search plays a pivotal role.

2.7.3 Limitations and Future research

The present work has several limitations that provide directions for future research. In the present study, we were unable to capture consumer haptic search across different categories. Future research can examine cross-category search behavior by consumers. This would be an interesting extension to our work and could shed light on understanding shopper behavior in the entire shopping journey. To further our studies, causality can be established by conducting an experiment and our results provide a starting point for such explorations. It would, therefore, be interesting to conduct field experiments based on our findings wherein retailers could make changes to their assortment (e.g., reduce products while increasing price dispersion) and gain further insights on in-store shopping behavior. We also assume that higher price paid by shoppers is better for retailers, however, it is dependent on the margins that retailers obtain from different brands as it is possible that lower priced brands could have higher margins and thus beneficial to retailers.

Our data collection using tracking technology helps us in figuring out exactly when consumers touch various products, however, we do not know the exact reason why a shopper is touching a product multiple times. We surmise that it is with an intention to investigate the product further, but it could also mean that they are unsure. Future studies could augment tracking data with shopper exit surveys to understand motivations for consumer actions. Finally, due to the constraints related to privacy (since shopper identifying information cannot be stored) we are unable to identify the same shoppers who returned to the category shelf after exiting the category. In other words, a shopper who picks up an item, leaves the category and comes back in a while to pick up another item would be considered a new shopper. However, in this study, we ensure that we minimize the impact of these returning shoppers by eliminating problematic records such as shoppers who returned an item first before picking up a product. In the future, it would be worthwhile to investigate such “revisit” behavior with more advanced technology. Notwithstanding these limitations, we believe that this research provides novel insights that will be of significant interest to both academicians and practitioners alike.

Chapter 3 Unintended Consequences of Recreational Cannabis Legalization Across Categories[5]

3.1 Introduction

Recent years have witnessed the substantial growth in cannabis sales in the US, reaching $17.5 billion in 2020, a 46% increase from 2019 (Yakowicz 2021). As the first state to allow the retail sale of recreational cannabis, Colorado has seen more than $12 billion in cumulative sales by 2021, with approximately $2.23 billion in the calendar year 2020 (Colorado Department of Revenue 2022a). Also, Colorado achieved cannabis tax revenue of about $2.02 billion by 2021 (Colorado Department of Revenue 2022b). The sales and associated tax revenues are positioned as "…a massive success and proof-of-concept for the future of the American cannabis industry" (Hoban 2021). Indeed, several states have followed suite and legalized cannabis in the last few years, with 18 states and Washington, D.C., legalizing recreational cannabis and a further 13 states decriminalizing its use.[6] In fact, a recent draft bill introduced in July 2021, called Cannabis Administration and Opportunity Act, aims to legalize cannabis in all states across US and could completely end federal prohibition.[7]

Surprisingly, the state of Colorado just passed a state bill (H 1317), which is applying limits to the state's medical marijuana industry (with requirements like reduction of sale of high potency marijuana to one-fifth of the current levels, real-time monitoring of sales, requirement of warning labels

[5] Researchers own analyses calculated (or derived) based in part on data from Nielsen Consumer LLC and marketing databases provided through the NielsenIQ Datasets at the Kilts Center for Marketing Data Center at The University of Chicago Booth School of Business. The conclusions drawn from the Nielsen data, as well as any errors or omissions, are those of the researchers and do not reflect the views of Nielsen. Nielsen is not responsible for, had no role in, and was not involved in analyzing and preparing the results reported herein.
[6] https://ballotpedia.org/Marijuana_laws_in_the_United_States
[7] https://www.reuters.com/world/us/us-senate-democrats-release-discussion-draft-federally-legalize-cannabis-2021-07-14/

etc.).[8] The debate for recreational cannabis legalization (RCL) has been ongoing with greater tax revenues, reduction in crime, stimulation of economy used as an argument in support of the legalization. In contrast, health concerns (due to addictiveness and cannabis acting like a gateway drug) are used as arguments against RCL (Dills et al. 2021).

However, beyond tax revenues, cannabis consumption can also have potential spillover effects due to crossfading, i.e., co-consumption of cannabis and alcohol / tobacco (Schauer et al. 2015, Barrett et al. 2006). Prior literature has demonstrated the importance of investigating the cross-category effects associated with consumers' purchase decisions (e.g., Manchanda et al. 1999, Song and Chintagunta 2006, Goli and Chintagunta 2021). Also, there is a potential for substantial health and policy implications due to spillover effects of cannabis legalization across categories. Indeed, Schlienz and Lee (2018) provide a detailed overview of the risks and health implications of co-use of cannabis, alcohol, and tobacco. While there is research highlighting that increased appetite due to cannabis consumption, research investigating the impact of co-use of cannabis and junk food consumption is limited (Hull 2019). According to Kruger et al. (2019), the majority of participants surveyed opted for junk food (i.e., chips) rather than healthy food (i.e., orange). Baggio and Chong (2020) conclude that the sales of high calorie food (i.e., ice cream, cookies, and chips) went up by 4.5% in volume post RCL. However, prior literaure relies on survey data (e.g., DiNardo and Lemieux 2001, Clements and Daryal 2005, Grant et al. 2018, Kruger et al. 2019) and online search data (e.g., Wang et al. 2019) to conduct analysis, indicating insufficient attention of

[8] https://www.hudson.org/research/17052-the-colorado-experiment-legalized-marijuana-s-impact-in-colorado

examining actual sales data in the marketing literature. Furthermore, the impact of RCL on the sales of related substances may be misleading without ruling out alternative explanations (e.g., adjustment in pricing or advertising).

Accordingly, we investigate the cross-category impact of RCL on retail sales of alcohol, tobacco, and munchies or junk food (i.e., candy and salty snacks). This is important because the impact of RCL on consumption across these different categories not only has revenue and tax implications but also has health care related consequences. In this study, we focus on the role of RCL in Colorado, the first state in the US where legal recreational marijuana sales began[9], resulting in the lack of expectations for consumers and marketers in Colorado. Consumers, therefore, are unlikely to alter their consumption of related substances via learning the behavior of customers from other states. Additionally, it is unlikely that how marketers would respond to the legislation change by adjusting pricing or advertising is affected by the reactions of marketers from other states. As such, RCL of Colorado, compared with RCL of other states, is considered as truly exogenous.

We obtain sales and pricing data from Nielsen retail scanner data as well as extracting advertising information from Nielsen ad intel database. Empirically, we utilize synthetic control method (SCM) which is "arguably the most important innovation in the evaluation literature in the last 15 years" (Athey and Imbens 2017, p. 9). The advantage of SCM is that it does not rely on the parallel trends assumption, a key assumption of Difference-in-Differences (DiD) approach. In addition, SCM accounts for time varying

[9] https://www.denverpost.com/2014/01/01/worlds-first-legal-recreational-marijuana-sales-begin-in-colorado/

confounders between the treated and the control groups by balancing the groups based on pre-treatment covariates and outcomes (Abadie et al. 2010).

Our empirical analysis reveals that RCL has a significant positive impact on both per capita dollar sales and unit sales of alcohol, candy, and salty snacks, but not on the sales of tobacco. Specifically, we find that by June 2014 (compared to the first week of 2014), the per capita dollar sales and unit sales for alcohol (28.92% and 19.18%), candy (25.73% and 27.71%) and salty snacks (19.09% and 20.96%) were significantly higher than what it would have been in the absence of RCL. To rule out the possibility that RCL has general spillover effects on all the categories, we show that RCL does not affect batteries category, which can be seen as a "null category". To further validate our findings, we also examine other underlying factors that could impact sales. Specifically, we find that weekly advertising spending and prices remain almost unchanged across all the categories in response to RCL availability, ruling out alternative reasons for sales increase.

Taken together our research provides guidance for policy makers and marketers. Specifically, policymakers need to pay close attention to the potential increase in healthcare expenses and social issues related to increased alcohol consumption and junk food consumption. Increase in alcohol consumption could lead to various issues including chronic diseases and impact on mental health. As detailed by Mayo clinic these issues include liver disease, heart problems, birth defects, digestive problems, bone damage, and neurological complications to just name a few.[10] The total cost associated with alcohol problems is $175.9 billion a year, with just the annual health care

[10] https://www.mayoclinic.org/diseases-conditions/alcohol-use-disorder/symptoms-causes/syc-20369243

expenditure for alcohol issues amounting to $22.5 billion.[11] In addition, cannabis use increases appetite and our findings of increased consumption of junk food (containing excessive sugar and salt) due to RCL can lead to obesity and other health risks like diabetes. Based on a National Heart, Lung, and Blood Institute (NHLBI) funded study bad eating habits cost about $50 billion a year in health care costs mainly attributable to heart disease, stroke, and diabetes.[12] It is therefore critical for policy makers trying to legalize cannabis to understand that the positive tax revenue impact from cannabis sales could be offset by the higher health costs due to increased sales and consumption of alcohol, candy, and salty snacks. Finally, while RCL increases tobacco online search volume (Wang et al. 2019), the increase in search does not translate into higher retail sales of tobacco in the case of Colorado. That is, we find tobacco is neither a complement nor a substitute for cannabis.

3.2 Background

This study draws on literature on the relationship between marijuana, alcohol, and tobacco. A stream of research suggests that marijuana and alcohol/ tobacco are likely to be substitutes (see DiNardo and Lemieux 2001, Clements and Daryal 2005, Wang et al. 2019, Miller and Seo 2018). DiNardo and Lemieux (2001) find that increases in the legal minimum drinking age lowered alcohol consumption based on a large sample consisting of students from 43 states between 1980–1989. Using student survey data, Clements and Daryal (2005) demonstrate that the legalization of marijuana reduced the consumption of beer, wine and spirits by 1%, 2% and almost 4%, respectively. Wang et al. (2019) illustrate that online search volume and advertising effectiveness for alcohol

[11] http://www.alcoholpolicymd.com/alcohol_and_health/costs.htm
[12] https://www.nhlbi.nih.gov/news/2019/americans-poor-diet-drives-50-billion-year-health-care-costs

decreased after RCL took effect. Miller and Seo (2018) find a 5% reduction in alcohol demand and a 12% decrease in tobacco demand after the implementation of RCL. However, other prior work suggests the complementary relationship between marijuana, alcohol, and tobacco (see Grant et al. 2018, Wang et al. 2019, Bhave and Murthi 2019). For instance, using a sample of pregnant and parenting women, Grant et al. (2018) find that alcohol use is associated with RCL of Washington which took effect in 2012. Wang et al. (2019) conclude that online search volume and advertising effectiveness for tobacco went up after RCL took effect. Similarly, Bhave and Murthi (2019) find that the implementation of RCL leads to a 4%-7% increase in cigarette consumption in Colorado. Given these mixed findings on the substitute or complementary relationship between marijuana, alcohol, and tobacco, it is crucial to obtain a better understanding of this discrepancy.

According to research investigating the impact of cannabis on appetite (Hull 2019), Tetrahydrocannabinol (THC), an active and main ingredient in cannabis, can bind to and activate cannabinoid receptor type 1 (CB1). CB1 is considered as one of the main reasons for increase in appetite post cannabis consumption. In addition, this increase in appetite is expected to influence preference for certain types of products like sweet, salty, or sour foods (Hull 2019). If RCL leads to an increase in sweet food consumption, then health issues like diabetes could be on the rise in the long-term. Similarly, increased consumption of salty snacks due to cannabis consumption could lead to various heart related diseases (National Institute on Alcohol Abuse and Alcoholism 2021). Given the significant health effects, it is critical to understand the relationship between cannabis and junk food or munchies categories.

More importantly, prior work mainly uses survey data (e.g., DiNardo and Lemieux 2001, Clements and Daryal 2005, Grant et al. 2018, Kruger et al. 2019) and online search data (e.g., Wang et al. 2019). While some economic literature has investigated the impact of RCL on related substances using actual sales data, there is a lack of attention in the marketing literature (Miller and Seo 2018, Bhave and Murthi 2019, Baggio and Chong 2020). In addition, merely examining the impact of RCL on sales of related substances could be misleading. For example, if alcohol manufacturers increase prices or cut advertising during the launch period, the decrease in sales could be attributed to both changes in marketing mix and launch of recreational cannabis. Hence, it is important to investigate the role of the launch of recreational cannabis along with the changes to the marketing mix undertaken by the manufacturers across various categories. Our study embarks on this route and extends the literature by examining the unintended consequences of RCL across various related categories using data of sales outcomes and marketing mix (e.g., pricing and advertising) to get a holistic view as policy makers make decisions about further legalization at the federal level.

3.3 Experimental Design and Data Collection

3.3.1 Quasi Experiment

Prior literature uses the timing of changes in legislations or regulations as a source of exogenous variation and the "stable unit treatment value assumption" (SUTVA) requires a valid control group which does not suffer from the spillover effects of the exogenous treatment (Goldfarb et. al. 2022, Angrist et al. 1996). Focusing on RCL of Colorado, the first state to implement RCL, provides us with a clean quasi-experimental research setting. If we choose RCL of any other states (e.g., Oregon), the new treated and control states may serve as the control states of RCL of Colorado and hence are likely to suffer from the spillover effects associated with RCL of Colorado. Also, consumers and marketers of these states may learn from the ones in Colorado when altering their consumption and marketing actions (e.g., pricing and advertising) of related substances, resulting in unclean effects.

While RCL of Colorado took effect from December 10th, 2012, we focus on January 1st, 2014, which is the date when marijuana was available for retail sales and therefore is more relevant to our outcome variables of interest (Miller and Seo 2018, Bhave and Murthi 2019). We use data from one year before and half year after this date as the pre- and post-RCL periods, respectively. As Colorado and Washington passed measures to legalize marijuana for recreational use in December 2012[13], the pretreatment period starts from January 2013 to ensure that there is no change in marijuana legislations. Furthermore, we stick to the half-year posttreatment period to eliminate potential spillover effects of Washington which allowed recreational marijuana for sale as of July 8th, 2014.[14]

3.3.2 Data

The two main data sources used in this research are Nielsen retail scanner data and Nielsen ad intel data. The retail scanner data includes pricing and quantities sold at UPC-store-weekly level across the US. The ad intel data contains information about ad spend at brand-weekly level in the US. We focus on alcohol, tobacco, candy, and salty snacks in our analysis and our sample period

[13] https://www.theguardian.com/world/2012/nov/09/colorado-washington-legalise-marijuana#:~:text=Colorado%20amendment%2064%20passed%20on,and%20sale%20of%20the%20drug.
[14] https://www.nytimes.com/2014/07/09/us/washington-to-begin-sales-of-recreational-marijuana.html

starts from January 2013 to June 2014.[15] Per capita dollar sales and per capita unit sales are the main outcome variables of interest in our analysis.[16] We also study the effect of RCL on unit price and ad spend to rule out alternative explanations and demonstrate the validity of our findings. All the variables are aggregated to the state-weekly level.[17]

3.3.3 Choice of Predictor Variables

Following Abadie et al. (2010, 2015), we use both state-level variables and variables that are likely to affect outcome variables as predictors in the main analysis. The state-level predictors include log-transformed per capita personal income (quarterly), unemployment rates (quarterly), percentage of population above 20 (annual), GINI index (annual), violent crime rates (annual), and property crime rates (annual)[18]. Per capita personal income is expected to be positively associated with all the outcome variables because high personal income per capita indicates higher purchase power. Unemployment rates are likely to reduce all the outcome variables as high unemployment rates signal low-income levels. The percentage of population above 20 is included as recreational marijuana can only be sold to adults who are above 21 years old[19]. Additionally, we include GINI index which is an indicator of social inequality. Finally, we include violent crime rates and property crime rates which are likely

[15] Given the offline setting, we remove subcategories with website sales. We focus on main categories and thus remove associated items such as tobacco accessories. We do not include e-cigarette in the tobacco category as it applies different tax rates. While there is no category called "salty snacks" in the database, we refer to a combination of crackers, snacks, and nuts, that are salty, as constituting this category.

[16] Per capita dollar sales refers to the ratio of dollar sales to total population. Per capita unit sales is measured as the ratio of unit sales to total population. We also use alternative measurement (i.e., the ratio of dollar sales or unit sales to population above 20) to show the robustness of our results.

[17] We first aggregate the data to state-weekly level and then deal with the missing values. Following Kan and Usher (2020), we replace missing dollar sales and unit sales with zero. We replace missing unit price with the moving average (Jindal et al. 2020). Specifically, we calculate the mean of unit price of one month before and after the missing observation. In terms of ad spend, we replace missing values with zero (Rao and Wang 2017).

[18] We obtain the per capita personal income, percentage of population above 20 and GINI index from US Bureau of Economic Analysis. We retrieve unemployment rates from Federal Reserve Economic Data. The data source of violent crime rates and property crime rates is Federal Bureau of Investigation.

[19] We use the percentage of population above 20 rather than the percentage of population above 21 due to data availability.

to increase marijuana sales. For per capita dollar sales and per capita unit sales, we further include unit price and ad spend as predictor variables to construct the synthetic controls. Finally, for all the outcome variables, we follow prior literature to use the mean of lagged outcome variables across the pretreatment period as an additional predictor (Lu et al. 2021, Tirunillai and Tellis 2017). The descriptive statistics for all the variables are provided in Table 7.

Table 7:
Descriptive Statistics

	Min	Max	Mean	Standard deviation
Panel A: Alcohol				
Per capita dollar sales	.00	.95	.15	.13
Per capita unit sales	.00	.09	.02	.01
Unit price (log)	2.17	3.10	2.45	.15
Ad spend (log)	.00	15.50	10.66	2.16
Per capita personal income (log)	10.43	11.09	10.70	.16
Unemployment rates	4.00	9.53	6.93	1.14
Percentage of population above 20	.72	.78	.75	.01
Gini index	.44	.51	.47	.02
Violent crime rates	.00	.01	.00	.00
Property crime rates	.02	.04	.03	.01
Panel B: Tobacco				
Per capita dollar sales	.00	.10	.03	.02
Per capita unit sales	.00	.02	.00	.00
Unit price (log)	2.02	3.17	2.52	.20
Ad spend (log)	.00	9.70	2.35	3.09
Per capita personal income (log)	10.43	10.96	10.67	0.14
Unemployment rates	4.00	9.57	7.14	1.19
Percentage of population above 20	.72	.77	.75	.01
Gini index	.44	.51	.47	.01
Violent crime rates	.00	.01	.00	.00
Property crime rates	.02	.04	.03	.01
Panel C: Candy				
Per capita dollar sales	.01	.36	.06	.04
Per capita unit sales	.01	.18	.03	.02
Unit price (log)	1.05	1.51	1.25	.07
Ad spend (log)	3.68	11.82	9.06	1.04
Per capita personal income (log)	10.43	11.09	10.69	.16
Unemployment rates	4.00	9.57	7.04	1.20
Percentage of population above 20	.72	.77	.75	.01
Gini index	.44	.50	.47	.01
Violent crime rates	.00	.01	.00	.00
Property crime rates	.02	.04	.03	.01
Panel D: Salty Snacks				
Per capita dollar sales	.01	.53	.12	.07
Per capita unit sales	.00	.18	.04	.02
Unit price (log)	1.28	1.51	1.43	.04
Ad spend (log)	.00	12.68	9.70	1.28
Per capita personal income (log)	10.43	11.09	10.69	.16
Unemployment rates	4.00	9.57	7.00	1.20
Percentage of population above 20	.72	.78	.75	.01
Gini index	.44	.50	.47	.01
Violent crime rates	.00	.01	.00	.00
Property crime rates	.02	.04	.03	.01

Notes: The above variables are calculated at state-weekly level. Unit price, ad spend and per capita personal income are log transformed. Per capita dollar sales is calculated as the ratio of dollar sales to total population of a state in a week. Per capita unit sales is measured as the ratio of unit sales to total population of a state in a week. We obtain the per capita personal income, percentage of population above 20 and GINI index from US Bureau of Economic Analysis. We retrieve unemployment rates from Federal Reserve Economic Data. The data source of violent crime rates and property crime rates is Federal Bureau of Investigation.

3.3.4 SCM Model Specification

We adopt Synthetic Control Methods (SCM) which relaxes the parallel trends assumption to build causal inference. To investigate the changes in the outcome variables (i.e., per capita dollar sales and per capita unit sales) during the post-RCL period, we use SCM to assess the effect of RCL on sales performance of related categories (i.e., alcohol, tobacco, candy, and salty snacks). Based on Abadie et al. (2010, 2015), we define the synthetic control state as a weighted average of the J states in the control group. Let $i = 1$ be the focal treated state (Colorado) and $i \in [2, \ldots, J+1]$ represent the potential control states. The predicted outcome in the synthetic control is given by $\sum_{i=2}^{J+1} w_i y_{it}$, where $\boldsymbol{W} = \left(w_2, \ldots, w_{J+1}\right)$ is a $J \times 1$ vector of state-specific nonnegative weights that sum up to 1 and minimize the difference between the characteristics of the treated state and the synthetic control during the pretreatment period, i.e.,

$$\boldsymbol{W} = argmin_{\boldsymbol{W} \in \boldsymbol{R}_+} \sum_{m=1}^{M} v_m (X_{1m} - \boldsymbol{X_{0m}} \boldsymbol{W}')^2 \quad s.t.\ \ \boldsymbol{1}_J' \boldsymbol{W} = 1$$

where $\boldsymbol{X_{1m}}$ is the value of the m[th] pretreatment characteristic of the treated state, $\boldsymbol{X_{0m}}$ is the corresponding vector of the same characteristic of the control states, and v_m is a weight measuring the relative importance of each characteristic in matching the treated unit and the synthetic control. Abadie et al. (2010, 2015) suggest that v_m can be chosen by minimizing the mean squared prediction error (MSPE) of the outcome variable in the pretreatment period so that a characteristic with greater prediction power should be assigned with a larger weight. Based on the estimated weights, we can assess the treatment effect by simply calculating the outcome gap between the treated unit and the synthetic control:

$$y_{1t} - \sum_{i=2}^{J+1} w_i y_{it}$$

3.3.5 Identifying Assumptions and Choice of Donor Pool

The formation of our donor pool constructing the synthetic control state should meet the following assumptions (Abadie et al. 2010, 2015). First, the synthetic control state should follow similar trends with the treated state (i.e., Colorado) during the pretreatment period. We demonstrate that this assumption is verified in the results section. Second, only the treated state (i.e., Colorado) undergoes the intervention in the posttreatment period. Washington is the first state to legalize recreational marijuana in the US despite Colorado being the first state where marijuana was available for purchase. Nevertheless, we remove Washington to ensure that none of the control states should undergo similar intervention by June 2014. The third assumption is that the outcome variables in control states constructing the synthetic control state should not be affected by RCL in Colorado and therefore there should not be any spillover effects. As such, we remove states that are within 1,000 miles away from Colorado to mitigate the effects of tourism. This is because people living in these states were likely to drive or fly to Colorado to purchase marijuana.[20] Finally, the treated state (i.e., Colorado) and control states should not undergo alternative changes affecting the outcome variables during the posttreatment period.

We not only investigate the changes in state exercise tax rates, but also examine additional policy or law changes. For alcohol category, we discard Rhode Island increasing state wine and spirits exercise tax rate, Tennessee

[20] Ideally, states that have direct flights to Colorado should be removed from the donor pool. Given that all the states have nonstop flights to Colorado, we get rid of states that are within 1,000 miles away from Colorado based on distance shown by Google Map. Generally, it should take fewer than 4 hours to fly and less than one day to drive through 1,000 miles. The purpose of doing so is to reduce the effects of visitors from other states.

increasing state beer exercise tax rate, and Washington decreasing state beer exercise tax rate during the posttreatment period. In addition, we also discard Texas which increased "ad valorem excise taxes" and Georgia which adopted child abuse/neglect provision. Similarly, for tobacco category, we remove Connecticut, Minnesota, New Hampshire, Oregon and Massachusetts which increased state cigarette exercise tax rate and Indiana which decreased state cigarette exercise tax rate during postintervention period. We have also examined other policy changes (e.g., state-level smoking bans and change in tobacco minimum purchase age) and found that Oregon started to ban smoking in vehicles with passengers under 18 inside as of January 1, 2014.

In terms of candy and salty snacks, we first discard Kansas and West Virginia which decreased state candy exercise tax rate, Maine which increased state candy exercise tax rate for candy category, and removed Kansas, Tennessee, and West Virginia decreasing state chips/ pretzels exercise tax rate during posttreatment period. Moreover, we get rid of California which amended existing law that required including a nutritious snack in the high school after-school programs and a physical activity element, Delaware which amended prior law to permit a wellness plan to be offered through health insurance if it meets certain requirements and New York which encouraged the production and consumption of fresh locally produced fruits and vegetables by elementary and secondary school aged children to help combat the increasing incidence of childhood obesity for both categories. Finally, given the lack of information about Alaska and Hawaii in the data, the donor pool of alcohol, tobacco, candy and salty snacks has 27, 24, 25, and 25 control states, respectively. Thus, while satisfying all the assumptions of the SCM, we still have a relatively large donor pools for each of the categories that we investigated. The detailed steps regarding the formation of our donor pool across categories are summarized in Table 8.

Table 8:
Selection of Control States to Form Donor Pool

	Number of States	List of States	Notes
Number of control states	50-1		Colorado is the treated state.
Lack of data	2	AK, HI	
Similar intervention	1	WA	We remove WA which legalized recreational use of marijuana in December 2012, despite that it allowed retail sale of recreational marijuana in July 2014.
Spillover effects regarding tourism	16	AZ, UT, WY, NE, KS, OK, TX, NM, NV, ID, MT, ND, SD, MN, IA, MO	We strive to eliminate spillover effects associated with people visiting Colorado to purchase marijuana by removing states that are within 1,000 miles away from Colorado based on estimated distance shown in Google Map.
Panel A: Alcohol			
State exercise tax rate change	3-1[21]	RI, TN, WA	During post-RCL period, RI increased state wine and spirits exercise tax rate, TN increased state beer exercise tax rate whereas WA decreased state beer exercise tax rate. We obtain information about tax rate change from Tax Policy Center (https://www.taxpolicycenter.org/statistics/state-alcohol-excise-tax-rates).
Other policy change	3-2[22]	TX, GA, TN	During post-RCL period, TX increased “ad valorem excise taxes” (taxes levied on the price of a beverage), GA adopted child abuse/neglect provision and TN changed laws specifying requirements or incentives for retail alcohol outlets to participate in server training programs. We obtain related information from Alcohol Policy Information System (https://alcoholpolicy.niaaa.nih.gov//s).
Number of control states in the donor pool		27	

[21] -1 means getting rid of the duplicated calculation of WA.
[22] -2 means getting rid of the duplicated calculation of TX and TN.

Table 8:
Selection of Control States to Form Donor Pool (Continued)

	Number of States	List of States	Notes
Number of control states	50-1		Colorado is the treated state.
Lack of data	3	AK, HI, ME	
Similar intervention	1	WA	We remove WA which legalized recreational use of marijuana in December 2012, despite that it allowed retail sale of recreational marijuana in July 2014.
Spillover effects regarding tourism	16	AZ, UT, WY, NE, KS, OK, TX, NM, NV, ID, MT, ND, SD, MN, IA, MO	We strive to eliminate spillover effects associated with people visiting Colorado to purchase marijuana by removing states that are within 1,000 miles away from Colorado based on estimated distance shown in Google Map.
Panel B: Tobacco			
State exercise tax rate change	6-1[23]	CT, MN, NH, OR, MA, IN	During post-RCL period, CT, MN, NH, OR and MA increased state cigarette exercise tax rate, whereas IN decreased state cigarette exercise tax rate. We obtain information about tax rate change from Tax Policy Center (https://www.taxpolicycenter.org/statistics/state-cigarette-tax-rates).
Other policy change	1-1[24]	OR	There are two major types of tobacco policy changes. First, state-level smoking bans. Effective from January 1, 2014, smoking is banned in vehicles with passengers under 18 inside in OR (https://legalbeagle.com/7392711-list-banned-smoking-cars-children.html). Second, change in tobacco minimum purchase age. There is no related changes took place in the post-RCL period.
Number of control states in the donor pool			24

[23] -1 means getting rid of the duplicated calculation of MN.
[24] -1 means getting rid of the duplicated calculation of OR.

Table 8:
Selection of Control States to Form Donor Pool (Continued)

	Number of States	List of States	Notes
Number of control states	50-1		Colorado is the treated state.
Lack of data	2	AK, HI	
Similar intervention	1	WA	We remove WA which legalized recreational use of marijuana in December 2012, despite that it allowed retail sale of recreational marijuana from July 2014.
Spillover effects regarding tourism	16	AZ, UT, WY, NE, KS, OK, TX, NM, NV, ID, MT, ND, SD, MN, IA, MO	We strive to eliminate spillover effects associated with people visiting Colorado to purchase marijuana by removing states that are within 1,000 miles away from Colorado based on estimated distance shown in Google Map.
Panel C: Candy			
State exercise tax rate change	3-1[25]	KS, WV, ME	During post-RCL period, KS and WV decreased state candy exercise tax rate, whereas ME increased state candy exercise tax rate. We obtain information about tax rate change from Bridging the Gap (http://www.bridgingthegapresearch.org/research/sodasnack_taxes/).
Other policy change	3	CA, DE, NY	As of Jan 1, 2014, CA amended existing law to include a requirement for including a nutritious snack in the high school after-school programs and a physical activity element, DE amended prior law to permit a wellness plan to be offered through health insurance if it meets certain requirements, NY encouraged the production and consumption of fresh locally produced fruits and vegetables by elementary and secondary school aged children to help combat the increasing incidence of childhood obesity (https://nccd.cdc.gov/dnpao_dtm/rdPage.aspx?rdReport=DNPAO_DTM.ExploreByTopic&islClass=OWS&islTopic=&go=GO).
Number of control states in the donor pool			25

[25] -1 means getting rid of the duplicated calculation of KS.

Table 8:
Selection of Control States to Form Donor Pool (Continued)

	Number of States	List of States	Notes
Number of control states	50-1		Colorado is the treated state.
Lack of data	2	AK, HI	
Similar intervention	1	WA	We remove WA which legalized recreational use of marijuana in December 2012, despite that it allowed retail sale of recreational marijuana from July 2014.
Spillover effects regarding tourism	16	AZ, UT, WY, NE, KS, OK, TX, NM, NV, ID, MT, ND, SD, MN, IA, MO	We strive to eliminate spillover effects associated with people visiting Colorado to purchase marijuana by removing states that are within 1,000 miles away from Colorado based on estimated distance shown in Google Map.
Panel D: Salty Snacks			
State exercise tax rate change	3-1[26]	KS, TN, WV	During post-RCL period, KS, TN and WV decreased state chips/ pretzels exercise tax rate. We obtain information about tax rate change from Tax Policy Center (http://www.bridgingthegapresearch.org/research/sodasnack_taxes/).
Other policy change	3	CA, DE, NY	As of Jan 1, 2014, CA amended existing law to include a requirement for including a nutritious snack in the high school after-school programs and a physical activity element, DE amended prior law to permit a wellness plan to be offered through health insurance if it meets certain requirements, NY encouraged the production and consumption of fresh locally produced fruits and vegetables by elementary and secondary school aged children to help combat the increasing incidence of childhood obesity (https://nccd.cdc.gov/dnpao_dtm/rdPage.aspx?rdReport=DNPAO_DTM.ExploreByTopic&islClass=OWS&islTopic=&go=GO).
Number of control states in the donor pool			25

[26] -1 means getting rid of the duplicated calculation of KS.

3.4 SCM Results

3.4.1 The Effect of RCL on Sales

Table 9 depicts the pre-RCL characteristics by comparing the predictor means of Colorado, Synthetic Colorado and the average of control states across the categories. As demonstrated by Table 9, Synthetic Colorado is more similar to Colorado than the average of control states, indicating that it plays a better role in matching the treated state.

Table 9:
Pre-Intervention Predictor Means

	Time Period	Colorado	Synthetic Colorado	Average of Control States
Panel A (1): Alcohol Dollar Sales				
Per capita personal income (log)	Quarterly,2013	10.76	10.74	10.68
Unemployment rates	Quarterly,2013	6.74	7.11	7.24
Percentage of population above 20	Annual,2013	.74	.74	.75
Gini index	Annual,2013	.46	.46	.47
Violent crime rates	Annual,2013	.00	.00	.00
Property crime rates	Annual,2013	.03	.03	.03
Unit price (log)	Weekly, 2013	2.25	2.30	2.46
Ad spend (log)	Weekly, 2013	10.59	10.82	10.79
lagged per capita dollar sales	Weekly, 2013	.02	.02	.16
Panel A (2): Alcohol Unit Sales				
Per capita personal income (log)	Quarterly,2013	10.76	10.81	10.68
Unemployment rates	Quarterly,2013	6.74	6.82	7.24
Percentage of population above 20	Annual,2013	0.74	0.75	0.75
Gini index	Annual,2013	0.46	0.46	0.47
Violent crime rates	Annual,2013	0.00	0.00	0.00
Property crime rates	Annual,2013	0.03	0.03	0.03
Unit price (log)	Weekly, 2013	2.25	2.32	2.46
Ad spend (log)	Weekly, 2013	10.59	11.06	10.79
lagged per capita unit sales	Weekly, 2013	0.00	0.00	0.02
Panel B (1): Tobacco Dollar Sales				
Per capita personal income (log)	Quarterly,2013	10.76	10.51	10.65
Unemployment rates	Quarterly,2013	6.74	7.78	7.46
Percentage of population above 20	Annual,2013	0.74	0.74	0.75
Gini index	Annual,2013	0.46	0.47	0.47
Violent crime rates	Annual,2013	0.00	0.00	0.00
Property crime rates	Annual,2013	0.03	0.03	0.03
Unit price (log)	Weekly, 2013	2.83	2.67	2.55
Ad spend (log)	Weekly, 2013	0.29	3.22	2.68
lagged per capita dollar sales	Weekly, 2013	0.06	0.06	0.02
Panel B (2): Tobacco Unit Sales				
Per capita personal income (log)	Quarterly,2013	10.76	10.59	10.65
Unemployment rates	Quarterly,2013	6.74	7.19	7.46
Percentage of population above 20	Annual,2013	0.74	0.75	0.75
Gini index	Annual,2013	0.46	0.46	0.47
Violent crime rates	Annual,2013	0.00	0.00	0.00
Property crime rates	Annual,2013	0.03	0.02	0.03
Unit price (log)	Weekly, 2013	2.83	2.70	2.56
Ad spend (log)	Weekly, 2013	0.29	2.54	2.68
lagged per capita unit sales	Weekly, 2013	0.01	0.01	0.00

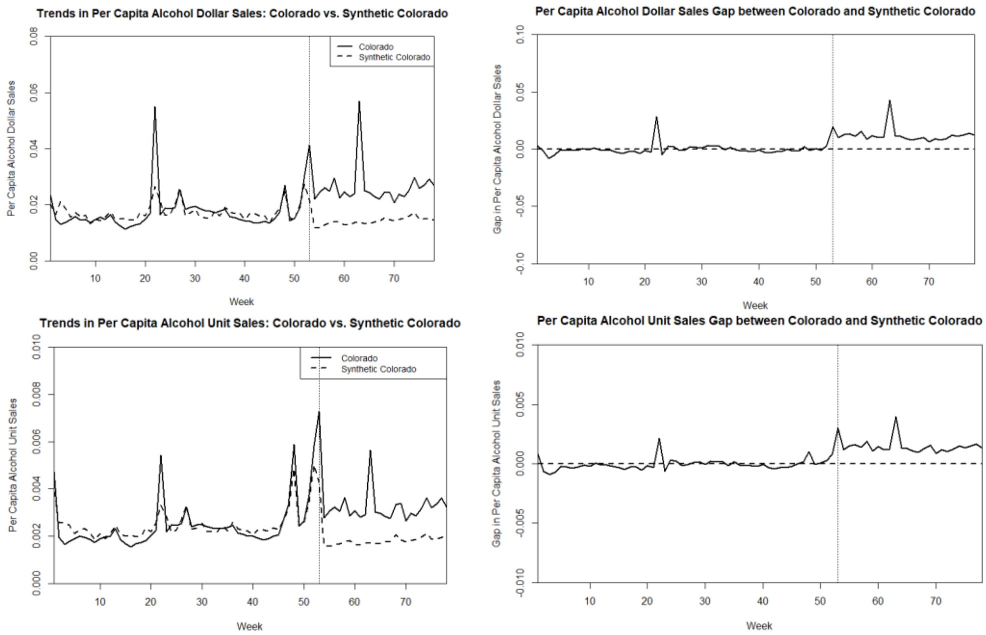
Table 9:
Pre-Intervention Predictor Means (Continued)

	Time Period	Colorado	Synthetic Colorado	Average of Control States
Panel C (1): Candy Dollar Sales				
Per capita personal income (log)	Quarterly,2013	10.76	10.77	10.67
Unemployment rates	Quarterly,2013	6.74	6.99	7.36
Percentage of population above 20	Annual,2013	0.74	0.75	0.75
Gini index	Annual,2013	0.46	0.46	0.47
Violent crime rates	Annual,2013	0.00	0.00	0.00
Property crime rates	Annual,2013	0.03	0.03	0.03
Unit price (log)	Weekly, 2013	1.29	1.29	1.25
Ad spend (log)	Weekly, 2013	8.68	8.93	9.02
lagged per capita dollar sales	Weekly, 2013	0.09	0.09	0.06
Panel C (2): Candy Unit Sales				
Per capita personal income (log)	Quarterly,2013	10.76	10.76	10.67
Unemployment rates	Quarterly,2013	6.74	7.05	7.36
Percentage of population above 20	Annual,2013	0.74	0.75	0.75
Gini index	Annual,2013	0.46	0.46	0.47
Violent crime rates	Annual,2013	0.00	0.00	0.00
Property crime rates	Annual,2013	0.03	0.03	0.03
Unit price (log)	Weekly, 2013	1.29	1.29	1.25
Ad spend (log)	Weekly, 2013	8.68	8.89	9.02
lagged per capita unit sales	Weekly, 2013	0.05	0.05	0.03
Panel D (1): Salty Snacks Dollar Sales				
Per capita personal income (log)	Quarterly,2013	10.76	10.76	10.67
Unemployment rates	Quarterly,2013	6.74	6.77	7.32
Percentage of population above 20	Annual,2013	0.74	0.75	0.75
Gini index	Annual,2013	0.46	0.46	0.47
Violent crime rates	Annual,2013	0.00	0.00	0.00
Property crime rates	Annual,2013	0.03	0.03	0.03
Unit price (log)	Weekly, 2013	1.46	1.45	1.42
Ad spend (log)	Weekly, 2013	9.15	9.68	9.68
lagged per capita dollar sales	Weekly, 2013	0.17	0.17	0.12
Panel D (2): Salty Snacks Unit Sales				
Per capita personal income (log)	Quarterly,2013	10.76	10.76	10.67
Unemployment rates	Quarterly,2013	6.74	6.77	7.32
Percentage of population above 20	Annual,2013	0.74	0.75	0.75
Gini index	Annual,2013	0.46	0.46	0.47
Violent crime rates	Annual,2013	0.00	0.00	0.00
Property crime rates	Annual,2013	0.03	0.03	0.03
Unit price (log)	Weekly, 2013	1.46	1.46	1.42
Ad spend (log)	Weekly, 2013	9.15	9.66	9.68
lagged per capita unit sales	Weekly, 2013	0.06	0.06	0.04

Figure 5 depicts the trajectory of per capita dollar sales and per capita unit sales in Colorado (solid) and synthetic Colorado (dashed) during both the pre-RCL and post-RCL periods for alcohol, tobacco, candy, and salty snacks categories. The trend patterns of Colorado and synthetic Colorado are very similar during the pretreatment period in all the categories, verifying the first assumption mentioned previously. More importantly, we observe a divergence of various levels between Colorado and synthetic Colorado over the post-RCL period across all the categories. According to Figure 5(a), the average gap in per capita alcohol dollar sales between Colorado and synthetic Colorado during the post-RCL period is .012, indicating approximately 28.92% (.0120/.0415) of the first week of 2014 baseline level. Per capita alcohol unit sales was enhanced by .001, which amounts to an increase of around 19.18% (.0014/.0073). As plotted in Figure 5(b), the magnitudes of the impact of RCL on per capita tobacco dollar sales and per capita tobacco unit sales are .003 and .001, respectively, representing an approximate increase of 3.61% (.0031/.0858) and 8.26% (.0009/.0109), respectively. Figure 5(c) illustrates that during the post-RCL period, per capita candy dollar sales went up by around 25.73% (.0184/.0715). The mean gap in per capita candy unit sales is .01, indicating an increase of 27.71% (.0115/.0415). The plots in Figure 5(d) indicate that the average gaps in per capita salty snacks dollar sales and per capita salty snacks unit sales between Colorado and synthetic Colorado after RCL are .038 and .015, respectively. Therefore, per capita salty snacks dollar sales and per capita salty snacks unit sales were enhanced by 19.09% (.0379/.1985) and 20.96% (.0149/.0711), respectively. The SCM results of per capita dollar sales and per capita unit sales are shown in Table 10.

Figure 5:
Trends and Gaps in Per Capita Dollar Sales and Per Capita Unit Sales between Colorado and Synthetic Colorado

(a) Alcohol



(b) Tobacco

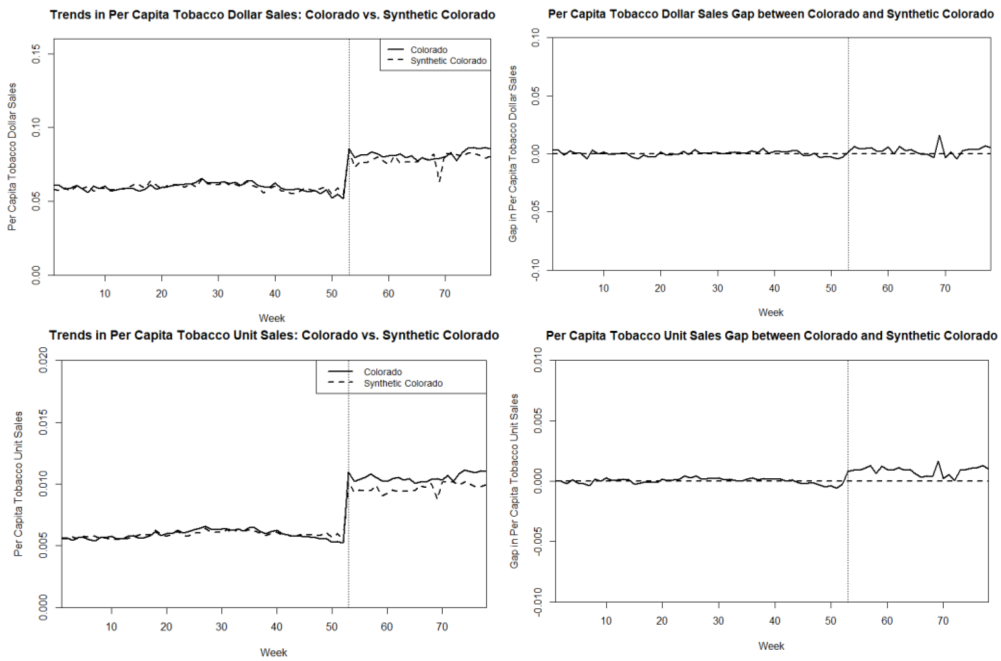
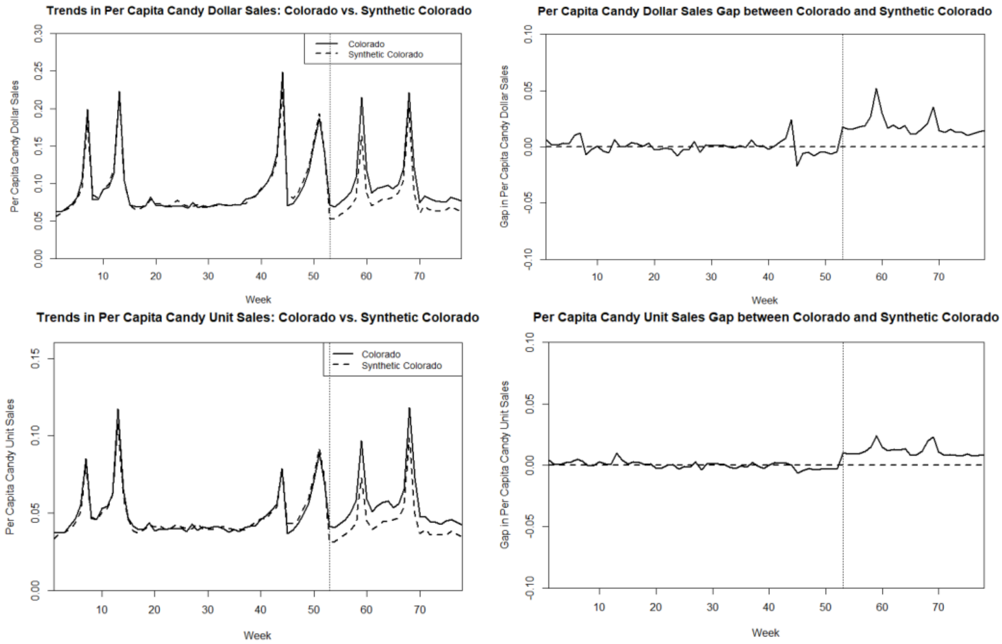


Figure 5:
Trends and Gaps in Per Capita Dollar Sales and Per Capita Unit Sales between Colorado and Synthetic Colorado (Continued)

(c) Candy



(d) Salty Snacks

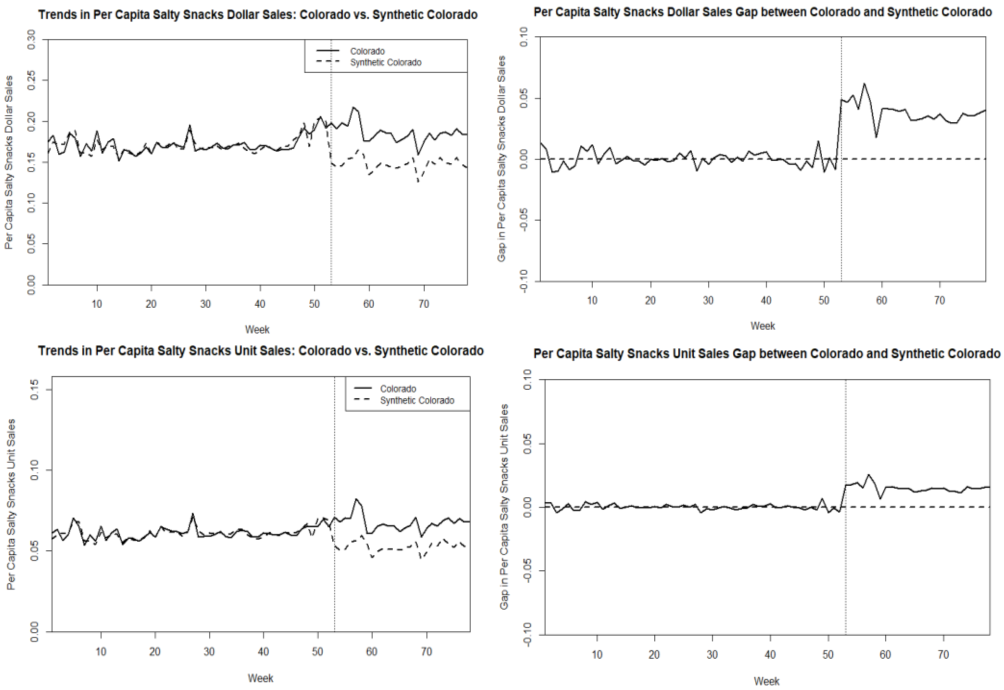


Table 10:
SCM Results of Per Capita Dollar Sales and Per Capita Unit Sales

	Treatment Effect	Estimated Change in Percentage	Post-/Pre-treatment MSPE (p value)
Panel A: Alcohol			
Per capita dollar sales	.012	28.92%	9.12 (.04)
Per capita unit sales	.001	19.18%	11.66 (.04)
Panel B: Tobacco			
Per capita dollar sales	.003	3.61%	5.41 (.64)
Per capita unit sales	.001	8.26%	18.11 (.44)
Panel C: Candy			
Per capita dollar sales	.018	25.73%	11.52 (.04)
Per capita unit sales	.012	27.71%	19.37 (.04)
Panel D: Salty Snacks			
Per capita dollar sales	.038	19.09%	41.13 (.04)
Per capita unit sales	.015	20.96%	45.81 (.04)

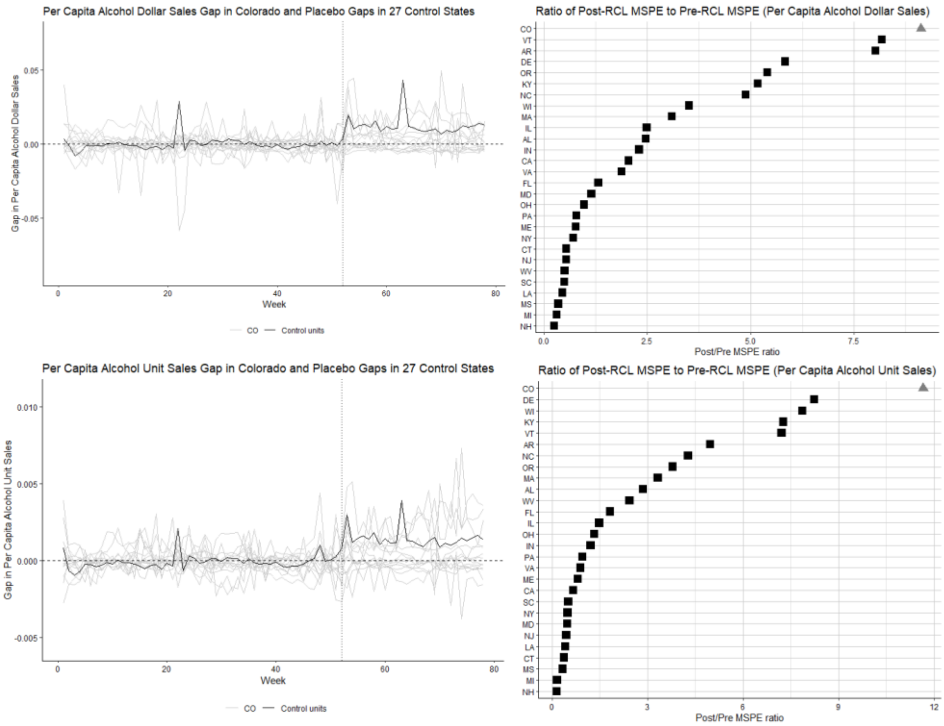
3.4.2 Placebo Tests

To ensure that our results are not driven by chance, we evaluate the significance of the results by conducting two placebo tests (Abadie et al. 2010, 2015). First, for each category, we conduct "in-space placebo" test by applying the approach of estimating the effect of RCL in Colorado to each of the control states and put Colorado back in the donor pool. As Figure 6(a) shows, the gaps in per capita alcohol dollar sales and per capita alcohol unit sales estimated for Colorado are larger than any of the 27 control states, indicating that our results are not obtained by chance. According to Table 10, post-/pre-treatment MSPE ratios of per capita alcohol dollar sales and per capita alcohol unit sales are 9.12 and 11.66, respectively. Therefore, if the intervention is randomly assigned to any of the 27 control states, the probability of obtaining a post-/pre-treatment MSPE ratio as large as Colorado's is 1/28 = .04. Second, we run "in-time placebo" study suggested by Abadie et al. (2015) through reassigning the RCL to the middle of the pretreatment period (i.e., week 27) which is half year earlier than the actual RCL. As shown in Figure 7(a), the synthetic Colorado resembles Colorado in per capita alcohol dollar sales and per capita alcohol unit sales between January and June of 2013 with little discrepancy in per capita alcohol dollar sales and per capita alcohol unit sales during June 2013 and Dec 2013, indicating no perceivable effects of the placebo RCL.

We run the same analyses for the remaining categories. As shown by Figure 6 (b), the gaps in per capita tobacco dollar sales and per capita tobacco unit sales estimated for Colorado are not larger than those of most of the 24 control states, demonstrating that RCL does not have significant impacts on per capita tobacco dollar sales (p value = .64) and per capita tobacco unit sales (p value = .44). Figure 6(c) and Figure 6(d) present that Colorado has the largest gaps in per capita dollar sales and per capita unit sales for candy and salty snacks among the 25 control states. Therefore, RCL has positive effects on per capita candy dollar sales (post-/pre-treatment MSPE ratio = 11.52, p value = .04), per capita candy unit sales (post-/pre-treatment MSPE ratio = 19.37, p value = .04), per capita salty snacks dollar sales (post-/pre-treatment MSPE ratio = 41.13, p value = .04) and per capita salty snacks unit sales (post-/pre-treatment MSPE ratio = 45.81, p value = .04), according to Table 10. Figure 7(b), (c), and (d) illustrate that placebo RCL has no perceivable effects during June 2013 and Dec 2013.

Figure 6:
"In-Space" Placebo Tests

(a) Alcohol



(b) Tobacco

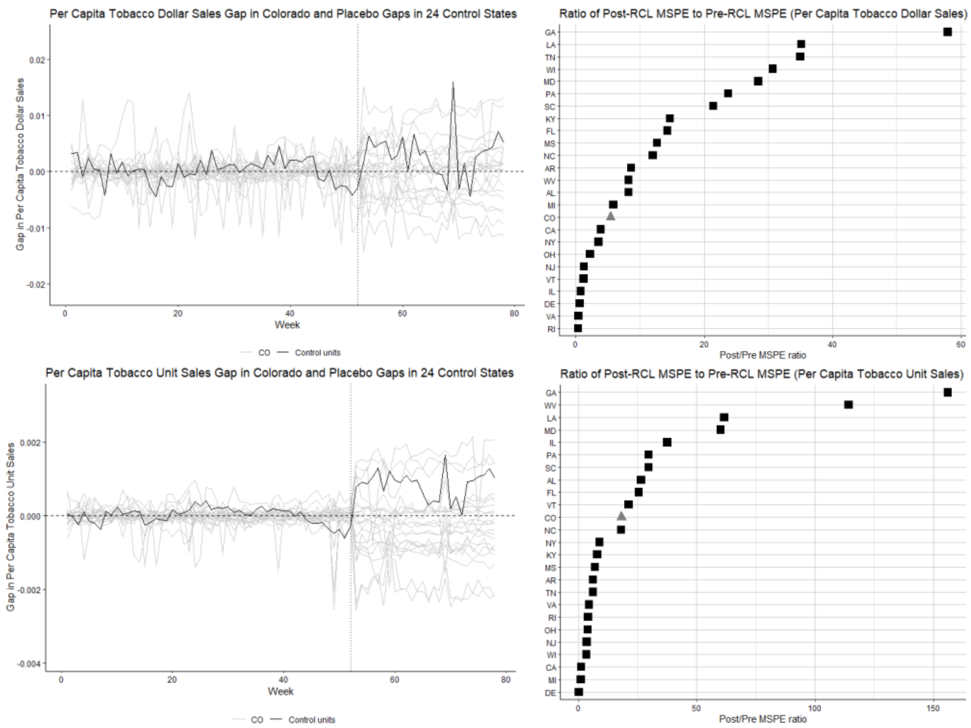
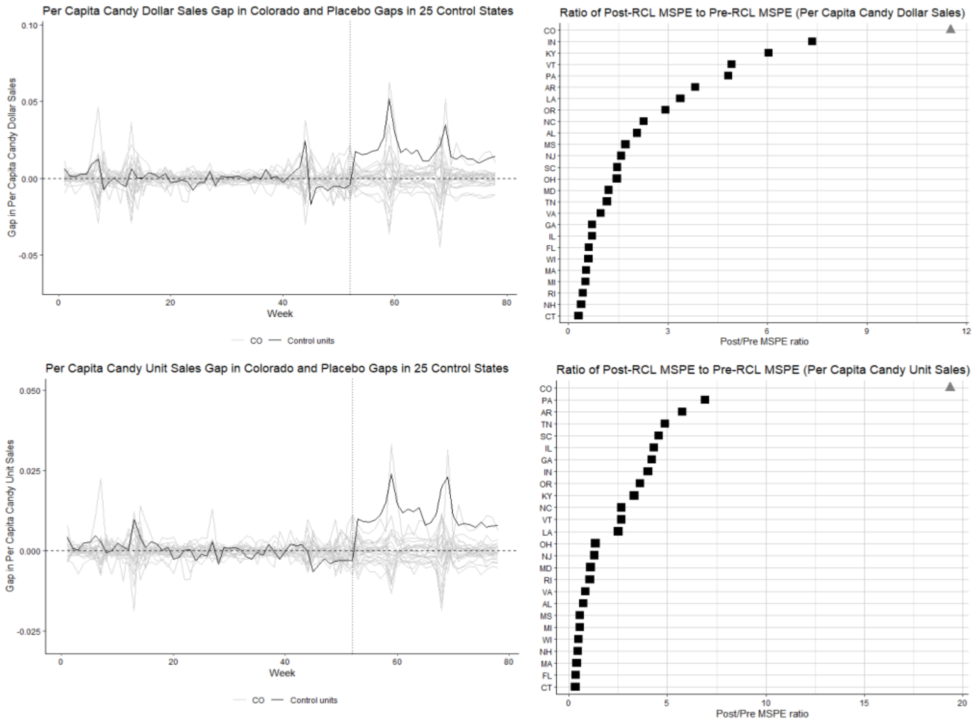


Figure 6:
"In-Space" Placebo Tests (Continued)

(c) Candy



(d) Salty Snacks

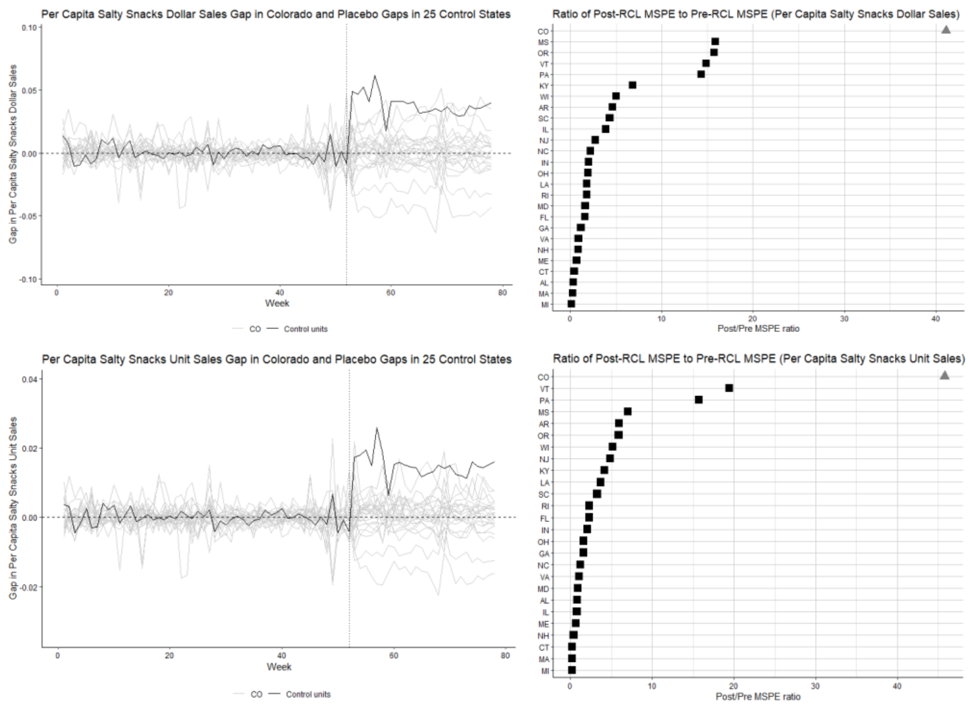
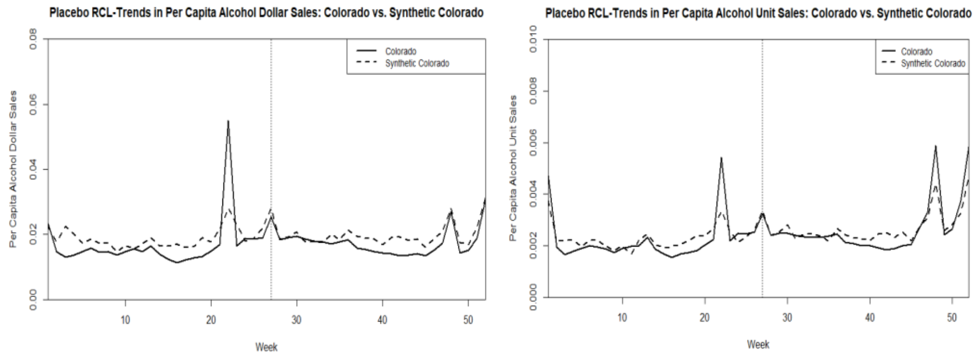
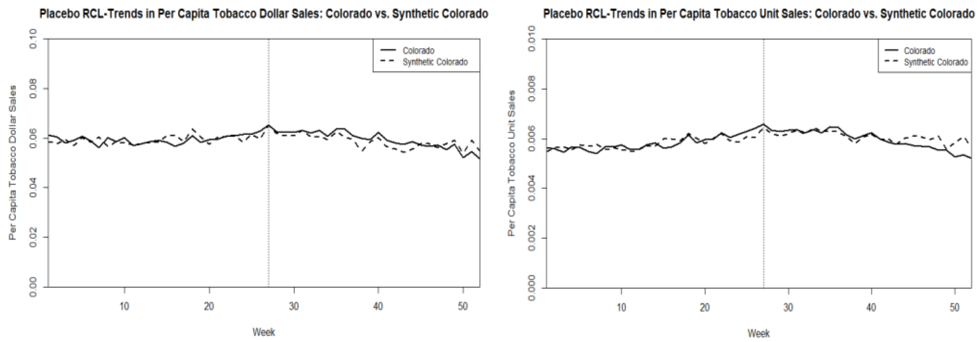


Figure 7:
"In-Time" Placebo Tests

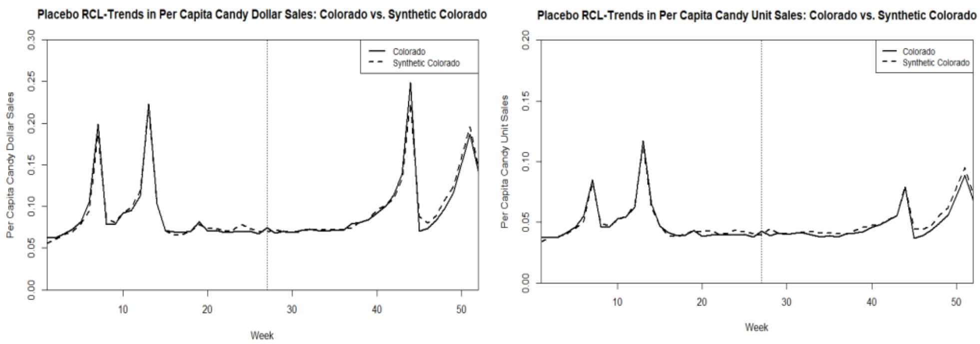
(a) Alcohol



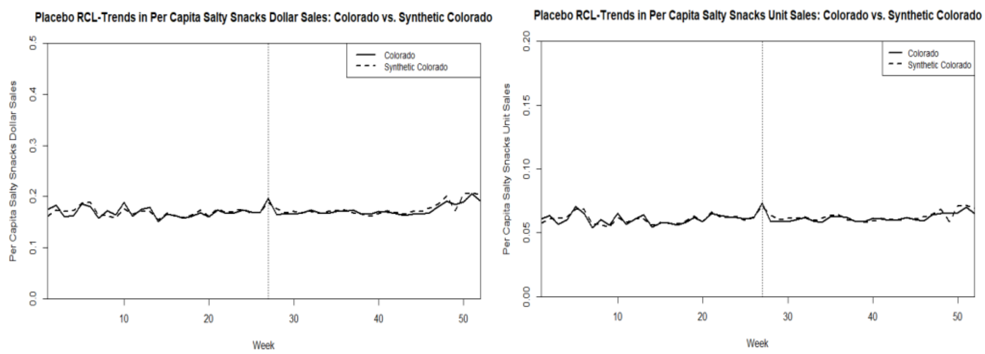
(b) Tobacco



(c) Candy



(d) Salty Snacks



3.5 Sensitivity Analysis

To demonstrate the robustness of our results, we conduct two different sensitivity analysis. The first one involves investigating alternative measures for the outcome variables and the second sensitivity analysis involves including additional predictor variables.

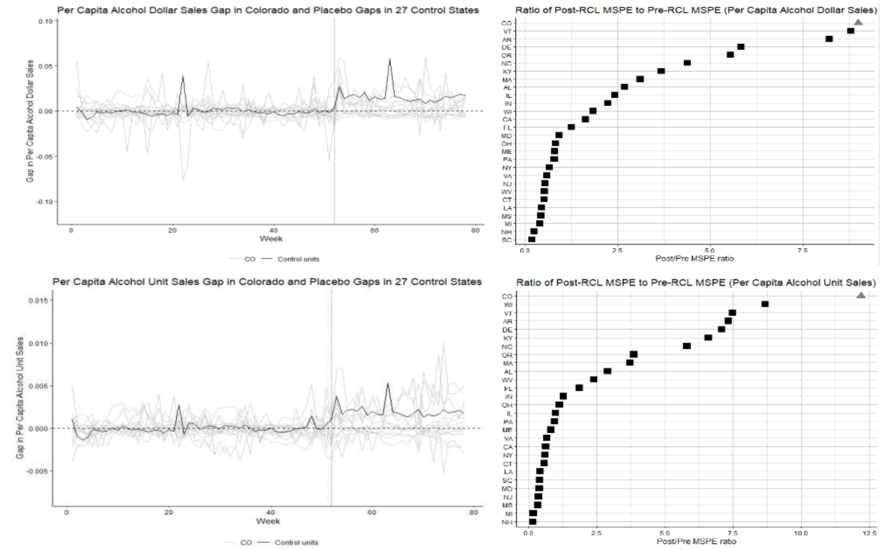
3.5.1 Alternative Measure of Outcome Variables

We use the ratio of dollar sales to the population above 20 years old[27] in a state, during a week, as the alternative measure of per capita dollar sales and the ratio of unit sales to the population above 20 years old in a state, during a week, as the alternative measure of per capita unit sales in all the categories. The results are shown in Figure 8. According to Figure 8(a), RCL leads to an increase in per capita alcohol dollar sales and per capita alcohol unit dollar sales by .016, .002, respectively. The results indicate an increase of 28.70% (p value = .04) and 20.41% (p value = .04), respectively. Figure 8(b) shows that RCL does not have significant impacts on per capita tobacco dollar sales (p value = .60) and per capita tobacco unit sales (p value = .48). As Figure 8(c) shows, RCL increases per capita candy dollar sales by 25.28% (p value = .04) and results in an increase in per capita candy unit dollar sales by 28.39% (p value = .04). Figure 8(d) indicates that per capita salty snacks dollar sales went up by 19.26% (p value = .04) and per capita salty snacks unit dollar sales rose by 20.50% (p value = .04) during post-RCL period. As such, the results using alternative measures indicate the robustness of our main results.

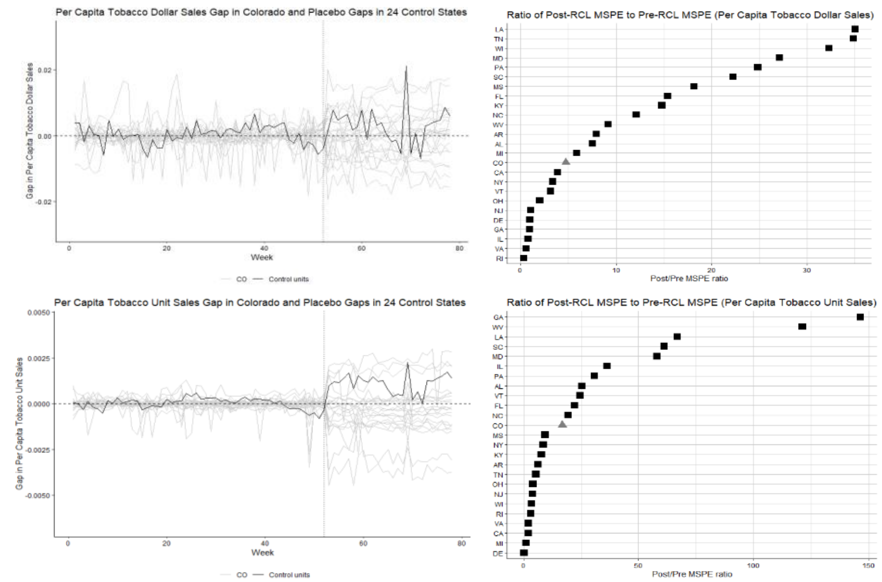
[27] The reason why we focus on population above 20 is that recreational marijuana is allowed to be sold to adults above 21. However, the data regarding the population above 21 is not available.

Figure 8: Sensitivity Analysis: Alternative Measure of Outcome Variables[28]

(a) Alcohol



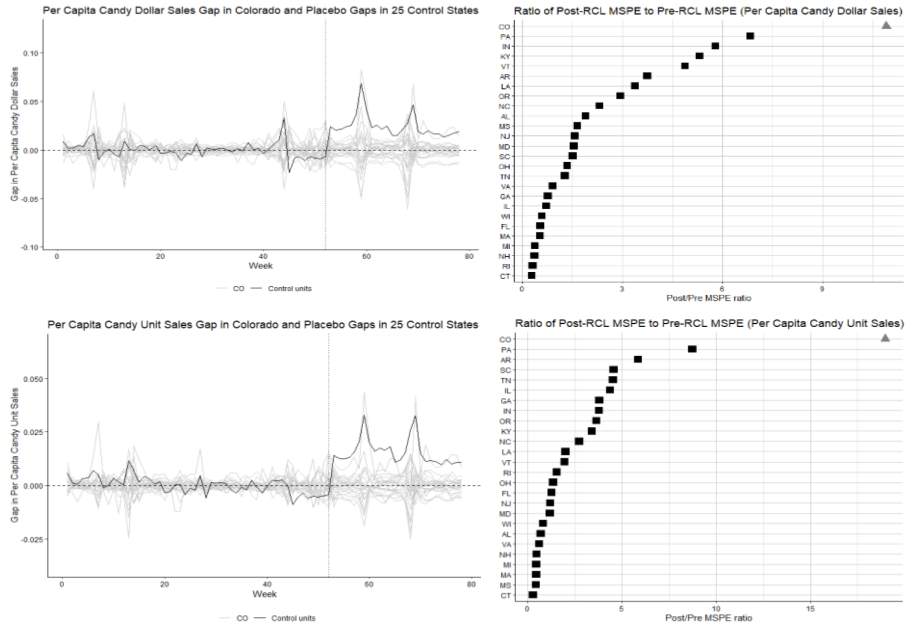
(b) Tobacco



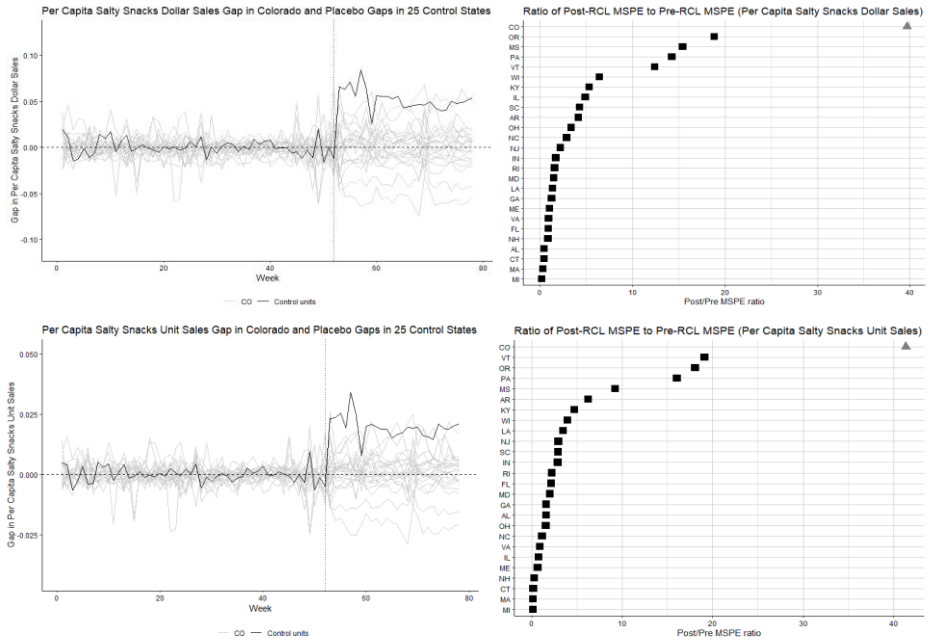
[28] We use the ratio of dollar sales to the population over 20 years old in a state during a week as the alternative measure of per capita dollar sales and the ratio of unit sales to the population over 20 years old in a state during a week as the alternative measure of per capita unit sales in all the categories.

Figure 8:
Sensitivity Analysis: Alternative Measure of Outcome Variables (Continued)[29]

(c) Candy



(d) Salty Snacks



---

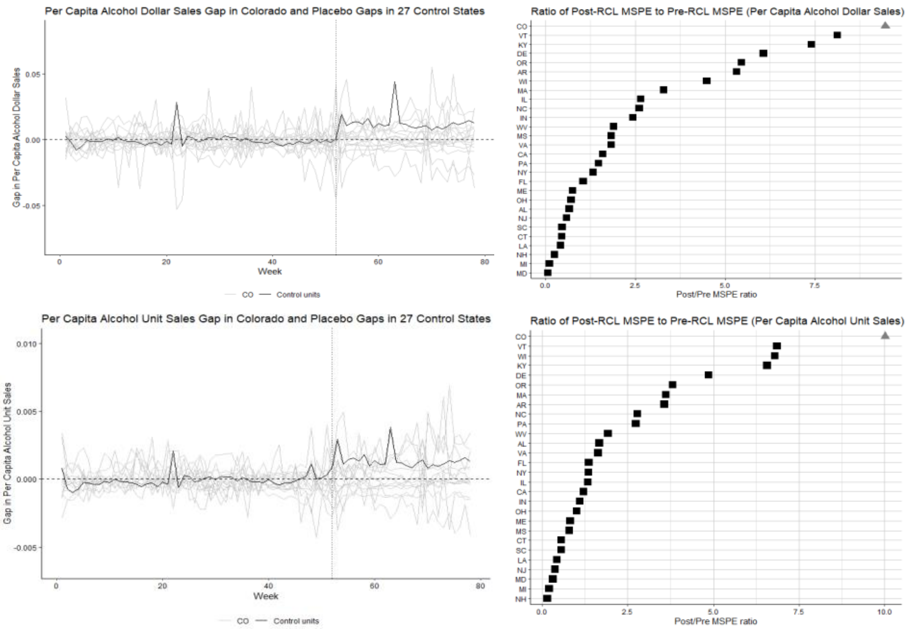
[29] We use the ratio of dollar sales to the population over 20 years old in a state during a week as the alternative measure of per capita dollar sales and the ratio of unit sales to the population over 20 years old in a state during a week as the alternative measure of per capita unit sales in all the categories.

3.5.2 Including Additional Predictor Variables

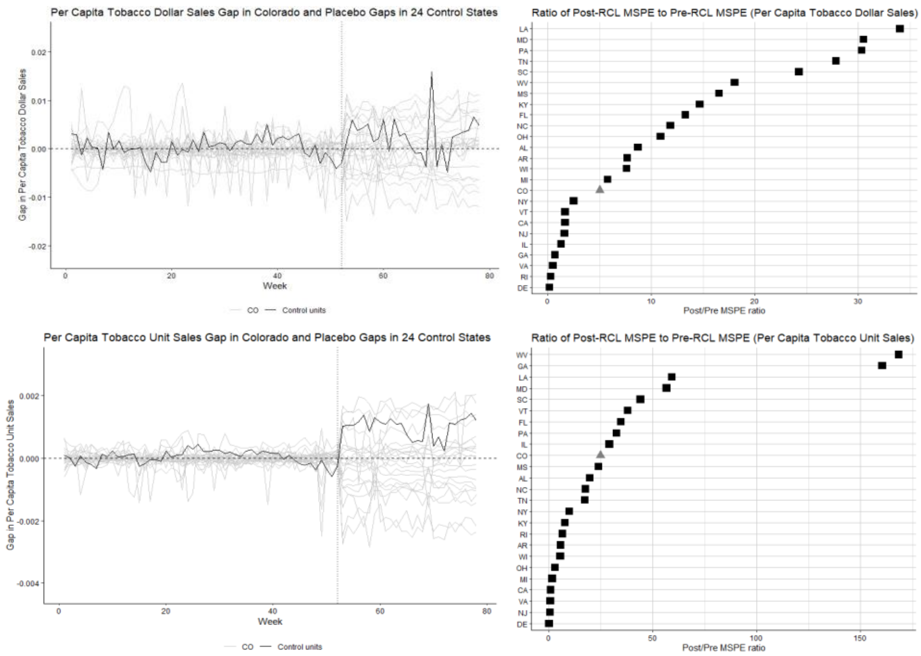
We add three additional predictors to show the robustness of our results. We obtain poverty rates, log-transformed population above 20 and gender ratio from US Bureau of Economic Analysis. The results are summarized in Figure 9. RCL leads to an increase in per capita alcohol dollar sales by 30.12% (p value = .04) and an increase in per capita alcohol unit dollar sales by 17.81% (p value = .04), according to Figure 9(a). Figure 9(b) shows that RCL has no impact on per capita tobacco dollar sales (p value = .64) and per capita tobacco unit sales (p value = .40). As can be seen from Figure 9(c), RCL leads to a 25.87% increase in per capita candy dollar sales (p value = .04) and a 27.47% increase in per capita candy unit dollar sales (p value = .04). Figure 9(d) shows that RCL contributes to a rise in per capita salty snacks dollar sales by .04, indicating an increase of 18.89% (p value = .04). Also, RCL leads to a 19.27% increase in per capita salty snacks unit dollar sales (p value = .04). Therefore, the sensitivity analysis results are consistent with our main results.

Figure 9:
Sensitivity Analysis: Additional Predictor Variables[30]

(a) Alcohol



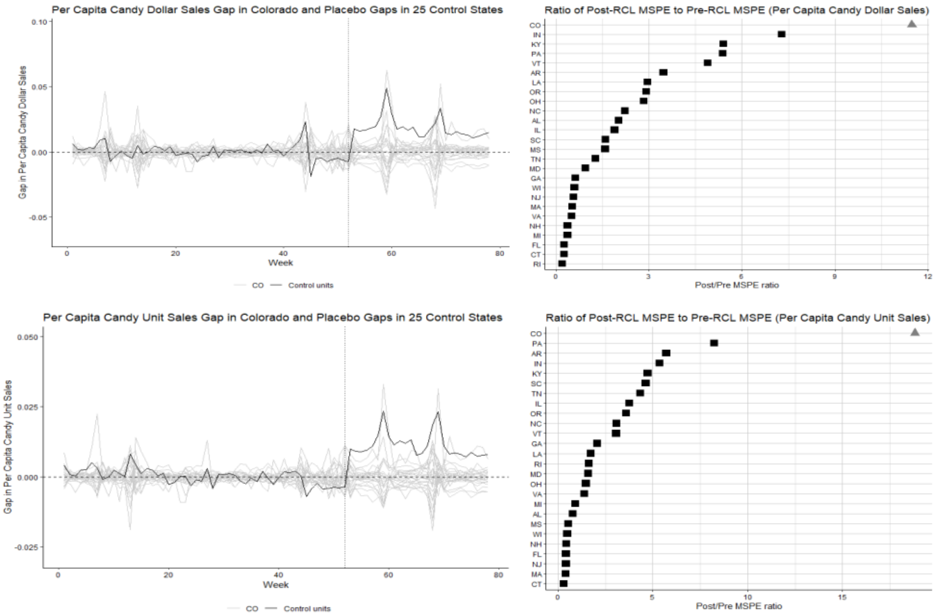
(b) Tobacco



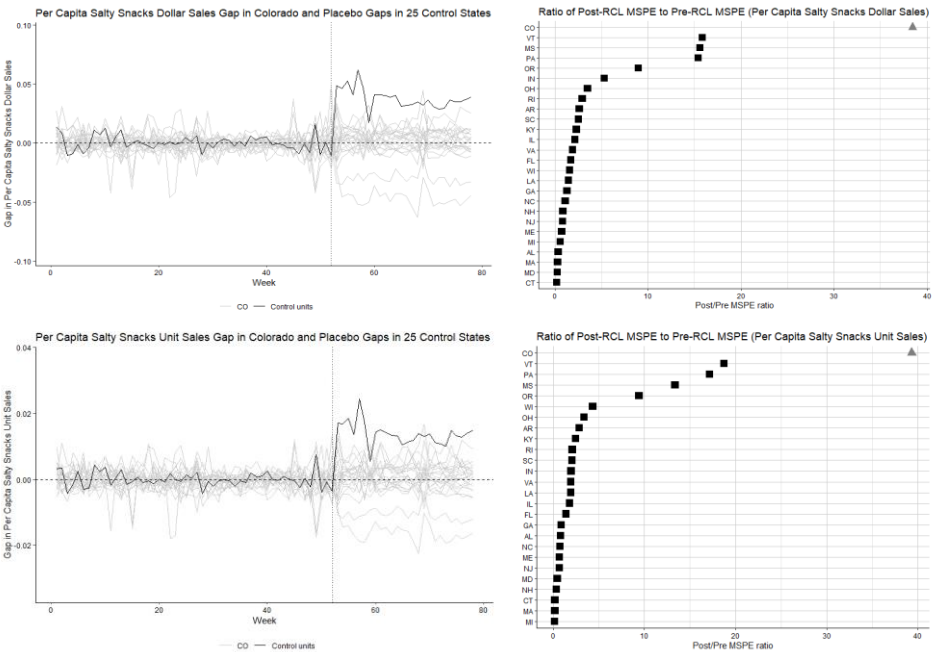
[30] We add three additional predictors to show the robustness of our results. We obtain poverty rates, population above 20 (log) and gender ratio from US Bureau of Economic Analysis.

Figure 9:
Sensitivity Analysis: Additional Predictor Variables (Continued)[31]

(c) Candy



(d) Salty Snacks



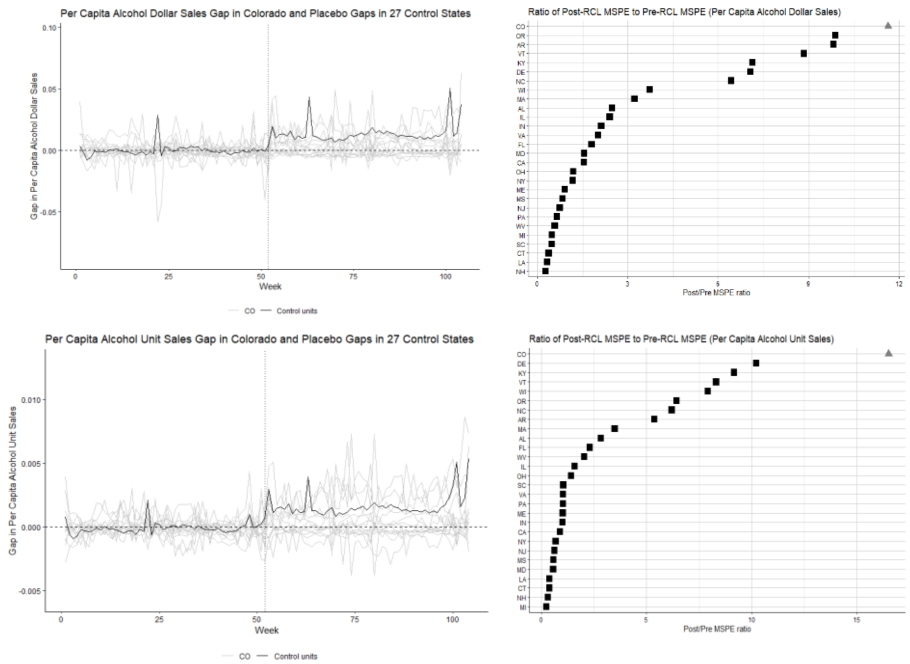
[31] We add three additional predictors to show the robustness of our results. We obtain poverty rates, population above 20 (log) and gender ratio from US Bureau of Economic Analysis.

3.5.3 Alternative Choices of Pre-/Posttreatment Period

We choose alternative pre-/posttreatment periods to conduct additional sensitivity analysis. Specifically, we use one year prior to and post RCL (see Figure 10) as well as half year before and after RCL (see Figure 11) to run the same analysis. As indicated by Figure 10, the positive effects of RCL on per capita dollar sales and per capita unit sales of alcohol are significant for alcohol (p value = .04), candy (p value = .04) and salty snacks (p value = .04), which is not the case for tobacco. The results shown in Figure 11 are similar except for the impact of RCL on per capita candy dollar sales. To sum up, the results of one-year and half-year pre-/post-RCL period are mostly consistent with our main result.

Figure 10. Sensitivity Analysis: One-Year Pre- and Posttreatment Period

(a) Alcohol



(b) Tobacco

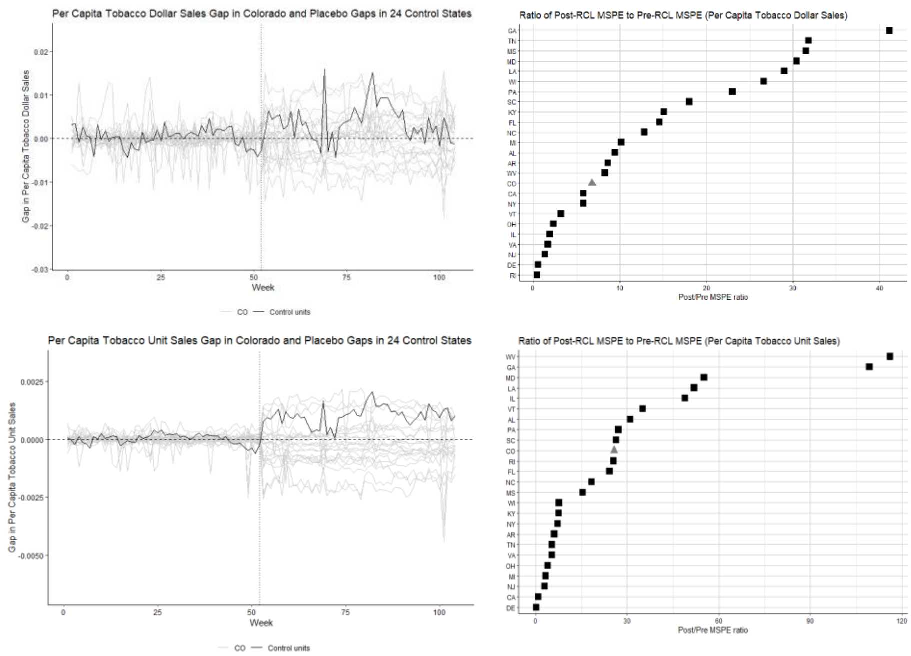
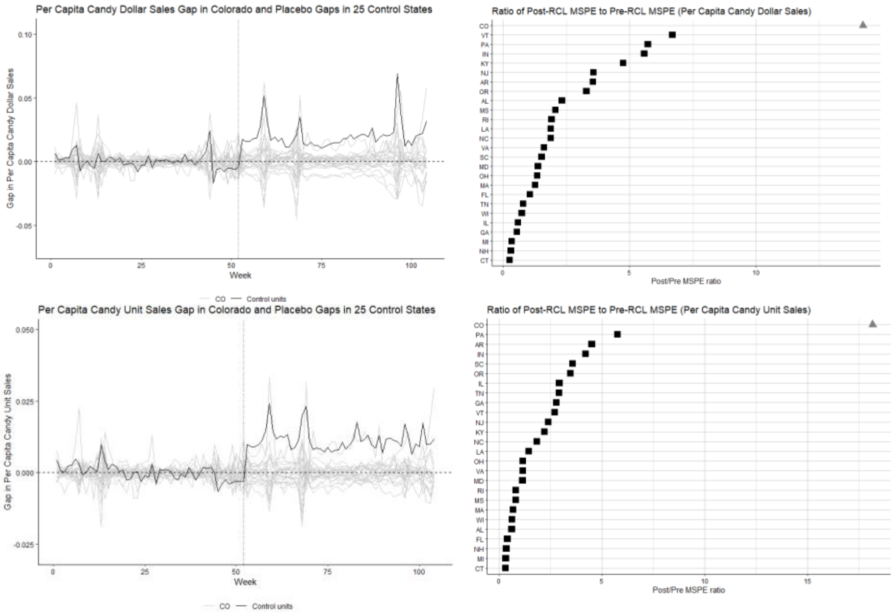


Figure 10. Sensitivity Analysis: One-Year Pre- and Posttreatment Period (Continued)

(c) Candy



(d) Salty Snacks

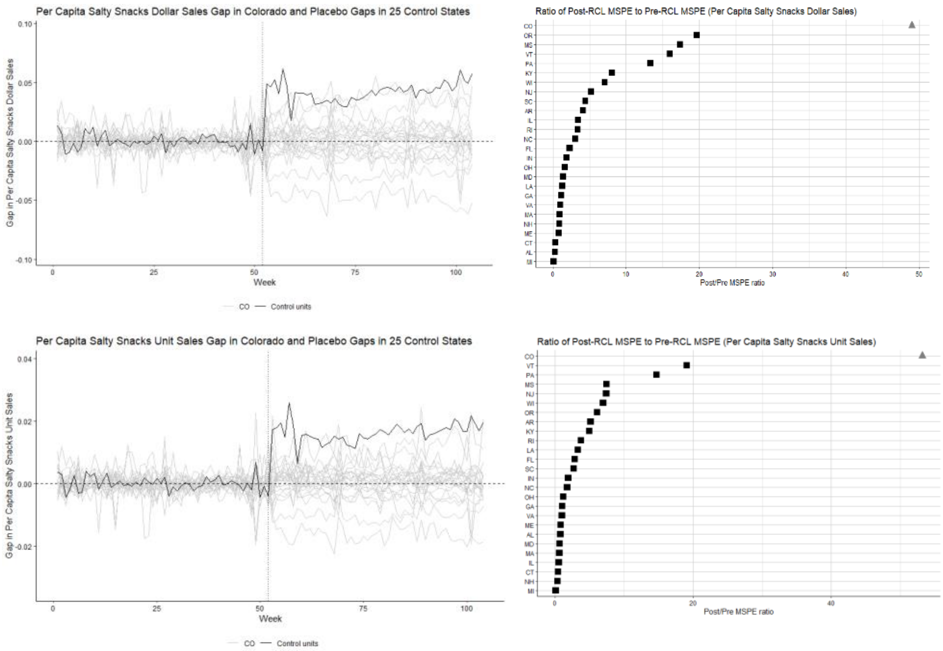
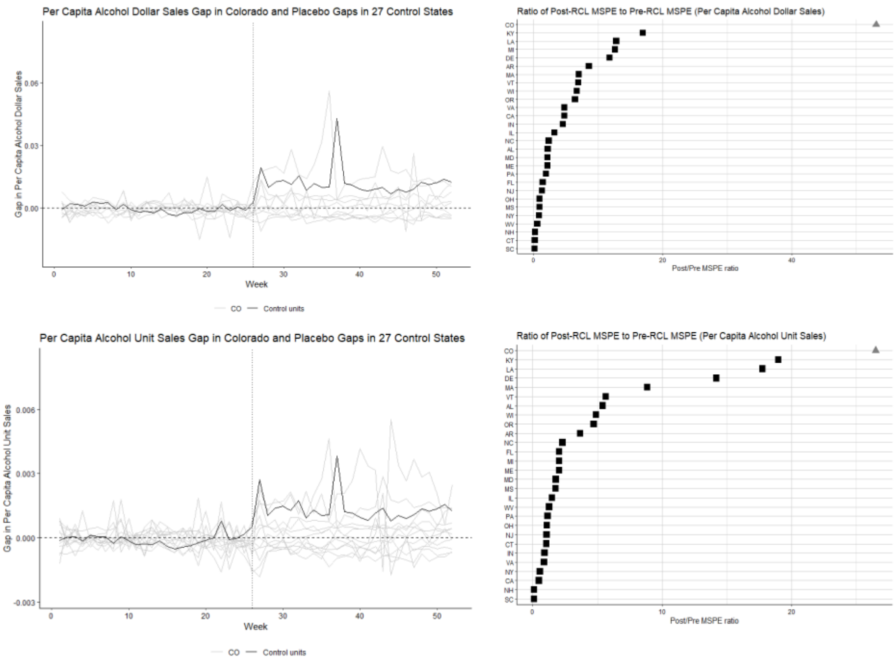


Figure 11. Sensitivity Analysis: Half-Year Pre- and Posttreatment Period

(a) Alcohol



(b) Tobacco

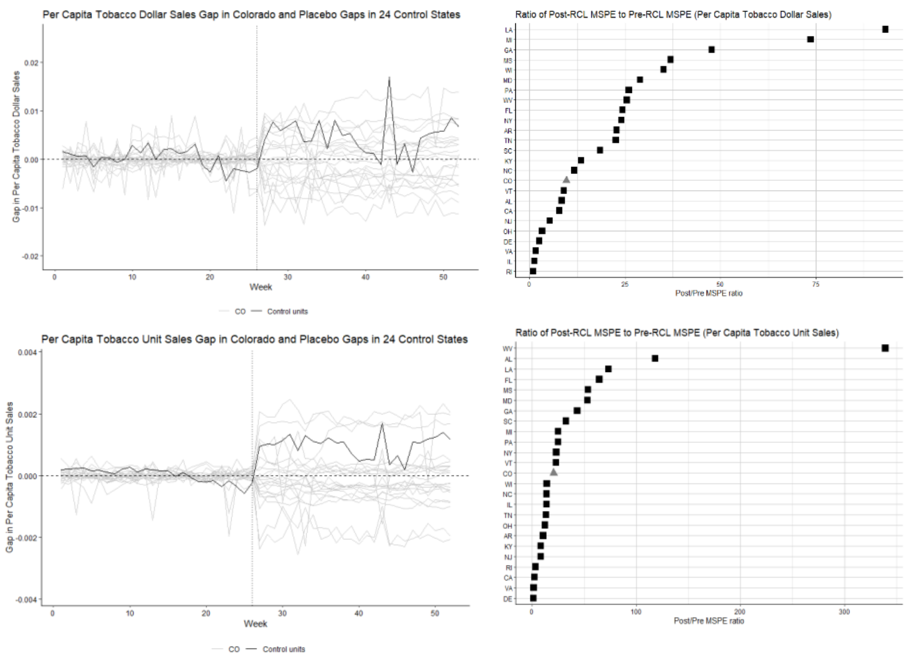
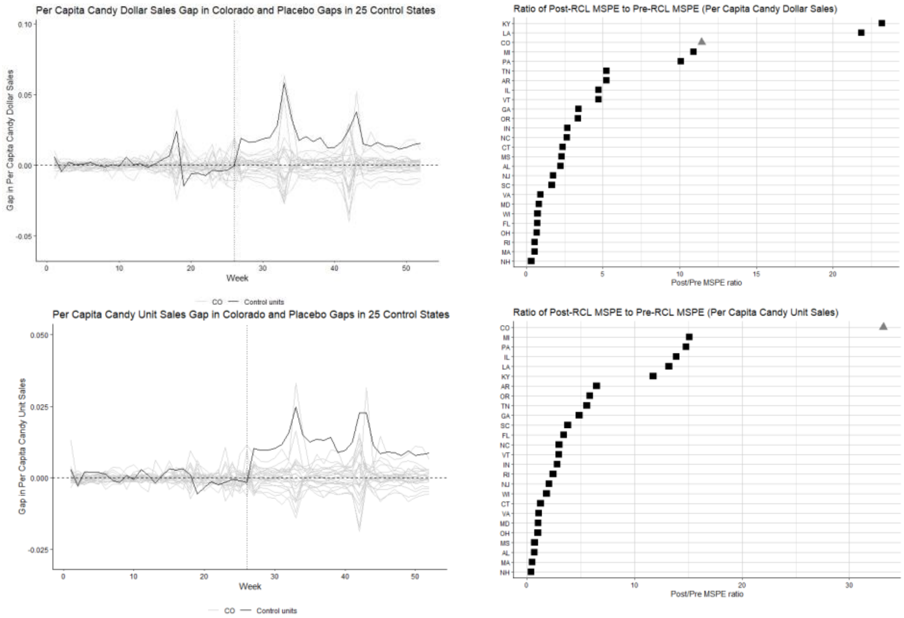
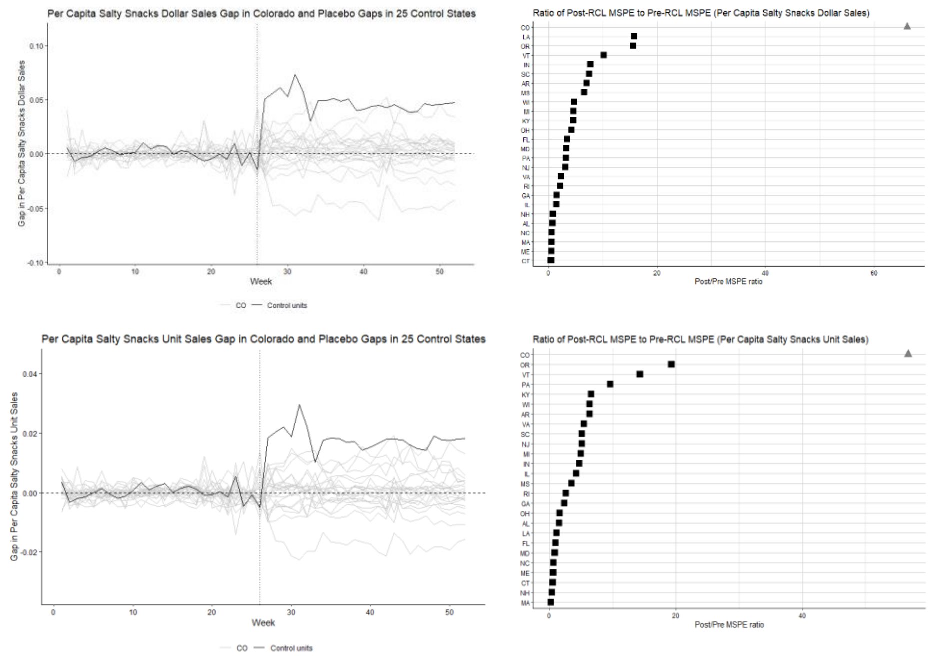


Figure 11. Sensitivity Analysis: Half-Year Pre- and Posttreatment Period (Continued)

(c) Candy



(d) Salty Snacks

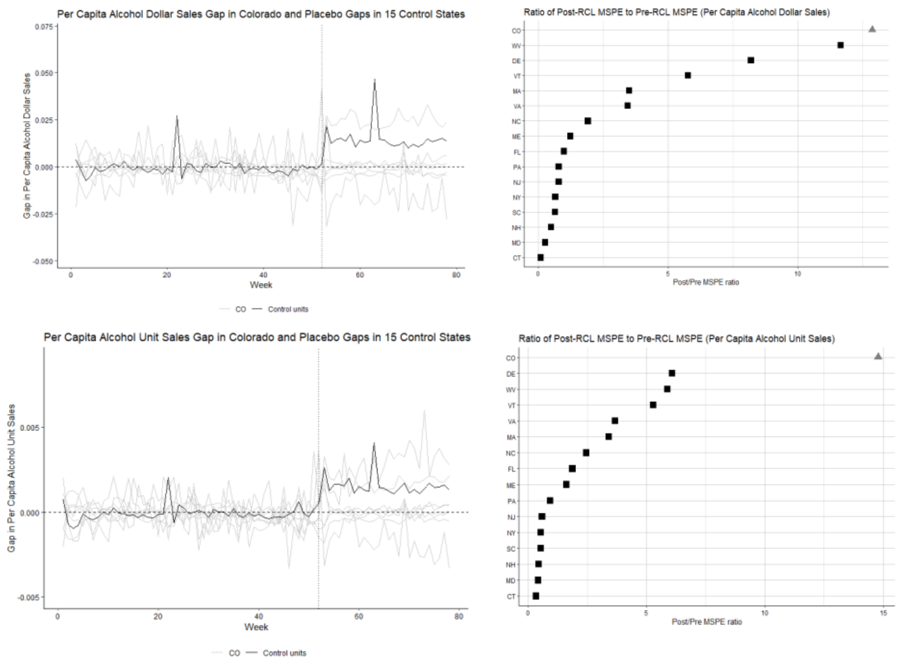


3.5.4 Alternative Choice of Donor Pool

To show the robustness of our results, we change each donor pool by deleting states that are within 1,500 miles away from Colorado. As such, the donor pools of alcohol, tobacco, candy and salty snacks are 15, 13, 13 and 14, respectively. According to Figure 12, RCL resulted in an increase in per capita dollar sales and per capita unit sales of alcohol (p value = .06), candy (p value = .07) and salty snacks (p value = .07). While p value is larger than .05, it is the highest p value we can achieve given the fact that Colorado ranks the top in terms of the Post-/Pretreatment MSPE ratio. However, the effect of RCL on per capita tobacco dollar sales (p value = .21) and per capita unit sales (p value = .14). In conclusion, our results are robust.

Figure 12. Sensitivity Analysis: Alternative Donor Pool

(a) Alcohol



(b) Tobacco

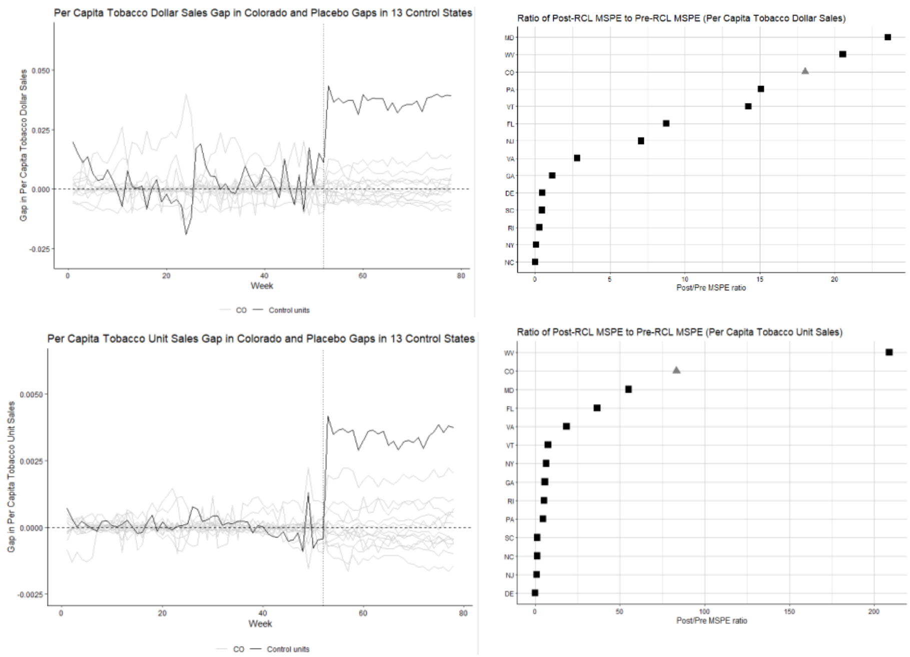
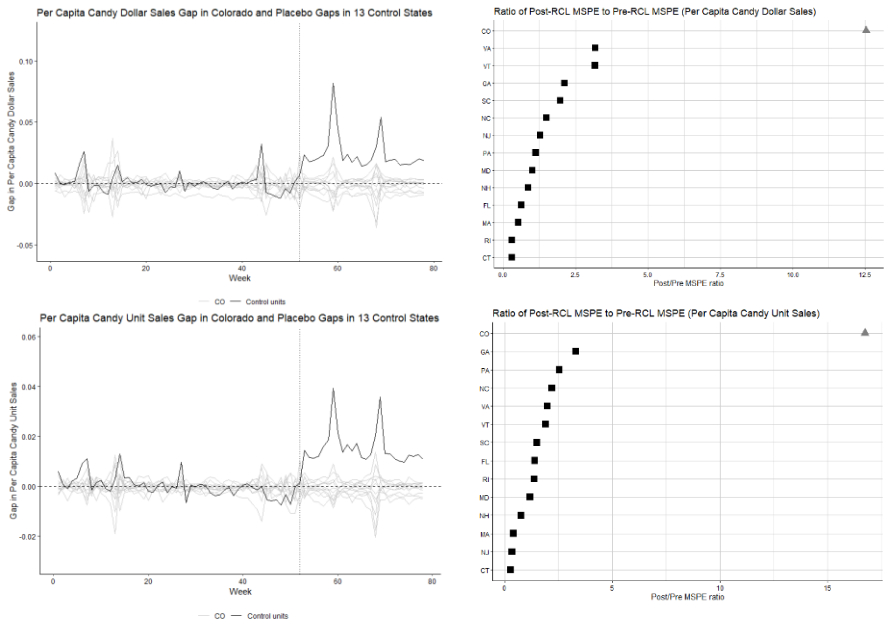
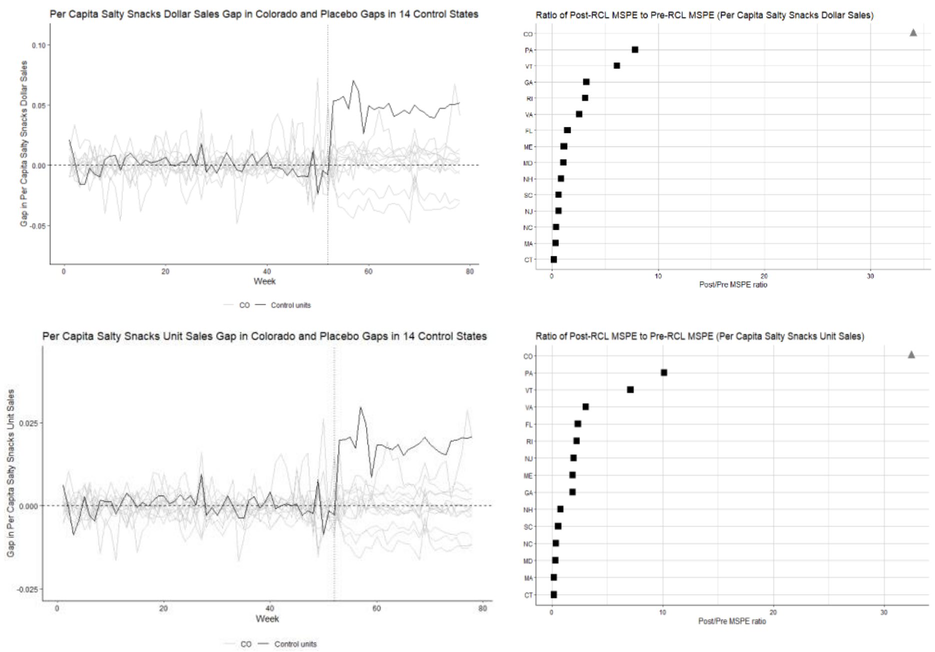


Figure 12. Sensitivity Analysis: Alternative Donor Pool (Continued)

(c) Candy



(d) Salty Snacks



3.5.5 DiD Estimation

Given that DiD model is a common methodology to analyze the effect of an exogenous event, we first conduct DiD analysis to demonstrate the effect of RCL on per capita dollar sales and per capita unit sales in alcohol, tobacco, candy and salty snacks. Specifically, we estimate the following DiD model to investigate the effect of RCL on per capita dollar sales and per capita unit sales across all these categories:

$$Y_{it} = \beta \times Treat_i \times Post_t + \theta_i + \sigma_t + \varepsilon_{it}$$

Where Y_{it} is outcome variable (i.e., per capita dollar sales and per capita unit sales) in state i week t. $Treat_i$ equals 1 if it is the treated state (i.e., Colorado) otherwise is 0. $Post_t$ takes the value of 1 during the posttreatment period and 0 otherwise. Here β is our focus and captures the average treatment effect of RCL. In addition, θ_i represents state fixed effects and σ_t refers to time fixed effects.

The estimated results are summarized in Table 11. Overall, the effects of RCL on per capita dollar sales and per capita unit sales are consistently positive across all the categories. As indicated by Table 11, RCL leads to an increase in per capita alcohol dollar sales (β =.04, $p<0.01$) and per capita alcohol unit sales (β =.00, $p<0.01$). In addition, the effect of RCL on per capita tobacco dollar sales (β =.02, $p<0.05$) and per capita tobacco unit sales (β =.00, $p<0.01$) are significantly positive. Similarly, RCL boosts the per capita candy dollar sales (β =.01, $p<0.01$) and per capita candy unit sales (β =.01, $p<0.01$). Finally, RCL has a positive impact on per capita salty snacks dollar sales (β =.03, $p<0.01$) and per capita salty snacks unit sales (β =.01, $p<0.01$). In summary, the DiD results are consistent with SCM results in all the categories except for the tobacco. However, the SCM results are more valid as SCM does not rely on parallel trends assumption which is a key assumption of DiD analysis.

Table 11. DiD Estimation Results

	Per Capita Dollar Sales	Per Capita Unit Sales
	Coeff. Sig.	**Coeff. Sig.**
Panel A: Alcohol		
Variables		
Treat×Post	.04 ***	.00 ***
State fixed effects	Yes	Yes
Time fixed effects	Yes	Yes
Model Fit		
F Statistic	33.53	36.15
Adjusted R-squared	.94	.95
Panel B: Tobacco		
Variables		
Treat×Post	.02 ***	.00 ***
State fixed effects	Yes	Yes
Time fixed effects	Yes	Yes
Model Fit		
F Statistic	114.10	122.89
Adjusted R-squared	. 94	.93
Panel C: Candy		
Variables		
Treat×Post	.01 ***	.01 ***
State fixed effects	Yes	Yes
Time fixed effects	Yes	Yes
Model Fit		
F Statistic	45.15	93.28
Adjusted R-squared	.87	.90
Panel D: Salty Snacks		
Variables		
Treat×Post	.03 ***	.01 ***
State fixed effects	Yes	Yes
Time fixed effects	Yes	Yes
Model Fit		
F Statistic	59.38	60.85
Adjusted R-squared	.94	.94

Notes. ***p < .001.

3.6 Ruling Out Alternative Explanations

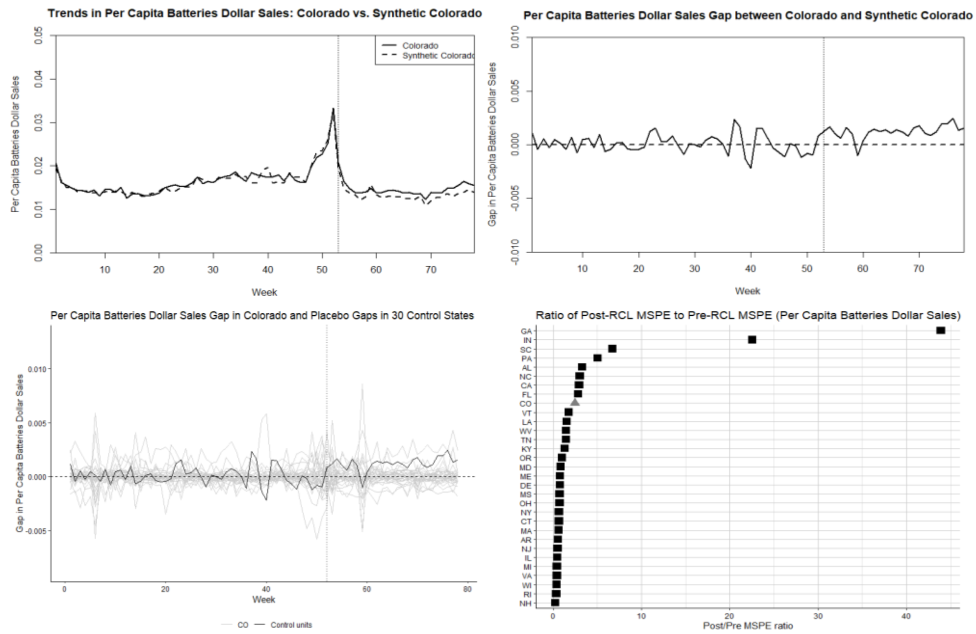
To rule out alternative explanations that may lead to an increase in sales, we conduct several additional analyses. First, we assess the effect of RCL on batteries category to ensure that there would not be spillover effects of RCL on all the categories. Second, we assess the impacts of RCL on pricing and advertising. Generally, if marketers cut price or increase ad spend, dollar sales and unit sales are likely to go up. Therefore, it is essential to check if RCL has any impact on unit price and ad spend based strategies adopted by manufacturers.

3.6.1 Assessing the Effect of RCL on Batteries Category

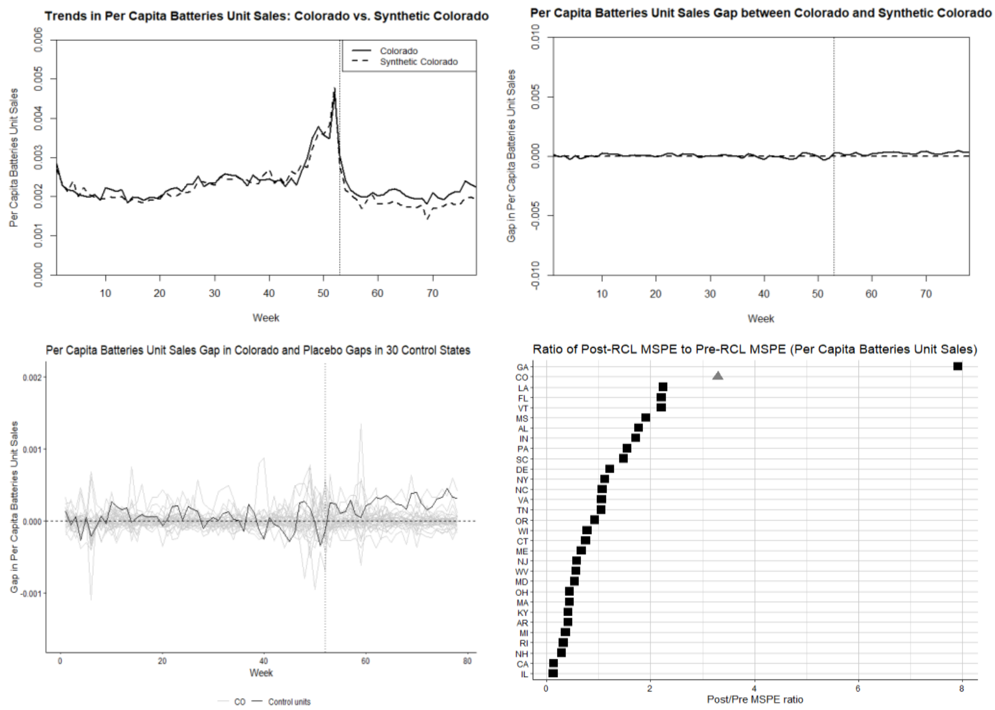
We decide to pick a category that is far removed from any of the other categories and should not have a plausible reason for impact due to RCL. As such, we investigate the impact of RCL on the batteries category. Figure 13 plots the SCM results regarding the implication of RCL on per capita batteries dollar sales and per capita batteries unit sales. We follow the same steps to form the donor pool for batteries category as done for the other categories. To the best of our knowledge, there were no alternative policy changes affecting battery sales during post-RCL period. Therefore, the donor pool consists of 30 control states. As show in Figure 13, it is clear that Colorado and synthetic Colorado follow a similar trend in per capita dollar sales and unit sales during both pre- and post-RCL period. The results of placebo tests also indicate that the effect of RCL on batteries sales are not significant, demonstrating that RCL does not have a general spillover effect on all the categories.

Figure 13. Rule Out Alternative Explanation: Batteries Category[32]

(a) Per Capita Dollar Sales



(b) Per Capita Unit Sales



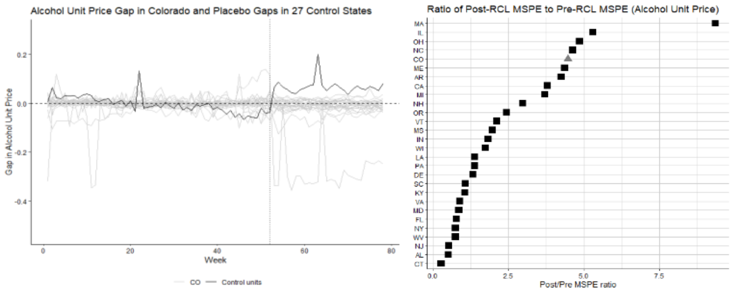
[32] We follow the same steps to form the donor pool for batteries category. To our best knowledge, there were no alternative policy changes affecting batteries sales during post-RCL period. Therefore, the donor pool consists of 30 control states (50-1-1-2-16 = 30).

3.6.2 Assessing the Effect of RCL on Unit Price and Ad Spend

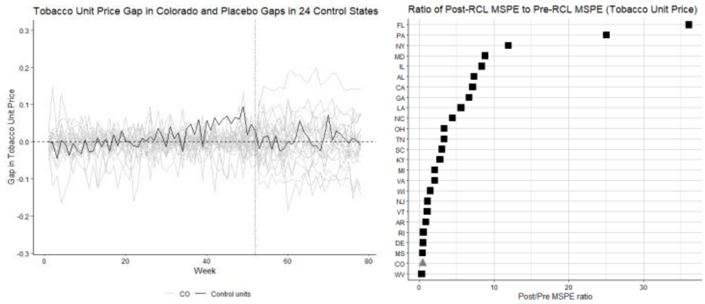
Figure 14 demonstrates that relative to the gaps generated by control states, the gap in unit price of Colorado is not large enough, indicating that RCL does not have significant impacts on unit price across the categories. In summary, RCL has no effect on unit price in these categories. According to Figure 15, the gap in ad spend of Colorado is not larger than the gaps obtained by most of the control states in the placebo test. Therefore, RCL does not affect ad spend during the post-RCL period for all the categories. The detailed SCM results are summarized in Table 12.

Figure 14. Rule Out Alternative Explanation: Unit Price[33]

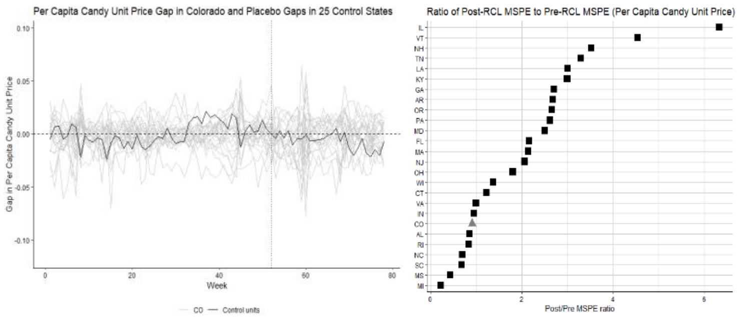
(a) Alcohol



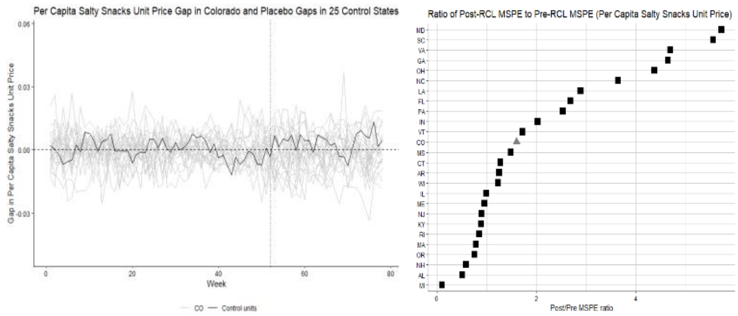
(b) Tobacco



(c) Candy



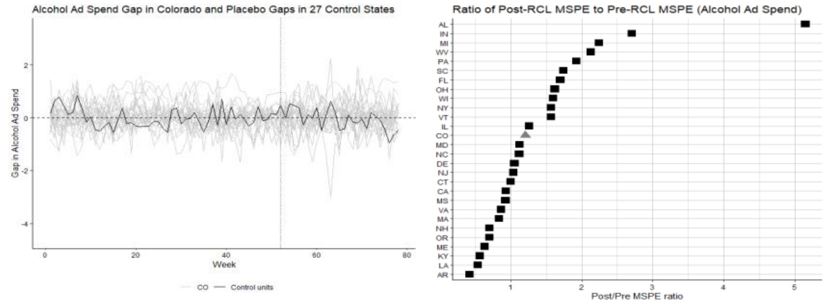
(d) Salty Snacks



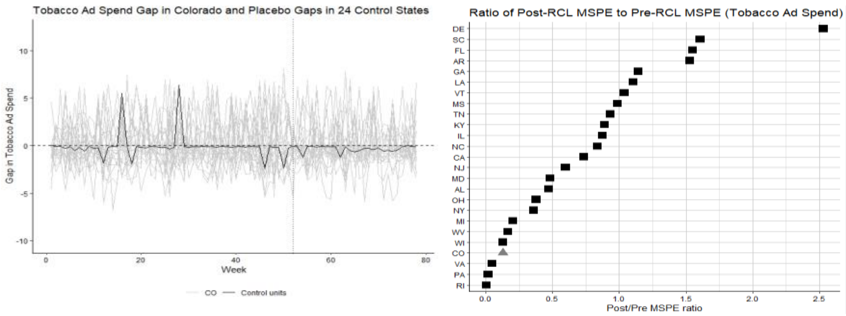
33 To rule out alternative explanation, we investigate the impact of RCL on unit price. The results show that RCL does not impact unit price across the categories, indicating that the increases in per capita dollar sales and per capita unit sales for alcohol, candy and salty snacks after RCL are not driven by the changes in pricing.

Figure 15. Rule Out Alternative Explanation: Ad Spend[34]

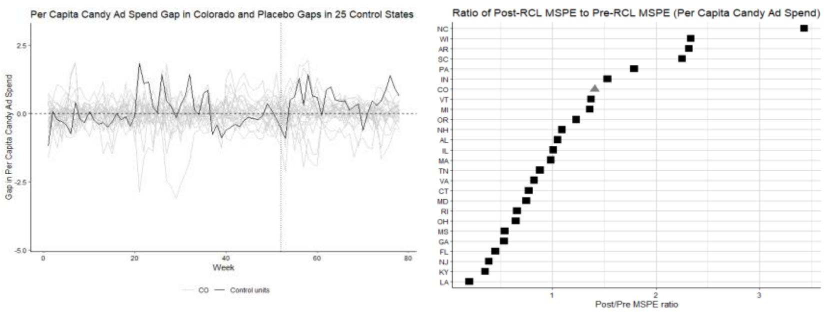
(a) Alcohol



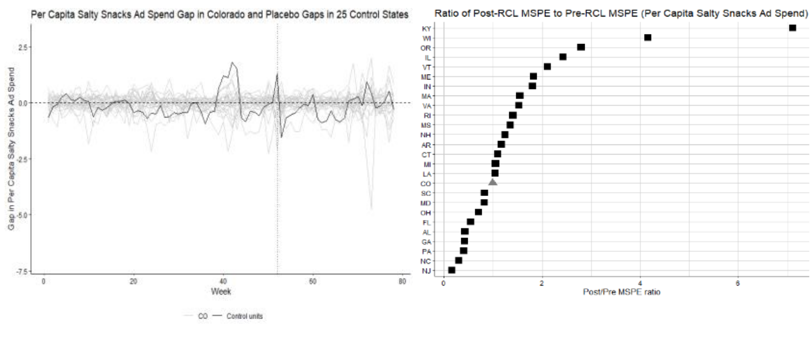
(b) Tobacco



(c) Candy



(d) Salty Snacks



[34] To rule out alternative explanation, we investigate the impact of RCL on ad spend. The results show that RCL does not impact ad spend across the categories, indicating that the increases in per capita dollar sales and per capita unit sales alcohol, candy and salty snacks after RCL are not driven by the changes in advertising.

Table 12:
SCM Results of Unit Price and Ad Spend

	Treatment Effect	Estimated Change in Percentage	Post-/Pre-treatment MSPE (p value)
Panel A: Alcohol			
Unit price (log)	.068	2.96%	4.47 (.18)
Ad spend (log)	-.081	-.73%	1.21 (.46)
Panel B: Tobacco			
Unit price (log)	.007	.27%	.44 (.96)
Ad spend (log)	-.356	NA[35]	.13 (.88)
Panel C: Candy			
Unit price (log)	-.007	-.57%	.91 (.77)
Ad spend (log)	.544	7.99%	1.41 (.27)
Panel D: Salty Snacks			
Unit price (log)	.004	.25%	1.61 (.46)
Ad spend (log)	-.204	-2.91%	.99 (.65)

[35] The percentage is NA because the baseline level of ad spend is zero.

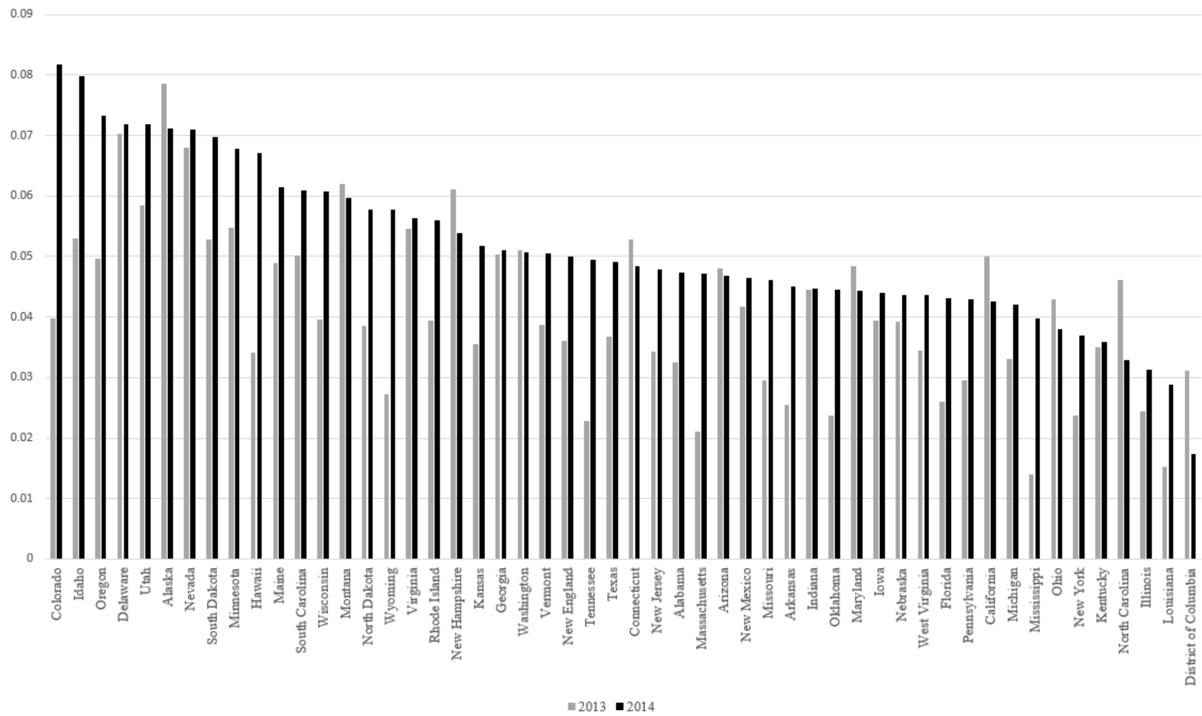
3.7 Discussion

3.7.1 Implications

According to Colorado Department of Public Health and Environment (CDPHE), most of the deaths in Colorado are attributable to chronic diseases and most of the chronic diseases are due to several common risk factors like: excessive alcohol use, obesity, tobacco use, diets high in sodium and saturated fats.[36] Our findings are consistent with the changes in U.S. personal health care expenditure between 2013 and 2014. As can be seen in Figure 16, Colorado raised medicare expenditure by 8.2% in 2014, ranking the top among all the 50 states. Even based on health care costs by state released as recently as 2022, Colorado (spending per capita of $10,254) ranks second in the US, only marginally behind Maine (spending per capita of $10,559).[37]

[36] https://docs.google.com/document/d/1siuIARtZ8VeyihfYSye_5RBK5yVpuNqi3NBEVDs587U/edit
[37] https://worldpopulationreview.com/state-rankings/health-care-costs-by-state

Figure 16. The Growth Rate of U.S. Medicare Spending in 2013 and 2014



Source: Centers for Medicare & Medicaid Services.

Given the complementary relationship between marijuna, alcohol, candy, and salty snacks sales, the potential problems associated with RCL such as excessive drinking and junk food consumption deserve more attention. Excessive drinking may not only alter decision-making and harm mental health but can also cause chronic disease and damage the heart (NIAAA 2021). Indeed, according to Behavioral Risk Factor Surveillance System (BRFSS) survey (2013, 2014), the percentage of respondents who reported angina or coronary heart disease increased after RCL, from 2.6% in 2013 to 3.0% in 2014.

When it comes to the relationship between marijuna and junk food, the received wisdom is that marijuna is associated with more high-calorie snacks consumption (The Economist 2019). States where recreational marijuana has been legalized observe a faster rise in the sales of the snacking and confectionery category than areas where recreational marijuana has not been legalized, according to the NielsenIQ data (NielsenIQ 2019). Specifically, the snack sales in states where cannabis is legal for recreational use achieve a growth rate of 7.2%, as compared to 6% in states where recreational cannabis is illegal. BRFSS surveys (2013, 2014) show that the percentage of adults aged 18 years and above who had an overweight classification increased from 35.1% prior to RCL to 36.1% after RCL. Also, the percentage of respondents who had diagnosed diabetes among adults went up from 6.3% in 2013 to 7.0% in 2014. These health related statistics further validate the findings of our research.

As excessive drinking is likely to be highly costly (e.g., it cost Colorado $5 billion in 2010), policymarkers are suggested to adopt a series of approaches to curb excessive drinking including setting limits for the density of alcohol stores as well as the operating days and hours as well as increasing alcohol price through raising tax rates (State Epidemiological Outcomes Workgoup 2018). In terms of policies tackling obesity, possible methods that policymakers can take include limiting promotions and using plain packaging for fast foods, restricting opening up fast food stores close to schools as well as reducing TV and online advertising (Tapper and McKie 2020). Finally, our research does not find significant increase in tobacco sales due to RCL and hence concerns about the complementary relationship of cannabis and tobacco may be unfounded.

3.7.2 Conclusion

With a recent bill aiming to legalize marijuana in all of US, it is imperative to understand the cross-category unintended consequences of marijuana legalization. Specifically, it is important for policymakers and marketers to understand the impact of RCL on actual sales across categories like alcohol, tobacco, candy and salty snacks. This research sheds light on the cross-category spillover effects of RCL and demonstrates an increase in retail sales of alcohol, candy and salty snacks categories post RCL. We validate these findings by also investigating changes to marketing mix (pricing and advertising) strategies of manufacturers in these categories. We do not find evidence of changes in marketing mix strategy further bolstering our analysis about the unintended consequences of RCL across categories. Finally, we find that RCL does not lead to an increase in tobacco sales.

Chapter 4 Summary of Conclusion

This dissertation explores various aspects of retail analytics through examining consumer search and marketing actions in a retail context using various forms of retail data. Given the crucial role of consumers in the brick-and-mortar environment, this dissertation first examines how they physically interact with the shelf space. Using a novel dataset collected by the state-of-the-art sensing technology, the first essay proposes key dimensions of consumer haptic search capturing the shoppers' speed, consideration set, and shopping path at the shelf space. This research further shows that the effects of consumer haptic search on price paid differ across food and non-food categories. The results offer managerial implications in terms of category assortment, pricing and allocation of shelf space. Apart from customers, marketers play a critical role in the retail sector especially when responding to changes in legislations. The spillover effects of laws and regulations are often of great interest to marketing researchers. Employing the traditional retail data, the second essay sheds light on the spillover effects of RCL on the sales performance and marketing actions (i.e., pricing and advertising) of related industries (i.e., alcohol, tobacco, candy and salty snacks). The conclusions not only assist marketers with identifying potential opportunities and competition, but also help policymakers to understand issues caused by RCL and may lead to rising health care expenses.

In sum, the findings in this dissertation provide significant implications regarding in-store category management and shelf layout as well as dealing with exogenous events to consumers, marketers and policymakers.

References

Abadie A, Diamond A, Hainmueller J (2010) Synthetic control methods for comparative case studies: Estimating the effect of California's tobacco control program. *J. Am. Sta.t Assoc*. 105(490): 493-505.

Abadie A, Diamond A, Hainmueller J (2015) Comparative politics and the synthetic control method. *Am. J. Pol. Sci*. 59(2): 495-510.

Angrist JD, Imbens GW, Rubin DB (1996) Identification of causal effects using instrumental variables. *J. Am. Sta.t Assoc*. 91(434):444-455.

Atefi Y, Ahearne M, Maxham III JG, Donavan DT, Carlson BD (2018) Does selective sales force training work? *J. Marketing Res.* 55(5):722-737.

Athey S, Imbens GW (2017) The state of applied econometrics: Causality and policy evaluation. *J. Econ. Perspect*. 31(2): 3-32.

Baggio M, Chong A (2020) Recreational marijuana laws and junk food consumption. *Econ. Hum. Biol.* 39:100922.

Barrett SP, Darredeau C, Pihl RO (2006) Patterns of simultaneous polysubstance use in drug using university students. *Hum. Psychopharmacol*. 21(4):255-263.

Behavioral Risk Factor Surveillance System (2013) BRFSS 2013 Survey (August 15), https://www.cdc.gov/brfss/annual_data/2013/pdf/Overview_2013.pdf.

Behavioral Risk Factor Surveillance System (2014) BRFSS 2014 Survey (September), https://www.cdc.gov/brfss/annual_data/2014/pdf/Overview_2014.pdf.

Bettman JR (1979) Memory factors in consumer choice. *J. Marketing.* 43(2):37-53.

Bhave A, Murthi BPS (2021) A Study of the Effects of Legalization of Recreational Marijuana on Sales of Cigarettes. Working paper, University of Texas at Dallas, Richardson, TX. Accessed April 9, 2022, https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3508422#:~:text=We%20two%20different%20methods%20%2D%20difference,checks%20and%20find%20similar%20results.

Bombaij NJ, Dekimpe MG (2020) When do loyalty programs work? The moderating role of design, retailer-strategy, and country characteristics. *Int. J. Res. Marketing.* 37*(*1):175-195.

Burmester AB, Becker JU, Van Heerde HJ, Clement M (2015) The impact of pre- and post-launch publicity and advertising on new product sales. *Int. J. Res. Marketing*. 32:408–417.

Carlson KA., Wolfe J, Blanchard SJ, Huber JC, Ariely D (2015) The budget contraction effect: How contracting budgets lead to less varied choice. *J. Marketing Res.* 52(3):337-348.

Chandon P, Hutchinson JW, Bradlow ET, Young SH (2007) Measuring the Value of Point-of-Purchase Marketing with Commercial Eye-Tracking-

Data. Michel Wedel and Rik Pieters, ed. *Visual Marketing: From Attention to Action* (Erlbaum, New York), 225–58.

---- (2009) Does in-store marketing work? Effects of the number and position of shelf facings on brand attention and evaluation at the point of purchase. *J. Marketing.* 73(6):1-17.

----, Wansink B, Laurent G (2000) A benefit congruency framework of sales promotion effectiveness. *J. Marketing.* 64(4):65-81.

Chen M, Burke RR, Hui SK, Leykin A (2021) Understanding lateral and vertical biases in consumer attention: An in-store ambulatory eye-tracking study. *J. Marketing Res.* 58(6):1120-1141.

Clements KW, Daryal M (2005) Exogenous shocks and related goods: drinking and the legalisation of marijuana. *Econ. Lett.* 89(1):101-106.

Cloud K (2015) Food Retail Industry Study Urges Shopper-Centric Approach Vs. Category Management. *The Shelby Repor*t (December 21), https://www.theshelbyreport.com/2015/12/21/food-retail-industry-study-urges-shopper-centric-approach-vs-category-management/.

Cobb CJ, Hoyer WD (1985) The influence of advertising at the moment of brand choice. *J. Advertising*. 14(4):5-27.

Colorado Department of Revenue (2022a) Marijuana sales reports (Accessed April 9, 2022), https://cdor.colorado.gov/data-and-reports/marijuana-data/marijuana-sales-reports.

Colorado Department of Revenue (2022b) Marijuana tax reports (Accessed April 9, 2022), https://cdor.colorado.gov/data-and-reports/marijuana-data/marijuana-tax-reports.

Cox E (2011) Retail Analytics: The Secret Weapon. (John Wiley & Sons).

Cronley ML, Posavac SS, Meyer T, Kardes FR, Kellaris JJ (2005) A selective hypothesis testing perspective on price-quality inference and inference-based choice. *J. Consumer Psychol*. 15(2):159-169.

Crost B, Guerrero S (2012) The effect of alcohol availability on marijuana use: Evidence from the minimum legal drinking age. *J. Health Econ*. 31(1):112-121.

Datta H, Ailawadi KL, Van Heerde HJ (2017) How well does consumer-based brand equity align with sales-based brand equity and marketing-mix response? *J. Marketing.* 81(3): 1-20.

----, Foubert B, Van Heerde HJ (2015) The challenge of retaining customers acquired with free trials. *J. Marketing Res.* 52(2):217-234.

Deloitte (2017) The retail profitability challenge a brave new world. Accessed September 1, https://www2.deloitte.com/content/dam/Deloitte/be/Documents/Retail%20Market%20Study%20Belgium.pdf.

Deloitte (2013). Analytics in retail: Going to market with a smarter approach. Deloitte. Accessed March, 1, https://www2.deloitte.com/content/dam/Deloitte/ch/Documents/consumer-business/ch-cb-en-Deloitte-Analytics-in-retail-0514.pdf.

Dickson PR, Sawyer AG (1990) The price knowledge and search of supermarket shoppers. *J. Marketing*. 54(3):42-53.

Dills AK, Goffard S, Miron J, Partin E (2021) The effect of state marijuana legalizations: 2021 update. *Policy Analysis* (February 2), https://www.fuoriluogo.it/wp-content/uploads/2021/02/effect-marijuana-cato.pdf.

DiNardo J, Lemieux T (2001) Alcohol, marijuana, and American youth: the unintended consequences of government regulation. *J. Health Econ.* 20(6):991–1010.

Eckhardt GM, Husemann KC (2018) The growing business of helping customers slow down. *Harvard Business Review* (December 7), https://hbr.org/2018/12/the-growing-business-of-helping-customers-slow-down.

Estelami H, De Maeyer P (2004) Product category determinants of price knowledge for durable consumer goods. *J. Retailing.* 80(2):129-137.

Gauri DK, Grewal D (2021) Navigating the Retailing Frontier through Academic and Practitioner Collaboration. *J. Retailing*. 97(1):2-5.

Gasparro A, Kang J (2020) Grocers wrest control of shelf space from struggling food giants. *The Wall Street Journal* (Feburary 19), https://www.wsj.com/articles/grocers-wrest-control-of-shelf-space-from-struggling-food-giants-11582108202.

Glandon PJ (2018) Sales and the (Mis) measurement of price level fluctuations. *J. Macroeconomics.* 58:60-77.

Goldfarb A, Tucker C, Wang Y (2022) EXPRESS: Conducting Research in Marketing with Quasi-Experiments. *J. Marketing*, ePub ahead of print April 1, https://doi.org/10.1177/00222429221082977.

Goli A, Chintagunta PK (2021) What happens when a retailer drops a product category? investigating the consequences of ending tobacco sales. *Marketing Sci*. 40(6):1169-1198.

Grant T, Graham J, Carlini B, Ernst C, Brown N (2018) Use of marijuana and other substances among pregnant and parenting women with substance use disorders: Changes in Washington State after marijuana legalization. *J. Stud. Alcohol Drugs*. 79(1):88-95.

Hoban R (2021) The success of Colorado’s marijuana tax dollars. *Forbes* (May 23), https://www.forbes.com/sites/roberthoban/2021/05/23/the-success-of-colorados-marijuana-tax-dollars/?sh=8446b09529df.

Hoyer WD (1984) An examination of consumer decision making for a common repeat purchase product. *J. Consumer Res.* 11(3):822-829.

Hui SK, Bradlow ET, Fader PS (2009) Testing behavioral hypotheses using an integrated model of grocery store shopping path and purchase behavior. *J. Consumer Res. 36*(3):478-493.

----, Fader PS, Bradlow ET (2009) Path data in marketing: An integrative framework and prospectus for model building. *Marketing Sci.* 28(2):320-335.

----, Huang Y, Suher J, Inman JJ (2013) Deconstructing the "first moment of truth": Understanding unplanned consideration and purchase conversion using in-store video tracking. *J. Marketing Res.* 50(4):445-462.

Hull M (2019) The science behind munchies: cannabis and your appetite. *Examine* (April 17), https://examine.com/nutrition/cannabis-munchies/.

Iyengar SS, Lepper MR (2000) When choice is demotivating: Can one desire too much of a good thing? *J. Per. Soc. Psychol.* 79(6):995–1006.

Jacobson R, Obermiller C (1990) The formation of expected future price: A reference price for forward-looking consumers. *J. Consumer Res.* 16(4):420-432.

Jindal P, Zhu T, Chintagunta P, Dhar S (2020) Marketing-Mix Response Across Retail Formats: The Role of Shopping Trip Types. *J. Marketing*. 84(2):114-132.

Kakatkar C, Spann M (2019) Marketing analytics using anonymized and fragmented tracking data. *Int. J. Res. Marketing.* 36(1):117-136.

Kalyanaram G, Little JD (1994) An empirical analysis of latitude of price acceptance in consumer package goods. *J. Consumer Res.* 21(3):408-418.

Kogan L, Papanikolaou D, Seru A, Stoffman N (2017) Technological innovation, resource allocation, and growth. *Q. J. Economics*. 132(2):665-712.

Krishna A, Schwarz N (2014) Sensory marketing, embodiment, and grounded cognition: A review and introduction. *J. Consumer Psychol.* 24(2):159-168.

Kruger JS, Blavos A, Castor TS, Wotring AJ, Wagner-Greene VR, Glassman T, Kruger DJ (2019) Manipulation checking the munchies: validating self-reported dietary behaviors during cannabis intoxication. *Hum. Ethol*. 34:10-16.

Kumar VK, Karande K, Reinartz WJ (1998) The impact of internal and external reference prices on brand choice: The moderating role of contextual variables. *J. Retailing.* 74(3):401-426.

Larson JS, Bradlow ET, Fader PS (2005) An exploratory look at supermarket shopping paths. *Int. J. Res. Marketing.* 22(4):395-414.

Lichtenstein DR, Bloch TH, Black WC (1988) Correlates of price acceptability. *J. Consumer Res.* 15(2):243–252.

Lim LG, Tuli KR, Dekimpe MG (2018) Investors' evaluations of price-increase preannouncements. *Int. J. Res. Marketing.* 35(3):359-377.

Lu S, Rajavi K, Dinner I (2021) The Effect of Over-the-Top Media Services on Piracy Search: Evidence from a Natural Experiment. *Marketing Sci*. 40(3): 548-568.

Malhotra NK (1982) Information load and consumer decision making. *J. Consumer Res.* 8(4):419-430.

Manchanda P, Ansari A, Gupta S (1999) The “shopping basket”: A model for multicategory purchase incidence decisions. *Marketing Sci*.18(2):95-114.

Mathys J, Burmester AB, Clement M (2016) What drives the market popularity of celebrities? A longitudinal analysis of consumer interest in film stars. *Int. J. Res. Marketing.* 32(2):428-448.

Mayhew GE, Winer RS (1992) An empirical analysis of internal and external reference prices using scanner data. *J. Consumer Res.* 19(1):62-70.

Mazumdar T, Papatla, P (2000) An investigation of reference price segments. *J. Marketing Res.* 37(2):246-258.

Mazumdar T, Raj SP, Sinha I (2005) Reference price research: Review and propositions. *J. Marketing.* 69(4):84-102.

----, Jun SY (1993) Consumer evaluations of multiple versus single price change. *J. Consumer Res.* 20(3): 441-450.

McKinsey (2020) Future of retail operations: Winning in a digital era. Retrieved March 1, https://www.mckinsey.com/~/media/McKinsey/Industries/Retail/Our%20Insights/Future%20of%20retail%20operations%20Winning%20in%20a%20digital%20era/McK_Retail-Ops-2020_FullIssue-RGB-hyperlinks-011620.ashx.

Mehta N, Rajiv S, Srinivasan K (2003) Price uncertainty and consumer search: A structural model of consideration set formation. *Marketing Sci.* 22(1):58-84.

Miller K, Seo B (2018) Tax revenues when substances substitute: Marijuana, alcohol, and tobacco. Working paper, University of Oregon, Eugene, OR. Accessed April 9, 2022, https://economics.nd.edu/assets/313828/miller_seo.pdf.

Mitra A, Lynch JG (1995) Toward a reconciliation of market power and information theories of advertising effects on price elasticity. *J. Consumer Res.* 21(4):644-659.

Monroe KB (2003) *Pricing: Making profitable decisions*, 3rd ed. (McGraw-Hill, Boston).

Mogilner C, Rudnick T, Iyengar SS (2008) The mere categorization effect: How the presence of categories increases choosers' perceptions of assortment variety and outcome satisfaction. *J. Consumer Res.* 35(2):202-215.

Nagle TT, Holden RK (1987) *The strategy and tactics of pricing*. (Englewood Cliffs, NJ).

National Institute on Alcohol Abuse and Alcoholism (2021) FY 2021 Budget (Accessed April 9, 2022), https://www.niaaa.nih.gov/sites/default/files/FY2021CJNIAAA-final.pdf.

Ngwe D, Ferreira KJ, Teixeira T (2019) The impact of increasing search frictions on online shopping behavior: Evidence from a field experiment. *J. Marketing Res.* 56(6):944-959.

NielsenIQ (2019) Monetizing the munchies: how legal marijuana use is affecting the U.S. snacking industry (July 17), https://nielseniq.com/global/en/insights/analysis/2019/monetizing-the-munchies-how-legal-marijuana-use-is-affecting-the-us-snacking-industry/.

Olayanju JB (2019) Top trends driving change in the food industry. Forbes (Feburary 16), https://www.forbes.com/sites/juliabolayanju/2019/02/16/top-trends-driving-change-in-the-food-industry/?sh=1d4403b16063.

Papies D, Ebbes P, Van Heerde HJ (2017) Addressing endogeneity in marketing models. *Advanced methods for modeling markets* (Springer, Cham), 581-627.

Park S, Gupta S (2012) Handling endogenous regressors by joint estimation using copulas. *Marketing Sci.* 31(4):567-586.

Peck J, Childers TL (2003) To have and to hold: The influence of haptic information on product judgments. *J. Marketing.* 67(2):35-48.

----, Wiggins J (2006) It just feels good: Customers' affective response to touch and its influence on persuasion. *J. Marketing.* 70(4):56-69.

----, Shu SB (2009) The effect of mere touch on perceived ownership. *J. Consumer Res.* 36(3):434-447.

PwC (2022) Data and Analytics in the Retail sector. Retrieved February 1, https://www.pwc.co.uk/issues/data-analytics/industries/data-and-analytics-in-the-retail-sector.html.

PwC (2015) Food trust Giving customers confidence in your food. Retrieved January 1, https://www.pwc.com/sg/en/publications/assets/food-trust.pdf.

Raghubir P, Valenzuela A (2006) Center-of-inattention: Possible biases in decision-making. *Organ. Behav. Hum. Decis. Processes.* 99(1):66-80.

Ratchford BT, Srinivasan N (1993) An empirical investigation of returns to search. *Marketing Sci.* 12(1):73-87.

Rao A, Wang E (2017) Demand for "healthy" products: False claims and FTC regulation. *J. Marketing Res.* 54(6): 968-989.

Rao AR, Sieben WA (1992) The effect of prior knowledge on price acceptability and the type of information examined. *J. Consumer Res.* 19(2):256-270.

Reutskaja E, Nagel R, Camerer CF, Rangel A (2011) Search dynamics in consumer choice under time pressure: An eye-tracking study. *American Econ. Rev.*101(2):900-926.

Roggeveen AL, Grewal D, Schweiger EB (2020) The DAST framework for retail atmospherics: The impact of in-and out-of-store retail journey touchpoints on the customer experience. *J. Retailing.* 96(1):128-137.

Schauer GL, Berg CJ, Kegler MC, Donovan DM, Windle M (2015) Assessing the overlap between tobacco and marijuana: Trends in patterns of co-use of tobacco and marijuana in adults from 2003–2012. *Addict. Behav.* 49:26-32.

Schlienz NJ, Lee DC (2018) Co-use of cannabis, tobacco, and alcohol during adolescence: policy and regulatory implications. *In.t Rev. of Psychiatry*. 30(3):226-237.

Schweidel D, Knox G (2013) Incorporating direct marketing activity into latent attribution models. *Marketing Sci.* 32(3), 471-487.

Schneider LG, Currim IS (1991) Consumer purchase behaviors associated with active and passive deal-proneness. *Int. J. Res. Marketing.* 8(3):205-222.

Seiler S, Pinna F (2017) Estimating search benefits from path-tracking data: measurement and determinants. *Marketing Sci.* 36(4):565-589.

----, Yao S (2017) The impact of advertising along the purchase funnel. *Quan. Marketing Econ.* 15:241-278.

Shafir E, Simonson I, Tversky A (1993) Reason-based choice. *Cognition*. 49(1-2):11-36.

Shannon CE, Weaver W (1949) The mathematical theory of information. *Urbana: University of Illinois Press.* 97.

Song I, Chintagunta PK (2006) Measuring cross-category price effects with aggregate store data. *Management Sci*. 52(10):1594-1609.

Sorce P, Widrick SM (1991) Individual Differences in Latitude of Acceptable Prices. *Adv. Consumer Res.* 18:802-805.

Sorensen H (2009) *Inside the Mind of the Shopper: The Science of Retailing*. (Pearson Prentice Hall).

Soster RL, Gershoff AD, Bearden WO (2014) The bottom dollar effect: The influence of spending to zero on pain of payment and satisfaction. *J. Consumer Res.* 41(3):656-677.

State Epidemiological Outcomes Workgoup (2018) Reduce Excessive Drinking To Build Healthier Communities (Accessed April 9, 2022), https://www.coloradoseow.org/wp-content/uploads/2019/08/SEOW-Alcohol-Policy-Brief_Volume-1_Reduce-Excesive-Drinking_April-2018.pdf.

Strack F, Deutsch R (2004) Reflective and impulsive determinants of social behavior. *Personality Soc. Psychol. Rev.* 8(3):220-247.

Tapper J, McKie R (2020) ‘Junk food is the new tobacco’: experts call for restrictions to tackle obesity. *The Guardian* (July 19), https://www.theguardian.com/society/2020/jul/19/junk-food-is-the-new-tobacco-experts-call-for-restrictions-to-tackle-obesity#:~:text=1%20year%20old-,'Junk%20food%20is%20the%20new%20tobacco'%3A%20experts%20call,for%20restrictions%20to%20tackle%20obesity&text=Ministers%20should%20regulate%20processed%20food,crisis%2C%20health%20experts%20have%20warned.

The Economist (2019) Legal weed is linked to higher junk-food sales (August 16), https://www.economist.com/graphic-detail/2019/08/16/legal-weed-is-linked-to-higher-junk-food-sales.

Tirunillai S, Tellis GJ (2017) Does offline TV advertising affect online chatter? Quasi-experimental analysis using synthetic control. *Marketing Sci*. 36(6):862-878.

Valenzuela A, Raghubir P (2009) Position-based beliefs: The center stage effect. *J. Consumer Psychol.* 19(2):185-196.

Van Den Bergh B, Heuvinck N, Schellekens GA, Vermeir I (2016) Altering speed of locomotion. *J. Consumer Res.* 43(3):407-428.

Van Nierop E, Bronnenberg B, Paap R, Wedel M, Franses PH (2010) Retrieving unobserved consideration sets from household panel data. *J. Marketing Res.* 47(1):63-74

Wang P, Xiong G, Yang J (2019) Frontiers: Asymmetric effects of recreational cannabis legalization. *Marketing Sci*. 38(6): 927-936.

Wang R (2018) When prospect theory meets consumer choice models: Assortment and pricing management with reference prices. *Manuf. Serv. Oper. Management.* 20(3):583-600.

Winer RS (1986) A reference price model of brand choice for frequently purchased products. *J. Consumer Res.* 13(2):250–256.

Yakowicz W (2021) U.S. cannabis sales hit record $17.5 billion as Americans consume more marijuana than ever before. *Forbes* (March 3), https://www.forbes.com/sites/willyakowicz/2021/03/03/us-cannabis-sales-hit-record-175-billion-as-americans-consume-more-marijuana-than-ever-before/?sh=477ee07f2bcf.

Zaichkowsky JL (1985) Measuring the involvement construct. J. Consumer Res. 12(3):341-352.

Zhang X, Li S, Burke RR, Leykin A (2014) An examination of social influence on shopper behavior using video tracking data. *J. Marketing*. 78(5):24-41.