UNIVERSITY OF CALIFORNIA

Los Angeles

State Dependence in Brand, Store, and Category Choice

A dissertation submitted in partial satisfaction

of the requirements for the degree

Doctor of Philosophy in Management

by

Julia Levine

2023

ABSTRACT OF THE DISSERTATION

State Dependence in Brand, Store, and Category Choice

by

Julia Levine

Doctor of Philosophy in Management

University of California, Los Angeles, 2023

Professor Stephan Seiler, Co-Chair

Professor Randolph E. Bucklin, Co-Chair

Across a wide variety of contexts, people who have experienced an event are more likely to experience that event in the future. This empirical regularity, hereafter referred to as state dependence, has two explanations, each with its own set of policy and managerial implications. One explanation, known as structural state dependence, is that the experience of an event alters the preferences or constraints that an individual would hold for that event in the future. A second explanation, spurious state dependence, is that people differ along some unobservable propensity to experience an event. For example, people that become unemployed once are more likely to be unemployed in the future. The structural explanation for this phenomenon is that unemployment has a sustained effect on the probability of future unemployment, while the spurious explanation argues that individuals vary on some unobservable variable, such as work ethic or skill set, that affects their probability of becoming unemployed at any time. These explanations have different implications: if state dependence is structural, short-term policies reducing unemployment can have large long-run effects. This dissertation aims to explore the effects of structural state dependence in three contexts: brand choice, store choice, and category consumption.

Using techniques in causal inference and structural modeling, and a rich database of transaction data, I find that structural state dependence 1) has no effect on brand choice in consumer packaged goods, 2) has a strong effect on where people shop for groceries, impacting nutritional intake, and 3) drives consumption in addictive categories to varying extents. These findings give us a better

understanding of why brand choice persists over time, why nutritional intake varies drastically across demographic groups, and how cigarette types vary in their addictive and habit-forming properties.

The dissertation of Julia Levine is approved.

Peter E. Rossi

Sylvia Hristakeva

Stephan Seiler, Committee Co-Chair

Randolph E. Bucklin, Committee Co-Chair

University of California, Los Angeles

2023

Contents

List of Figures

List of Tables

ACKNOWLEDGMENTS

I am deeply indebted to my committee members and the Marketing faculty at UCLA for their guidance and support throughout the doctoral program. My fourth year in the program was interrupted by COVID-19, at which point I would have been floundering and lost without the kindness of this group. I would like to thank Elisabeth Honka for her generosity with her time and honest feedback, and Brett Hollenbeck for his thoughtful comments and suggestions. I am grateful to Randy Bucklin for his unwavering optimism and support throughout many phases of my research interests, and to Peter Rossi for his kind advice and exacting questions that helped me learn to communicate my work. I would also like to thank my co-author and committee member, Sylvia Hristakeva, for her friendship and guidance – I hope to emulate her resolve, work ethic, and honesty. Lastly, I am indebted to my advisor, Stephan Seiler. I met Stephan during his time as a visiting professor at UCLA, and during the early days of quarantine we began working on what would become the first chapter of this dissertation. After his move to Imperial College London he was no less attentive, patient, and communicative. I admire his ability to see the potential in ideas and to think of simple but effective ways to conduct research. I feel incredibly lucky to have had him as a mentor.

I would also like to acknowledge the students who I shared this journey with: Jonathan Bogard, Ipek Demirdag, Sherry He, Pedro Makhoul, Joey Reiff, and David Zimmerman. I cannot imagine doing this without you, and I owe much of my happiness throughout the doctoral program to the formation of this game night crew. To Joey, I owe all of the rest.

Chapter 1 is a version of Levine and Seiler (2022). We are very thankful to Andrey Simonov, Jean-Pierre Dubé, Günter Hitsch, and Peter Rossi for sharing their code for the estimation of a model with state dependence that accounts for consumers' initial conditions and to Andrey Simonov in particular for helping us better understand their code.

Chapter 2 is work in progress in collaboration with Sylvia Hristakeva. Funding for this research partially came from the Morrison Center for Marketing & Data Analytics.

All of the research presented in this dissertation utilizes NielsenIQ datasets made available through the Kilts Center. 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 NielsenIQ data are those of the researchers and do not reflect the views of NielsenIQ. NielsenIQ is not responsible for, had no role in, and was not involved in analyzing and preparing the results reported herein.

VITA

2013-2017	B.A. Economics, University of Florida
2016-2017	Teaching Assistant, University of Florida
2017-2022	Anderson Fellowship, UCLA Anderson School of Management
2018-2022	Teaching Assistant, UCLA Anderson School of Management

1 Identifying State Dependence in Brand Choice: Evidence from Hurricanes

Abstract: We analyze structural state dependence in brand choice using variation from brand switching during stock-outs caused by hurricanes. We derive a simple test for structural state dependence based on the time-series of choice persistence for households affected by the stock-outs. Using data from the bottled water category, we show that demand increases substantially before hurricanes, causing households to purchase different brands. We find that purchase behavior reverts back to its pre-hurricane trajectory immediately after a hurricane and we are not able to reject the null hypothesis of no structural state dependence. By contrast, the common approach of estimating structural state dependence based on temporal price variation via a discrete choice model yields a positive effect using data for the same category. We argue that our approach is better suited to identify the causal impact of past choices because it requires fewer assumption and is based on more plausibly exogenous variation in brand switching due to stock-outs.

1.1 Introduction

A large literature in marketing and economics (e.g., Jones and Landwehr (1988), Keane (1997), Seetharaman et al. (1999), Dubé et al. (2010a)) documents that consumers are persistent in their choices and are more likely to purchase products they purchased in the past. Such persistence can be explained either by time-invariant preference heterogeneity or by a causal effect of past choices on current purchase behavior. The distinction between these two explanations, also referred to as *spurious* and *structural* state dependence, (Heckman (1981)) respectively, is important for understanding the dynamics of consumer choice and has implications for optimal firm policies such as pricing (Dubé et al. (2008)).[1] In this paper, we provide a novel framework for identifying structural state dependence, and we show in an application based on data from a consumer packaged goods category that consumers do not exhibit structural state dependence.

Our approach involves the collection of new data and the development of a new and simple test

[1]Similar to Dubé et al. (2009), Dubé et al. (2010a), and related papers, we focus on the impact of consumers' choices in the preceding period on current period choices. We do not consider other forms of temporal dependence, such as learning, where current choices depend on choices in multiple earlier time periods.

for the presence of structural state dependence. In terms of data, we gather information on the location and timing of hurricanes that cause demand spikes and therefore stock-outs in consumer packaged goods (CPG) categories and combine it with consumer-level purchase data.[2] We observe fourteen hurricanes over the course of twelve years that affect thousands of households, leading to increased brand switching behavior. We use these data to test for the presence of structural state dependence based on the time-series of choice persistence and its evolution in reaction to the exogenous shock induced by a hurricane. Our approach allows us to test for structural state dependence without making assumptions about the distribution of preference heterogeneity and without modeling consumers' initial conditions, both of which are important assumptions in prior work on structural state dependence (e.g., Simonov et al. (2020)). The key idea of our identification strategy is that, under the null hypothesis of no structural state dependence, brand choice during the hurricane will have no impact on future choices and therefore purchase probabilities will revert back to their pre-hurricane levels immediately after the hurricane.

We apply our framework to data from the bottled water category, for which we observe a large demand spike in the period leading up to a hurricane. We find that the purchase probability for products and brands purchased prior to this demand spike decreases significantly around the time of the hurricane, but reverts back to its pre-hurricane trajectory immediately after the hurricane. We are thus not able to reject the null hypothesis of no structural state dependence. Due to slight seasonal fluctuations in purchase behavior for bottled water, we implement a test that only analyzes behavior in a short window around the hurricane in addition to an analysis based on a generalized synthetic control approach and a two-way fixed effects model. All tests generate similar results and the null effect is precisely estimated.

Our empirical findings differ from most prior papers (e.g., Keane (1997), Seetharaman et al. (1999), Dubé et al. (2010a), Simonov et al. (2020)) that estimate a structural model of consumer choice and tend to find that consumer behavior is characterized by some degree of structural state dependence. Interestingly, when we estimate a demand model with state dependence (following the the approach in Simonov et al. (2020)) using data from the bottled water category,[3] we also find a

[2]We do not observe stock-outs directly, but we observe a demand increase and increased brand switching behavior around the time of a hurricane. We exploit the increased brand switching behavior which is likely caused by stock-outs to study state dependence in choices.

[3]We use a different set of households for the estimation of the structural demand model, because we need to impute prices for non-purchased products. The imputation of prices is only possible for households that visit stores

positive and significant impact of past choices. The estimated average structural state dependence term is similar in magnitude to the estimated effect in Simonov et al. (2020) based on margarine data. We also find that the impact of a stock-out implied by the structural model estimates lies far outside of the confidence interval of the estimate based on our approach and hence our null result is not driven by a lack of statistical power.

In order to reconcile the differences in results between the two approaches, we first analyze whether the specific setting of hurricane-induced stock-outs might affect our findings. To this end we show that our results are not driven by longer-term disruptions in purchase behavior due to a hurricane. We also show that the estimated null effect is not due to unusual purchase behavior during the hurricane such as switching to niche products or bulk buying. Finally, several data patterns suggest that brand switching during hurricanes is not driven by context-specific purchase behavior when preparing for a hurricane. Taken together these robustness checks provide evidence that hurricanes only affect consumer brand choice behavior through stock-outs and not through any other direct channel.

Having ruled out these alternative explanations, we argue that two key advantages of our approach might be driving the difference in results. First, our approach identifies structural state dependence based on hurricane-induced stock-outs, whereas other papers typically rely on price variation due to discounts. Identification in either setting requires past prices or past stock-outs to affect current choices only through their impact on past choices. We believe this assumption is more likely to be fulfilled in the case of hurricane-driven stock-outs whereas past prices conceivably correlate with marketing activity such as advertising or preferential shelf placement that might be persistent over time and affect current choices.

Second, the prior literature on state dependence requires the researcher to model preference heterogeneity flexibly in order to separate structural state dependence from spurious state dependence. Paulson (2012) argues that functional form assumptions on preference heterogeneity can make it difficult to separately identify the lagged choice term, i.e., structural state dependence. Dubé et al. (2010a) therefore allow for flexible functional forms (mixtures of normals) of heterogeneity. Moreover, Simonov et al. (2020) show that not modeling consumers' initial conditions correctly can lead

that are also observed in the Nielsen store-level data. We provide more details on sample construction in Section 1.5 and Appendix 1.7.5.

to biased estimates of structural state dependence. A key advantage of our approach is that we do not need to specify preference heterogeneity nor do we need to explicitly account for initial conditions. Our approach therefore avoids possible model mis-specification that could arise from a failure to correctly model the initial condition or an insufficiently flexible distribution of preference heterogeneity.

Apart from the literature on structural state dependence cited above, this paper is also related to the literature on product availability and stock-outs (e.g., Anupindi et al. (1998), Bruno and Vilcassim (2008), Musalem et al. (2010), Vulcano et al. (2012), Conlon and Mortimer (2013)). In our setting, we do not observe stock-outs directly, but we show that demand increases strongly in the weeks leading up to a hurricane, followed by an increase in brand switching behavior. We surmise that the demand spike leads to stock-outs which, in turn, trigger subsequent brand switching. We exploit the observed increase in brand switching to study structural state dependence. In a related paper Sudhir and Yang (2014) study structural state dependence based on data from rental car upgrades where consumers obtain a different car from the one they originally booked. Figueroa et al. (2019) study the effects of stock-outs by analyzing the impact of an earthquake that damaged the factories of two leading beer brands in Chile and led to stock-outs that spanned several weeks. The paper finds that the stocked-out brands had lower market shares in the post-stock-out period, whereas smaller brands increased their market shares, which the paper interprets as a shift in consumers' valuations of different brands. Our setting involves short-term stock-outs that affect most consumers on only one purchase occasion, making it better suited for the identification of structural state dependence rather than longer-term brand preference effects.

The remainder of the paper is organized as follows. In Section 1.2 we present the data and descriptive statistics. In Section 1.3 we outline our empirical framework and illustrate our identification strategy using simulations based on a consumer choice model with and without structural state dependence. In Section 1.4 we present our main empirical analysis and robustness checks. In Section 1.5 we estimate state dependence based on a structural model of consumer choice. We show that such an approach leads to different results with regards to structural state dependence and discuss differences relative to our estimation approach. We provide concluding remarks in Section 1.6.

1.2 Data

We rely on three sources of data for our empirical analysis. We use HURDAT2, a hurricane tracking data set collected by the National Hurricane Center, in conjunction with the store-level Nielsen Retail Scanner data in order to identify geographical areas that were affected by hurricanes as well as the precise timing of when those areas were affected. We then select households from the Nielsen Consumer Panel dataset who lived in these locations and study how their purchase behavior is affected by the hurricanes. Below, we first describe how we select households that were affected by a hurricane (which we simply will refer to as the "treatment group" going forward) and how we match treated households with a set of control households. We then describe how the panel data set used for our main analysis is constructed and how we define key variables.

1.2.1 Household Selection

We use storm location data (so-called "best track" data) to identify households that were affected by hurricanes. Through post-storm analysis, best track data provides the best estimates of location and intensity at each point in the storm's track. These data are usually compiled through a combination of aircraft reconnaissance ("Hurricane Hunters") and satellite remote sensing.[4] We use HURDAT2, a well known best track data set collected by the National Hurricane Center. These data include coordinates of each active storm at three times each day, as well as information on wind intensity, wind radii, and pressure. We limit the data to storms that eventually became hurricanes and made landfall somewhere in the continental U.S.

For the purpose of our analysis, we want to identify households whose purchase behavior changed due to stock-outs that occurred following hurricane preparations, but do not require that households were directly affected by the presence of a storm. Thus we aim to identify households that were located in areas that *anticipated* a hurricane rather than areas that were actually hit. Due to imperfect forecasts, the former and the latter are not necessarily identical. To the best of our knowledge there is no record of the forecasts that were made prior to each hurricane and we therefore have to resort to a more indirect technique that combines the hurricane data with the Nielsen Retail Scanner data, which records purchases at the store-level across a large set of stores. We use the

[4] https://www.air-worldwide.com/publications/air-currents/2013/Best-Track-Data/

Hurricane	Month	Year	# Treated States	# Treated Counties	# Treated Households
Sandy	Oct	2012	16	224	6,537
Irma	Sep	2017	9	99	2,611
Ernesto	Sep	2006	4	34	1,487
Harvey	Sep	2017	3	9	1,138
Isaac	Aug	2012	5	80	1,002
Gustav	Sep	2008	5	48	675
Irene	Aug	2011	6	49	569
Matthew	Oct	2016	3	23	563
Ike	Sep	2008	6	29	536
Hermine	Sep	2016	3	17	244
Hanna	Sep	2008	3	18	191
Dolly	Jul	2008	1	8	170
Arthur	Jul	2014	1	5	109
Humberto	Sep	2007	1	1	6
					15,838

Table 1.1: **Hurricanes.** Counts of affected states, counties, and households for each hurricane.

Nielsen Retail Scanner (RMS) data to identify counties where stores exhibited preparation behavior in the week of a storm.[5]

Based on exploratory analysis we identify three product groups that are likely to experience demand spikes in anticipation of a storm: bottled water, canned soup, and batteries/flashlights. We consider a product group as experiencing a demand spike if the total units sold in a county during a storm week is at least two standard deviations above the average weekly units for that county and product group. We then define counties as treated if they experienced demand spikes for at least two out of the three groups of hurricane staples. To rule out idiosyncratic demand spikes that are unrelated to the hurricane, we drop counties that are far away from the storm.[6] Our final sample contains households that were affected by at least one out of fourteen hurricanes. Table 1.1 reports a list of these hurricanes and the number of households that lived in affected counties.

Next, we select control households from the set of all untreated households in the Nielsen data.

[5]Many stores in the consumer panel data are not observed in the RMS data. We therefore define affected counties, instead of affected stores, based on the store-level data and then identify households that live in those counties in the consumer-level data. A county is the most granular measure of where stores are located in the RMS data.

[6]We retain counties that fall within the storm's most inclusive radius, or where the distance to the center of the storm is less than the median distance of counties with demand spikes. The most inclusive wind radius is defined as the maximum distance from the center of the storm where a wind intensity of at least 34 knots (the lowest wind intensity reported in data) is recorded. The radius is reported separately for four directions (NE, NW, SE, SW).

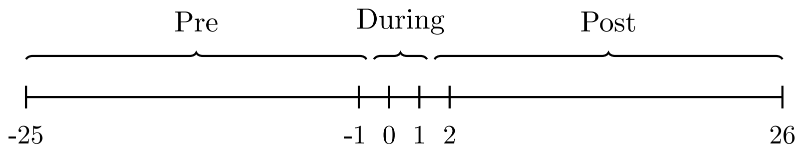


Figure 1.1: **Timeline of the Estimation Sample.** The graph shows a running counter of weeks. Week 0 is defined as the week that ends in the hurricane. The "during" period comprises week 0 and week 1. The pre and post periods comprise 25 weeks each.

Specifically, we select households that live in a county that was at least 100 miles from the storm and where there were no demand spikes for any of the three groups of hurricane staples in a storm week. We randomly select two controls for each household treated on a given date.[7] Our data contain 15,838 treated households between 2006 and 2017 and 31,676 control households.[8]

We track each treated household in the sample for a period of one year surrounding a hurricane event and track each control household for the same time period as the treated household they are assigned to. The unit of observation in our estimation sample is a household (i) / week (t) combination. For ease of exposition, we define a set of time periods for each household. We consider the week leading up to the hurricane as well as the week following the hurricane as weeks that are likely to generate different purchases due to stock-outs. We also retain data for the 25 weeks before and after the two weeks affected by the hurricane. Together, the pre- / during- / post-hurricane periods constitute a sample of 52 weeks per household.[9] Figure 1.1 displays the timing and notation for our main estimation sample. We denote the week leading up to the hurricane and the week after as weeks 0 and 1.[10] The pre- and post-period comprise weeks -25 to -1 and weeks 2 to 26 respectively.

[7]Controls are sampled without replacement from the pool of all eligible controls in that year. We choose a relatively conservative radius of 100 miles when selecting control households to assure that they are not affected by the hurricane. Because the set of possible control households in the Nielsen data is large relative to the number of treated households, this selection rule does not impact the size of our control group.

[8]A small number of households experience, or serve as controls for, multiple hurricanes. For these households, we construct a separate time series of 52 weeks around each hurricane event. For simplicity we refer to households throughout the text, when more precisely it should be a household / hurricane combination. There are 15,838 treated and 31,676 control household-hurricane combinations. There are 15,047 distinct treated and 28,381 control households.

[9]When calculating choice persistence on a particular shopping trip, we need to compare purchases on the specific trip with purchases made on the previous trip. In order to define choice persistence on the first trip during the main sample period, we use previous trips during weeks –50 to -26.

[10]For each household, we define week 0 so that the final day of week 0 coincides with the final day of the hurricane.

1.2.2 Choice Persistence

We define choice persistence within a category as

$$Persist_{it} = \frac{1}{\#J_{it}} \sum_{j \in J_{it}} 1(j \in J_{it}^{last}). \tag{1.1}$$

where J_{it} denotes the list of brands purchased by consumer i in week t in a given category and J_{it}^{last} denotes the list of brands purchased in the previous week in which the consumer made a purchase in the category. This variable measures how many of the brands purchased in a given week are identical to brands that the consumer also chose the last time she purchased in the category. Consumers usually purchase only one brand within the focal category during one shopping trip per week, in which case the variable is simply an indicator that is equal to one if the current purchase is identical to the brand purchased previously. Our formulation allows for the fact that consumers occasionally buy multiple brands on a given shopping trip and we also aggregate purchases from shopping trips that occur within the same week. We need to aggregate the data at this level because we later analyze the data using a generalized synthetic control approach, which does not allow for multiple observations for a given household within the same time period. Going forward we simply use the terminology "previous shopping trip" instead of "previous week with a purchase in the category". We analyze choice persistence both at the brand and the product level. Depending on the level of analysis, J_{it} and J_{it}^{last} therefore either refer to lists of brands or lists of UPCs.

1.2.3 Category Selection & Descriptive Statistics

In our main empirical analysis, we focus on the bottled water product category.[11] Bottled water is purchased heavily in preparation for hurricanes and is therefore likely to experience a stock-out, causing a disruption in households' product choices.[12] For most of our analysis (and unless stated otherwise) we select households that made at least one purchase during weeks 0 and 1 and were therefore affected by a hurricane. We also condition on households that made at least 4 purchases in the category in the pre-hurricane period to focus on households that purchased frequently in our

[11]The bottled water data used throughout the paper does not include *carbonated* water which is categorized separately in the Nielsen-Kilts data.

[12]We choose bottled water because many consumers purchase frequently in this category and the degree of brand switching due to stock-outs during hurricanes is relatively large. We experimented with data from other categories and found that hurricanes triggered less brand switching and/or the sample of affected households was smaller.

	Focal Category: Bottled Water	Margarine	Orange Juice
Top Brands	Poland Spring, Nestle Pure Life, Deer Park	Imperial, Blue Bonnet, Smart Balance	Simply Orange, Tropicana, Minute Maid
# Brands	657	76	189
# Brands (>3% market share)	7	9	4
# UPCs	4,608	914	1,671
# UPCs (>0.5% market share)	31	52	53
Share of Weeks With a Purchase	0.431	0.317	0.460
Av. Choice Persistence (Brand level)	0.661	0.691	0.655
Av. Choice Persistence (UPC level)	0.436	0.549	0.464

Table 1.2: **Descriptive Statistics.**

focal category.[13] Out of all treated households that were affected by a hurricane we retain 2,201 households for our main analysis.[14]

Table 1.2 contains basic descriptive statistics for the bottled water category. For comparison, we also report the same set of descriptive statistics for two other commonly studied CPG product categories: margarine and orange juice. Bottled water contains 657 brands and 4,608 UPCs, but only a small number of brands (UPCs) have a market-share of over 3% (0.5%). Table 1.2 also describes choice persistence for each product category, calculated as shown in equation (1.1), averaged across all treated households and shopping trips in the pre-hurricane period. For all product categories, choice persistence at the UPC level is naturally lower because consumers might switch to a different product that belongs to the same brand. For bottled water choice persistence is equal to 0.661 at the brand level and 0.436 at the UPC level. This level of choice persistence is comparable with that of margarine and orange juice.

Before turning to our main analysis, we illustrate the nature of the variation we aim to exploit. In the top graph of Figure 1.2 we plot the evolution of weekly average expenditure per household in the bottled water category over time.[15] The graph is centered around the hurricane event for each

[13] These criteria are used for all of our main empirical analysis in Section 1.4 except for one robustness check that uses a different sample selection criterion.

[14] Based on the same criteria we retain 3,866 control households.

[15] The graph plots *unconditional* average spending of households in our sample. In most weeks a share of households does not purchase in the category. We use a larger sample than our main estimation sample in this graph, namely all households that purchased bottled water at least once during the sample period.

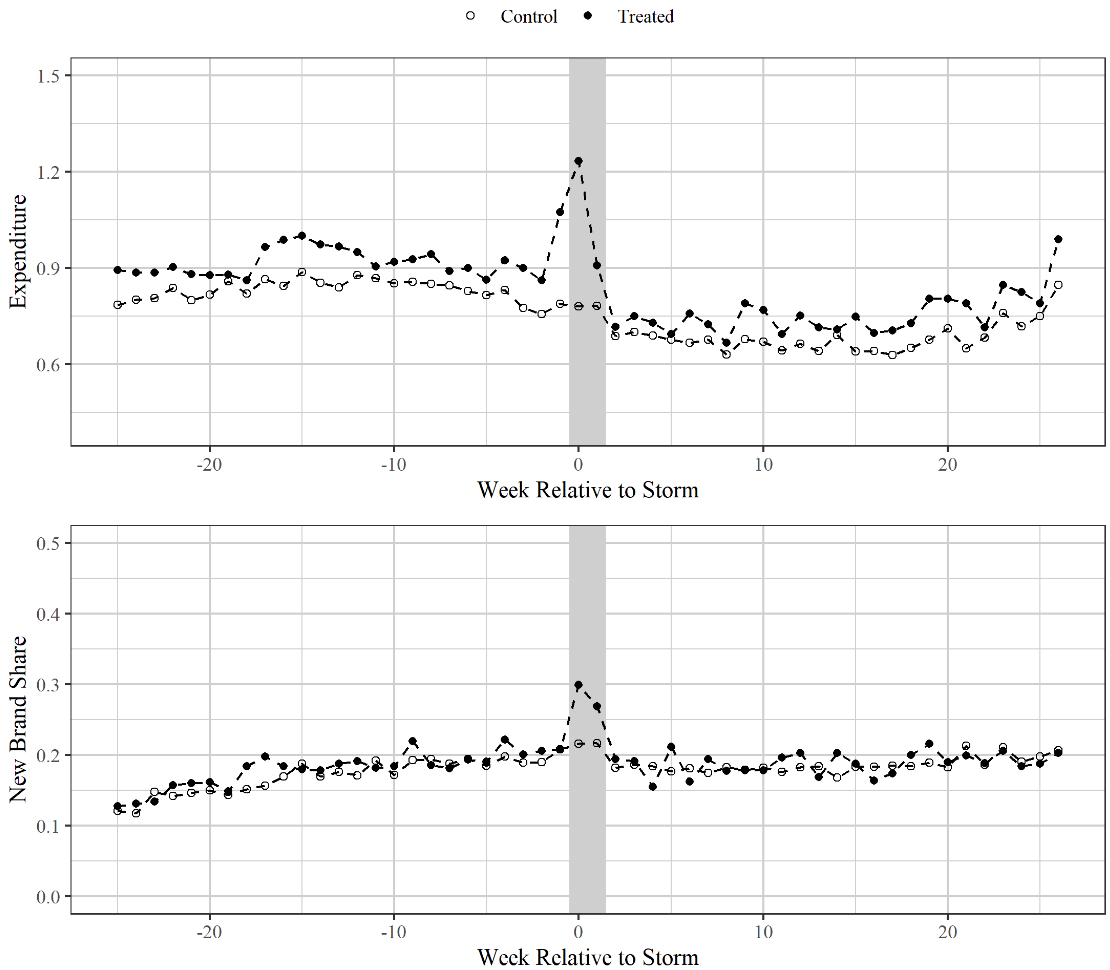


Figure 1.2: **Bottled Water Expenditure and Purchases of New Brands.** The top graph displays average weekly expenditure (in dollars) per household in the bottled water category for the treatment and control group. New brand share is the share of unique brands purchased on a given shopping trip that were not purchased during a six month period preceding the main sample. The bottom graph displays the average value of this variable. The vertical gray bars indicate weeks 0 and 1 which are likely to be affected by stock-outs.

household and shows that expenditure increased substantially in the week of the hurricane (week 0) as well as in the week before (week -1) when households were likely preparing for the hurricane.[16] The spike in demand around the hurricanes leads to the shock to purchase behavior that we aim to exploit for our empirical analysis: During the weeks leading up to a hurricane, household expenditure rises and therefore stock-outs of individual brands become more likely, resulting in different purchases because households are unable to purchase their preferred brand.

[16] We also observe a slightly higher-than-usual expenditure pattern in week 1 after the hurricane in the treatment group, possibly due to imperfect data on the exact timing of the hurricane.

In the bottom graph of Figure 1.2 we show evidence of this sequence of events, plotting the evolution of the share of new brands purchased averaged across households. We define "new brand share" as the share of unique brands purchased on a given shopping trip that were not purchased during a six month period preceding the main sample. We find that the average share of new brands purchased displays a large increase from 20% to 30% during weeks 0 and 1 which are highlighted by the gray bar. Taken together, the two graphs show that the expenditure spikes in weeks 0 and -1 are lagged by one period relative to the two weeks that we consider to be affected by stock-outs. This pattern is consistent with the sequence of events driving brand switching, namely that hurricanes lead to higher demand in weeks -1 and 0 which leads to stock-outs that occur in weeks 0 and 1, whereas stores are able to refill stocks by week 2. We emphasize that the share of new brands purchased increases only in weeks 0 and 1, but not in week -1 despite the observed increase in demand. This pattern suggests that brand switching is not merely due to different behavior when preparing for a hurricane but is rather driven by stock-outs.[17] We re-iterate that we do not directly observe stock-outs, but we harness the higher likelihood of stock-outs due to hurricanes and their impact on consumer switching behavior to study the impact of product switches on subsequent choices.

In Section 1.4.3 we analyze whether the hurricanes affect other dimensions of choice behavior and find that consumers' choices during the hurricane are similar in terms of average product popularity and price level compared to products purchased prior to the hurricane. Consumers therefore do not appear to switch to more niche products or exhibit different sensitivity to price during the hurricane.

1.3 Conceptual Framework

In this section we show how we can use brand switching induced by hurricanes to identify a causal effect of past choices on current choices, i.e. *structural state dependence.* Contrary to other approaches in the literature, we do not estimate a model of consumer choice and instead base our analysis on the consequences of an underlying model of consumer choice (with or without structural

[17]In Section 1.4.4 we analyze the reasons underlying the observed switching behavior in more detail. Based on the timing of the expenditure increase and the subsequent brand switching behavior as well as a series of other data patterns, we conclude that stock-outs are the more likely driver of brand switching rather than different behavior when preparing for a hurricane relative to regular shopping trips.

state dependence) for the aggregate time-series pattern of persistence in consumers' choices. Going forward we will use the terms "structural state dependence" and "state dependence" interchangeably. We refer to "choice persistence" as the persistence observed in the data which could originate from either structural or spurious state dependence.

To provide intuition for our empirical analysis and identification strategy, we consider a simple model of consumer choice that allows for preference heterogeneity as well as structural state dependence. We use a set of simulations of consumer behavior based on this choice model to illustrate brand choice dynamics in steady state and to analyze brand choice patterns in reaction to a shock such as the hurricane-induced stock-outs that we study in our empirical application. We assume a consumer can choose from 3 products and the utility consumer i obtains when purchasing product j on a trip in week t is given by

$$u_{ijt} = \delta_{ij} + \gamma \times \mathbf{1}(j \text{ purchased on last trip}) + \varepsilon_{ijt},$$

where δ_{ij} denotes a consumer-specific product intercept. The second term captures structural state dependence by allowing utility to differ when product j was purchased on the previous shopping trip. Finally, ε_{ijt} is a standard normal taste shock that is independent across consumers, products, and time periods. For simplicity we do not explicitly model price, but consider price movements to be part of the error term ε_{ijt}. The population of consumers consists of 3 types (with equal share in the population) and each type prefers one of the three available products. For each consumer type, we set $\delta_{ij} = \delta^* > 0$ for the preferred product and $\delta_{ij} = 0$ for the other two. In the simulations below we analyze consumer choices when varying the degree of state dependence (γ) and preference heterogeneity. We capture changes in preference heterogeneity in a simple way by altering the difference in preferences for each consumer's preferred product (δ^*) relative to the other two products (whose intercepts are normalized to zero). The simulations are set up to mimic actual consumer behavior in our data.

In order to capture choice persistence around the hurricane shock we plot a modified measure of choice persistence that is given by:

$$\widetilde{Persist}_{it} = \begin{cases} \frac{1}{\#J_{it}} \sum_{j \in J_{it}} 1(j \in J_{it}^{pre-hurricane}) & \text{if "first trip after the hurricane"} \\ \frac{1}{\#J_{it}} \sum_{j \in J_{it}} 1(j \in J_{it}^{last}) & \text{otherwise.} \end{cases} \tag{1.2}$$

This measure of choice persistence is identical to the standard definition of choice persistence in equation (1.1) in most cases and measures whether on a given trip, the consumer purchases the same product she purchased previously. The modified measure differs from the standard definition only on the first purchase of a given household after the hurricane. On these trips, we compute choice persistence in reference to the last pre-hurricane purchase, i.e. we measure whether the product purchased on the first trip after the hurricane is identical to the product purchased on the last trip before the hurricane. As will become clear below, this modified variable makes it easier to analyze changes in behavior after the hurricane. All reported analyses use this modified choice persistence variable, and we refer to it as choice persistence and modified choice persistence interchangeably.

We start by plotting consumer behavior for a scenario with no structural state dependence in choice ($\gamma = 0$). We set $\delta^* = 1.67$ in order to generate a degree of choice persistence that is similar to the one in our data. We simulate behavior for a large set of consumers and arbitrarily set an initial condition for the first purchase and then simulate behavior for several weeks. The first 100 periods are discarded as burn-in and the next 52 weeks constitute the time window over which we study the evolution of the choice persistence variable. We assume that each consumer makes a choice in 43% of weeks to reflect the frequency with which we observe purchases in our data. To capture a stock-out effect similar to that observed in the data, we remove two randomly selected products from the choice sets of several consumers in the middle of the sample period (indicated by the vertical grey bars). We apply such a stock-out event to 25% of consumers, causing consumers to switch to available products that they may not have otherwise purchased.

The scenario without structural state dependence is illustrated by the closed dots in the top graph of Figure 1.3 and leads to an average choice persistence of around 0.65 in the pre-hurricane period (the left half of the graph). As a consequence of the stock-out, choice persistence decreases during the two affected weeks. In the absence of structural state dependence, the modified choice

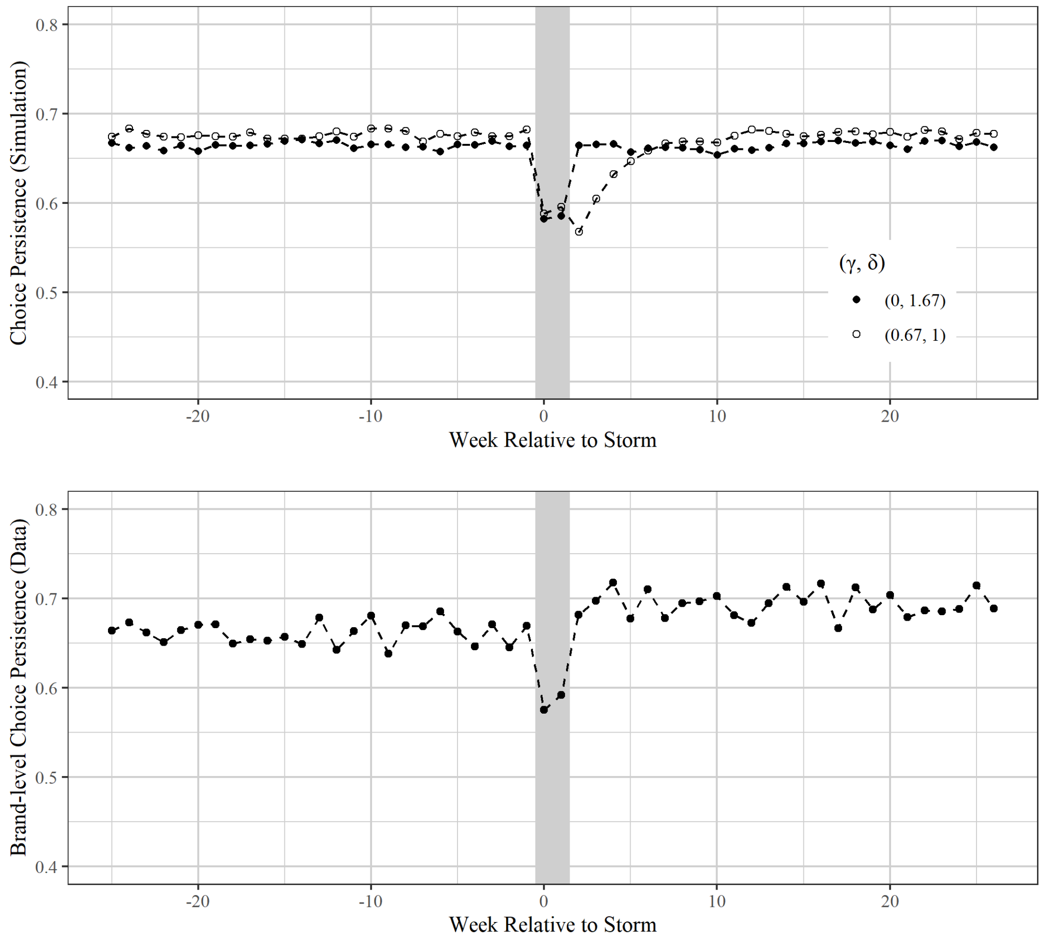


Figure 1.3: **Average Choice Persistence: Simulated Data and Empirical Patterns.** The top graph shows how average choice persistence evolves in response to a stock-out shock based on simulations with and without structural state dependence. The bottom graph plots average choice persistence in the data before and after a hurricane. In both graphs the vertical gray bars indicate weeks 0 and 1 which are affected by stock-outs.

persistence variable jumps back to its pre-hurricane level immediately. Without a causal effect of past choices, the product switches during the stock-out have no lasting impact and on the first trip after the hurricane, consumers' purchase probabilities and therefore average choice persistence are identical to their pre-stock-out values. We note that the use of the *modified* choice persistence variable is necessary to generate this pattern. When using a standard definition of choice persistence, the first trip after the hurricane would be characterized by lower persistence because the consumer has to "switch back" from the original switch during the hurricane.

Next, we analyze consumer behavior in the presence of structural state dependence by setting $\gamma = 0.67$ and $\delta^* = 1$. Structural state dependence coupled with a lower degree of preference heterogeneity generates a similar level of choice persistence in the pre-hurricane period as the scenario without state dependence discussed in the previous paragraph. The identical patterns of choice persistence illustrates the fundamental problem of identifying structural state dependence (Heckman (1981)): different combinations of preference heterogeneity and structural state dependence can generate identical patterns in observed choice persistence and therefore data on persistence in choices is not sufficient to identify structural state dependence separately from heterogeneity in preferences. The key idea of our identification strategy is that behavior in reaction to a shock to purchase behavior is different in the presence of structural state dependence. As the open dotted line in the top graph of Figure 1.3 shows, choice persistence decreases during the stock-out and then stays at a lower level for several weeks after the stock-out before slowly converging back to the pre-stock-out level. Contrary to the scenario without structural state dependence illustrated by the closed dot line, switches during the stock-out have an impact on choices beyond the period of stock-out.

An important aspect of the comparison of a scenario with and without structural state dependence is that these scenarios behave differently in the short-run after an external shock. However, in the long-run, the effect of the shock will dissipate even in the presence of structural state dependence and the lines corresponding to choice persistence in the two scenarios in Figure 1.3 therefore eventually converge. This insight informs our empirical analysis below, where we focus on the short-term impact of the hurricane on choice persistence to test for structural state dependence. If (modified) choice persistence jumps back to its pre-hurricane level immediately after the hurricane, such a behavior would suggest an absence of structural state dependence. We therefore take the equality of pre-hurricane and immediately post-hurricane choice persistence as our null hypothesis that corresponds to a model of consumer behavior without structural state dependence. We then test whether we can reject this null hypothesis, which would allow us to conclude that there is structural state dependence in consumers' choices.

In Appendix 1.7.1 we present an additional simulation based on a more realistic data-generating process. Specifically, we simulate data based on the estimates from a discrete choice model with structural state dependence that we implement in Section 1.5 based on bottled water data. Contrary to the simulations described above, this additional simulation is based on a continuous distribution of preference heterogeneity, allows for heterogeneity in the state dependence parameter, and includes price in the utility function. We find that when exposing consumers with such preferences to a stock-out shock of equal size as the one in our data, the post stock-out pattern of choice persistence looks very similar to the one for the setting with structural state dependence represented by the open dotted line in the top graph of Figure 1.3.

1.3.1 Identifying Assumptions

The basic idea behind our empirical test is the fact that under the null hypothesis of no structural state dependence, consumers' choices are independent across time periods. Therefore, the distribution of choice shares for each consumer is the same in each period and choice persistence at the consumer level is given by $Pr(choice_t = choice_{t'}) = \sum_j Pr_i(j)^2$, where $Pr_i(j)$ denotes the single-period choice probability of consumer i for product j, which is identical for any pair of periods t and t'. Based on this reasoning, choice persistence when comparing the first trip after the hurricane to the last trip before the hurricane will be identical to choice persistence between any of the pre-hurricane periods. This equality of choice persistence holds for each consumer and hence also holds for the average value of choice persistence. It follows that if average choice persistence reverts back to its pre-hurricane level immediately after the hurricane, we should conclude that choices in different time periods are independent.

This property of choice behavior holds regardless of the distribution of preference heterogeneity across consumers. An immediate reversion to pre-hurricane choice persistence therefore establishes an absence of structural state dependence regardless of how preferences are distributed in the population. We also assume that average choice persistence reflects consumers' choices in steady state and therefore our framework does not require us to explicitly account for consumers' initial conditions. Modeling preference heterogeneity in a sufficiently flexible fashion and accounting for initial conditions is typically required when estimating structural state dependence based on a discrete choice model of demand (e.g. Keane (1997), Dubé et al. (2008), Dubé et al. (2010a), Simonov et al. (2020)). We return to a more detailed comparison to this alternative approach in Section 1.5.

We also note that our approach is also related to an older literature on state dependence that tests for "zero-order" choice behavior at the individual consumer level (Frank (1962), Massy (1966), Bass et al. (1984)). Our approach similarly tests for a zero-order choice process, i.e. independent choices in different time periods, but does so by analyzing how choice persistence reacts to a stock-out shock.

1.3.2 First Look at the Data

We plot out average weekly choice persistence in the bottled water category over the one year time horizon surrounding a hurricane in the bottom graph in Figure 1.3. This graph shows that the observed choice persistence pattern does not exhibit any short-term change after the hurricane event. Instead, choice persistence appears to revert to its pre-hurricane value immediately after the storm.[18] The empirical patterns therefore look similar to the simulated patterns displayed in the top graph for a scenario without structural state dependence. The lower graph of Figure 1.2 that plots the share of new brands (defined relative to the brands purchased in a six month period preceding the main sample) purchased in each week tells a similar story: In weeks 0 and 1 consumers buy a larger share of products that they did not previously purchase. However, those choices are not persistent and the share of new brands decreases back to its pre-hurricane level immediately after the hurricane.

1.4 Empirical Analysis

Our empirical analysis closely follows the framework laid out in the previous section and analyzes the time series of choice persistence before, during, and after the hurricane. Because the time series pattern of choice persistence exhibits a small amount of seasonal fluctuation,[19] we add data from control households that are unaffected by the hurricane and employ a synthetic control approach. We also present estimates from a two-way fixed effect model which yields very similar results. However, because treated and control households deviate slightly in their pre-hurricane trends,[20]

[18] Visual inspection suggests a small increase in average choice persistence in the post-hurricane period. This likely relates to seasonal fluctuations in demand for bottled water as shown in Figure 1.2. We also observe a slight decrease in discounts on bottled water in the second half of our sample period, which might lead to an increase in choice persistence.

[19] Although our sample is not based on calendar time because the data contains households affected by hurricanes at different points in the year, some seasonality is nevertheless likely to affect our data. As shown in Table 1.1, most hurricanes occur in a similar time period of the year, usually around September and hence many observations are centered around this time of the year.

[20] Recall from Section 1.2.1 that by construction the treated and control groups consist of households that live in different geographic regions. It is therefore not unreasonable to expect seasonal trends that may lead to different patterns of choice persistence, i.e. demand for bottled water in Florida is higher in the winter months than in states with colder climates. In Appendix 1.7.2 we analyze time trends in the treatment and control group in detail.

the synthetic control approach constitutes our preferred specification.

The goal of our estimation approach is to analyze whether choice persistence reverts back to its pre-hurricane value immediately after the hurricane or whether it displays a gradual adjustment pattern over time. As outlined in Section 1.3, studying these adjustment patterns allows us to test for the presence of structural state dependence. We first outline the synthetic control method and present results for this preferred specification. We then proceed to a set of robustness checks in Section 1.4.2, and an analysis of which products consumers switch to and whether they alter other aspects of their behavior in Section 1.4.3. Finally, we analyze possible other ways in which hurricanes can impact consumers apart from stock-outs in Section 1.4.4.

1.4.1 Synthetic Control Method

We use the generalized synthetic control method proposed by Xu (2017) to impute counterfactuals for treated units. This method imputes the counterfactual evolution of the outcome variable based on an interactive fixed effects model (Bai, 2009). Specifically, we assume the following estimation equation:

$$\widetilde{Persist}_{it} = \delta_{it} D_{it} + \alpha_t + \boldsymbol{\lambda}_i' \boldsymbol{f}_t + \epsilon_{it} \tag{1.3}$$

where week fixed effects are represented by α_t and $\boldsymbol{f}_t$ is an $r \times 1$ vector of unobserved factors common across units in week t, where r is determined by cross-validation. The unobserved factors are weighted by an $r \times 1$ vector of factor loadings $\boldsymbol{\lambda}_i$ specific to unit i. Idiosyncratic shocks to unit i in week t are represented by ϵ_{it}. The treatment indicator D_{it} is equal to 1 if household i is part of the treated group and if the trip made in week t is during or after the hurricane. The effect of the treatment on the treated unit i in week t is represented by δ_{it}. The functional form in equation (1.3) assumes that both treated and control units are affected by the same set and number of unobserved factors.[21] In order to identify the causal treatment effects δ_{it} we require ϵ_{it} to be independent of D_{it}, α_t and $\boldsymbol{f}_t$.[22]

[21] Note that equation (1.3) nests the two-way fixed effects model when the model includes one factor that is equal to 1 for all t. In this case, the fixed effect structure is equal to a week and a household fixed effect.

[22] The model also requires weak serial dependence of error terms and a set of regularity conditions (see Xu (2017) for details). Moreover, the assumption that error terms are cross-sectionally independent and homoscedastic is needed for valid inference based on a block bootstrap procedure.

Estimation proceeds in three steps. First, we use only control units to estimate α_t, $\boldsymbol{f}_t$, and $\boldsymbol{\lambda}_i$ for all control units. Second, given estimates $\hat{\alpha}_t$ and $\hat{\boldsymbol{f}}_t$, we use pre-hurricane data for all treated units to estimate factor loadings $\boldsymbol{\lambda}_i$ in the treatment group. Finally, we construct a synthetic control observation for each treated unit by applying the estimates of $\hat{\alpha}_t$ and $\hat{\boldsymbol{f}}_t$ from the first step and the estimated factor loadings for treated units $\hat{\boldsymbol{\lambda}}_i$ from the second step and plugging them into the interactive fixed effect model:

$$\widehat{Persist}_{it}(0) = \hat{\alpha}_t + \hat{\boldsymbol{\lambda}}_i{}' \hat{\boldsymbol{f}}_t \tag{1.4}$$

where $\widehat{Persist}_{it}(0)$ denotes the counterfactual choice persistence value for treated unit i in time period t in the absence of treatment. This framework allows us to estimate the treatment effect for each household i and week t as the difference between the observed value of the choice persistence variable and its counterfactual value. We can then recover the average treatment effect on the treated (ATT) by taking the average of this difference across households in each period of the sample. We compute standard errors based on a non-parametric block-bootstrap, where we sample treated units with replacement from the data. Throughout the paper we report significance levels and confidence intervals based directly on the boostrap draws and not on normal approximations.

In our setting, we are particularly interested in the treatment effect for the weeks immediately after the hurricane, because these weeks capture consumers' first purchases after the stock-out forced them to switch brands. To analyze behavior after a hurricane, we start by displaying the full time series of average choice persistence for treated units and for the synthetic controls in Figure 1.4. We find that choice persistence in the treatment group decreases in weeks 0 and 1 relative to the control group. However, after the hurricane, choice persistence in the treatment group immediately reverts back to its counterfactual time trend given by the synthetic control group. The pattern is similar at the brand level and the UPC level which are displayed in the top and bottom graph respectively.[23] As outlined in Section 1.3, in the presence of state dependence, persistence would transition gradually back to its steady state level whereas in the absence of state dependence, persistence will revert back immediately. The graphs in Figure 1.4 therefore suggest that consumers' choices do not exhibit structural state dependence.

[23]We re-iterate that our analysis uses the modified persistence measure defined in equation 1.2, which defines persistence on the first purchase after the hurricane in relation to the last purchase before the hurricane.

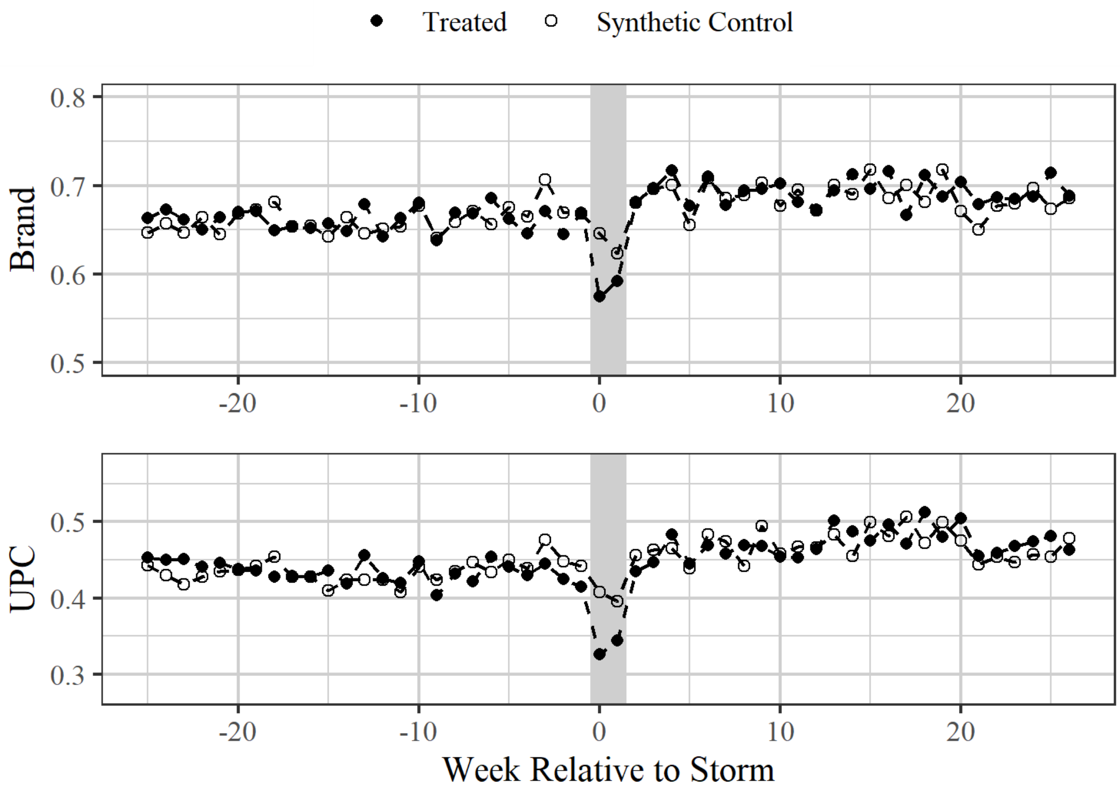


Figure 1.4: **Average Choice Persistence: Treatment Group and Synthetic Control.** The graphs display average choice persistence at the brand- and UPC-level. Closed and open dots represent choice persistence in the treatment group and the synthetic control value respectively. The vertical gray bars indicate weeks 0 and 1 which are affected by stock-outs.

Next, in order to quantify the statistical precision of these results, we report the treatment effect with its corresponding standard error for the weeks immediately after the hurricane. We focus on choice persistence during weeks 2 to 5 because simulations based on estimates from a structural model with state dependence (see Section 1.5 and Appendix 1.7.1) suggest that persistence will remain below its steady-state level for about 4 weeks following a shock like the one in our data. In column (1) of Table 1.3 we report the pooled effect for weeks 2 to 5 and find that it is not statistically significant and the point estimate takes on a small positive value. The decrease in choice persistence during the hurricane of -0.071 is relatively large compared to the impact immediately after the hurricane, which even when evaluated at the lower bound of the 95% confidence interval is equal to only -0.010. In column (2) we decompose the post-hurricane effect at the weekly level. All weekly effects are small in magnitude and have a positive sign. We cannot reject the null hypothesis that all 4 weekly differences are equal to zero.

			Robustness Checks		
	(1)	(2)	(3)	(4)	(5)
Level of Aggregation	Brand	Brand	UPC	Brand	Brand
Estimation Approach	Synthetic Control	Synthetic Control	Synthetic Control	Two-way Fixed Effects	Two-way Fixed Effects
Dependent Variable	$\widetilde{Persist}$	$\widetilde{Persist}$	$\widetilde{Persist}$	$\widetilde{Persist}$	$\widetilde{Persist}$
Week 0	-0.071***	-0.071***	-0.081***	-0.077***	-0.073***
	(0.018)	(0.018)	(0.017)	(0.014)	(0.015)
Week 1	-0.032*	-0.032*	-0.051***	-0.042***	-0.037**
	(0.019)	(0.019)	(0.017)	(0.016)	(0.010)
Weeks 2 -5	0.010		-0.004	0.001	0.006
	(0.012)		(0.012)	(0.011)	(0.012)
Week 2		0.001			
		(0.021)			
Week 3		0.001			
		(0.019)			
Week 4		0.017			
		(0.019)			
Week 5		0.022			
		(0.019)			
Weeks 6-26	0.004	0.004	0.003	-0.006	-0.005
	(0.007)	(0.007)	(0.007)	(0.007)	(0.013)
Quadratic Time Trend (interacted with Treatment)	n/a	n/a	n/a	No	Yes
Treated Observations	38,044	38,044	38,044	38,044	38,044
Treated Households	2,201	2,201	2,201	2,201	2,201
Control Observations	67,982	67,982	67,982	67,982	67,982
Control Households	3,866	3,866	3,866	3,866	3,866

Table 1.3: **Average Treatment Effect across Weeks.** Columns (1) to (3) report average treatment effects for specific weeks (or groups of weeks) based on the generalized synthetic control method. Standard errors and significance levels in columns (1) to (3) are based on 500 bootstrap samples. Significance levels are calculated based on the distribution of bootstrap estimates and not based on a normal approximation. Columns (4) and (5) report coefficients on the interaction of time period dummies with treatment status. Significance codes: *p<0.1; **p<0.05; ***p<0.01.

1.4.2 Robustness Checks

As a first robustness check, we replicate the brand-level specification in column (1) at the UPC level in column (3). We find that results are broadly similar. The observed decrease in choice persistence during the hurricane is slightly larger at the UPC-level and we do not observe a significant difference between treatment and synthetic control in the long-run. Most importantly, we observe no significant difference in choice persistence in weeks 2 to 5. When decomposing the effect at the weekly level (not reported in the table), we find no significant effect for any of the four weeks and we are not able to reject the null hypothesis that all weekly differences in weeks 2 to 5 are equal to zero. The absence of a post-hurricane effect on choice persistence at the UPC-level is also visible in the lower graph in Figure 1.4 which plots choice persistence in the treatment group and the synthetic control.

In the final two columns of Table 1.3 we report results from two regression specifications that include full sets of time period and household fixed effects. The two-way fixed effect model is specified as follows:

$$\begin{aligned}\widetilde{Persist}_{it} &= \bar{\beta}_i + \bar{\gamma}_t \\ &\quad + \mathbf{1}(Treated_i = 1) \times [\beta_0 \times \mathbf{1}(Week = 0) + \beta_1 \times \mathbf{1}(Week = 1) \\ &\quad + \beta_{2-5} \times \mathbf{1}(2 \leq Week \leq 5) + \beta_{6+} \times \mathbf{1}(Week \geq 6)] + \mu_{it},\end{aligned}$$

where $\mathbf{1}(Treated_i = 1)$ denotes a dummy that is equal to one for a consumer in the treatment group. Consumer and week fixed effects are denoted by $\bar{\beta}_i$ and $\bar{\gamma}_t$ respectively. The impact of the hurricane in the during / short-run / long-run period represent differences in behavior in the treatment group relative to the control group: β_0 and β_1 capture the immediate impact of the hurricane shock on choice persistence, and β_{2-5} and β_{6+} measure the short-run and long-run impact of the hurricane on choice persistence. The error term is denoted by μ_{it}. We cluster standard errors at the household level.

We show in Appendix 1.7.2 that the trends in persistence diverge between treatment and control group. As an additional robustness check we therefore report a version of the two-way fixed effect model that also includes an interaction of treatment with a quadratic time trend in column (5) of Table 1.3.[24]

[24] As we show in Appendix 1.7.2, the differential evolution in persistence between treatment and control is characterized by a gap that first slowly widens and then closes towards to end of the sample period. We therefore believe that a quadratic differential time trend constitutes a reasonable functional form to correct for the difference in time

Results from both specifications are very similar to the synthetic control results. Both regressions show a significant decrease in choice persistence during the hurricane and we do not find a significant impact on choice persistence in the weeks immediately after the hurricane in either of the two specifications. The estimated coefficients in both regressions are similar in magnitude to the treatment effects estimated in our synthetic control specification. When we include a quadratic time trend (interacted with treatment status) in order to remedy the diverging pre-trends in the treatment and control group, the estimated coefficients of the two-way fixed effect model become more similar to the synthetic control estimates.

In our final robustness check, we implement an analysis that only analyzes the first choice made after the hurricane by a given household regardless of when the first purchase in the category occurs. In particular, we compare choice persistence on the last trip of a given household prior to the hurricane with the first trip after the hurricane. We then test whether average choice persistence before the hurricane is significantly different from the average (modified) choice persistence variable on the first trip after the hurricane. The key idea of this test is the same as the one underpinning the synthetic control approach: in the absence of structural state dependence consumers will revert back to their pre-hurricane behavior immediately, whereas structural state dependence will cause a decrease in the choice persistence variable on the first trip after the hurricane relative to the last trip before the hurricane. Contrary to the analysis presented in Table 1.3, this additional test is based on a balanced panel of consumers and focuses specifically on the short-run effect on the first trip after the hurricane. In our earlier analysis of the time series of choice persistence, the composition of consumers in each week changed due to different purchase frequencies across consumers.

Table 1.4 reports results for the comparison just outlined based on a panel of all consumers that purchased at least once in the 4 weeks before and the 4 weeks after the hurricane, and also made at least one purchase during the hurricane. We choose a four week window to roughly replicate the 4-week window used in Table 1.3 to define the time-period shortly after the hurricane. As we discuss in more detail below, our results are robust over a range of alternative choices for the window in which we need to observed a purchase in order for a household to be included.

Before turning to consumer behavior after the hurricane, we first analyze the change in choice persistence that is caused by the hurricanes. In the second row of the table, we compare choice

trends.

		Average $\widetilde{Persist}$	Diff. in Means	S.E.
Brand-level				
Last Trip Before Hurricane	0.661			
First Trip During Hurricane		0.600	-0.061***	(0.014)
First Trip After Hurricane		0.682	0.021	(0.013)
UPC-level				
Last Trip Before Hurricane	0.424			
First Trip During Hurricane		0.349	-0.075***	(0.014)
First Trip After Hurricane		0.435	0.011	(0.014)
Observations (Households)	1,430			

Table 1.4: **Choice Persistence Comparison Before versus After a Hurricane.** The analysis in this table is based on all consumers that purchased bottled water at least once during the hurricane period as well as once in the 4 weeks before and after the hurricane. Significance codes: *p<0.1; **p<0.05; ***p<0.01.

persistence on the first trip during the hurricane to choice persistence in the last trip before the hurricane. We find that at the brand-level choice persistence drops from 0.661 to 0.600 and the change is statistically significant. The next row of the table provides our primary piece of analysis: here we compare choice persistence before the hurricane to the modified choice persistence measure on the first trip after the hurricane. We re-iterate that the modified measure calculates choice persistence in reference to the last trip before the hurricane. In the absence of structural state dependence we would expect the two measures of choice persistence to be identical. Our results show that we cannot reject the null hypothesis of equal means across the two variables. Choice persistence is slightly larger after the hurricane, but the difference is not statistically significant. Even at the lower end of the 95-percent confidence interval, choice persistence post-hurricane is smaller by only $0.021 - 1.96 \times 0.013 = -0.005$. This difference is small relative to the decrease in choice persistence during the hurricane of -0.061.

Results at the UPC level are reported in the lower panel of Table 1.4 and are very similar to the brand-level results. We find that choice persistence decreases by a larger amount at the UPC-level and the change is statistically significant. Choice persistence post-hurricane is estimated to be slightly larger than pre-hurricane choice persistence, but the difference is not statistically significant.

In Table 1.9 in Appendix 1.7.3 we show that allowing for a larger or smaller window before and after the hurricane leads to similar results. Widening the window allows us to include additional households whose first post-hurricane purchase occurs later. However, a larger window is more likely to be affected by the small amount of seasonal fluctuation in choice persistence documented earlier. Specifically, we vary the time window in the before and after period between 1 and 10 weeks. We find that the pattern presented in Table 1.4 for a 4 week window holds consistently regardless of the width of the time window both at the brand- and the UPC-level. In all specifications we find a significant decrease in choice persistence during the hurricane and no statistically significant difference in choice persistence when comparing the last trip before the hurricane to the first trip after the hurricane.

1.4.3 Consumer Purchase Behavior During the Hurricane

In this section, we explore what types of products consumers tend to purchase during a hurricane and whether consumers alter their purchase behavior along other dimensions apart from an increase in brand switching. We start by analyzing how purchases during the hurricane differ from pre-hurricane purchases in terms of product popularity. To this end, we rank brands by their pre-hurricane market-share and calculate the change in purchase share during the hurricane relative to the pre-hurricane period. We plot the change in purchase share by brand in Figure 1.5. The top graph plots out the brand-level market-share before and during the hurricane, whereas the bottom graph plots the percentage change in market-share for each brand. We separately plot behavior for the top 17 brands that make up 90 percent of total market share. The right-most data-point in both graphs represents a residual category of all other brands that make up the bottom 10% of brands in terms of their market-share.[25] Taken together the two graphs show that switches do not exhibit any particular pattern in terms of popularity and pre-hurricane popularity does not appear to predict the change in purchase share during the hurricane.

Next, we explore changes in consumer behavior along a series of other dimensions by re-estimating our synthetic control specification using a series of different outcome variables. We first analyze changes in total expenditure during the hurricane in column (1) of Table 1.5 and find

[25]We treat all private label products as one brand in this analysis. Together they make up the largest purchase share, represented by the left-most points in Figure 1.5.

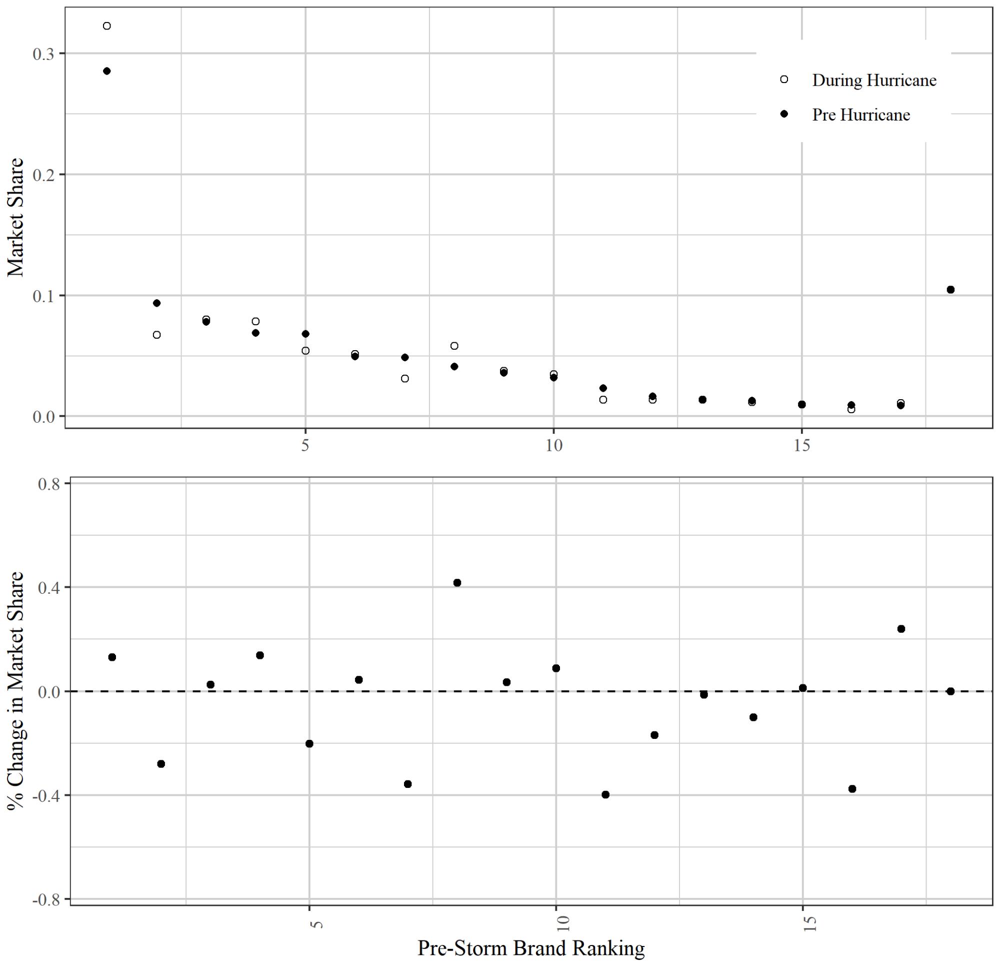


Figure 1.5: **Market Share of Top Brands Before and During the Hurricane.** The top graphs displays market-shares for the top 17 brands (ranked from largest the smallest) and a residual category of all other brands (the right-most points) before the hurricane (solid dots) and during the hurricane (open dots). The lower graph displays the percentage change in market-shares for each brand during the hurricane relative to the time period before the hurricane.

that expenditure increased significantly during the hurricane. We then decompose the expenditure effect into its price and quantity components in columns (2) and (3). We find that consumers purchase similar products in terms of their price level, but quantity purchased increased significantly. Finally, we analyze the number of unique brands purchased on a given trip in column (4) of Table 1.5 and find a small but significant increase in the number of brands purchased. The average number of unique brands purchased is equal to 1.18 in the pre-hurricane period and increases by 0.05 during week 0. We also note that we do not find evidence for changes in consumer behavior in the long-run along any of the outcomes analyzed in Table 1.5, a point that we will return to in

Dependent Variable	(1) Expenditure	(2) Price / Oz	(3) Ounces Purchased	(4) # Brands Purchased
Mean of DV (in the Pre-Hurricane Period)	2.30	0.02	537.57	1.18
Week 0	1.011***	0.017	110.27***	0.047***
	(0.296)	(0.030)	(16.15)	(0.018)
Week 1	-0.223	-0.000	34.29**	0.029**
	(0.466)	(0.024)	(16.93)	(0.015)
Weeks 2 -5	-0.012	0.012	-14.59	-0.003
	(0.216)	(0.018)	(10.03)	(0.011)
Week 6-26	-0.091	0.000	5.92	0.000
	(0.304)	(0.002)	(6.61)	(0.007)
Treated Observations	38,044	38,044	38,044	38,044
Treated Households	2,201	2,201	2,201	2,201

Table 1.5: **Impact of the Hurricane on Purchase Behavior.** All columns report average treatment effects for specific weeks (or groups of weeks) based on the generalized synthetic control method. Standard errors and significance levels are based on 500 bootstrap samples. Significance levels are calculated based on the distribution of bootstrap estimates and not based on a normal approximation. Significance codes: *p<0.1; **p<0.05; ***p<0.01.

the next sub-section.

Next, we analyze whether the unusual behavior in terms of purchase quantity and multi-brand purchases documented above might impact our results. For this purpose we use the synthetic control approach introduced in Section 1.4.1, and report differences between the observed weekly choice persistence and the counterfactual for specific subsets of households. In the first column of Table 1.6 we replicate our baseline results for the full sample as a benchmark. Columns (2) and (3) display results separately for households that purchased an above / below median quantity of bottled water during the hurricane.[26] We find that both groups of households behave similarly in terms of their choice persistence after the hurricane and for both groups we are not able to reject the null hypothesis of no structural state dependence. In columns (4) and (5) of Table 1.6 we

[26]The median split is based on all purchases made during the hurricane.

investigate whether the small number of households that buy multiple brands on the same shopping trip during the hurricane behave differently from households that purchase only one brand. The results from these regressions show that for both groups we do not find a significant change in choice persistence after the hurricane. In Appendix 1.7.4 we provide additional robustness checks related to multi-brand purchases.

Finally, we analyze behavior for the subset of households that purchased popular brands during the hurricane. In column (6) we select only households that purchase one of the top 10 brands and find that these households exhibit a similar decrease in choice persistence during the hurricane and post-hurricane choice persistence in the treatment group is not significantly different from choice persistence in the synthetic control. In column (7) we narrow the sample down further to households that purchased one of the top 5 brands during the hurricane and continue to find no change in choice persistence after the hurricane. We also analyze whether consumers behave differently when purchasing more or less expensive products by analyzing behavior separately for consumers that purchase above / below median price brands during the hurricane and find a null effect for both sub-groups.[27]

In summary, we conclude that consumers do not purchase unusual products in terms of their popularity or price during the hurricane and the null effect is not driven by subgroups with unusual purchase behavior during the hurricane such as purchases of niche products, bulk buying, or purchases of multiple brands on the same shopping trip.

1.4.4 Hurricanes & Other Channels of Impact

There are several ways in which hurricanes might affect consumer behavior apart from generating stock-outs that trigger brand switching. In this section we assess the evidence for possible other channels through which hurricanes impact consumers. One way to conceive of our identification strategy is that we would like consumers to switch brands because they face a stock-out on a particular store visit, but this stock-out does not correlate with any other factors that might impact demand. Because we rely on stock-outs induced by hurricanes we need to consider the possibility that the hurricane affects consumers in other ways.

[27] We find that the post-hurricane effect is not statistically significant for either group and the coefficient estimate (standard error) is equal to 0.023 (0.015) and -0.003 (0.015) for consumer with low- and high-price purchases respectively.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Dependent Variable	$\widetilde{Persist}$	$\widetilde{Persist}$	$\widetilde{Persist}$	$\widetilde{Persist}$	$\widetilde{Persist}$	$\widetilde{Persist}$	$\widetilde{Persist}$
Sample	Full Sample	$\geq$ Median Purchase Quantity	$<$ Median Purchase Quantity	Single Brand Purchased	Multiple Brands Purchased	Top 10 Products	Top 5 Products
Brand-level							
Week 0	-0.071***	-0.092***	-0.048***	-0.049***	-0.166*	-0.074***	-0.074***
	(0.018)	(0.023)	(0.022)	(0.019)	(0.035)	(0.018)	(0.019)
Week 1	-0.032*	-0.042	-0.021	-0.030	-0.039	-0.048	-0.040
	(0.019)	(0.025)	(0.024)	(0.020)	(0.037)	(0.020)	(0.022)
Weeks 2 -5	0.010	0.014	0.007	0.019	-0.021	0.010	0.009
	(0.012)	(0.015)	(0.015)	(0.013)	(0.023)	(0.012)	(0.013)
UPC-level							
Week 0	-0.081***	-0.108***	-0.051***	-0.074***	-0.112**	-0.090***	-0.091***
	(0.017)	(0.022)	(0.021)	(0.018)	(0.034)	(0.018)	(0.019)
Week 1	-0.051***	-0.070***	-0.033**	-0.055***	-0.038	-0.072***	-0.075***
	(0.017)	(0.022)	(0.023)	(0.018)	(0.036)	(0.018)	(0.019)
Weeks 2 -5	-0.004	-0.011	0.003	0.005	-0.036	-0.006	-0.010
	(0.012)	(0.016)	(0.014)	(0.012)	(0.024)	(0.012)	(0.013)
Treated Observations	38,044	19,136	18,908	30,132	7,912	33,684	28,068
Treated Households	2,201	1,056	1,145	1,814	387	1,939	1,581

Table 1.6: **Subgroup Analysis.** All columns report average treatment effects for specific weeks (or groups of weeks) based on the generalized synthetic control method. Standard errors and significance levels are based on 500 bootstrap samples. Significance levels are calculated based on the distribution of bootstrap estimates and not based on a normal approximation. Significance codes: *p<0.1; **p<0.05; ***p<0.01.

Longer-term Impact of Hurricane It is possible that hurricanes lead to longer term changes in behavior due to the general disruption and possible financial shocks associated with a hurricane. We note, however, that our analysis is based on consumers that were preparing for a hurricane, but many of those consumer were never affected or only mildly affected by the actual hurricane. Moreover, the most likely effect of a permanent financial shock due to a hurricane would be for consumers to permanently purchase a different brand (most likely a less expensive one). Therefore, the presence of long-term shocks might generate a permanent change in brand choice which one might then incorrectly attribute to structural state dependence. It is less likely that a long-term effect of a hurricane would cause us to falsely estimate a null effect with regards to structural state dependence.

We can test for the presence of longer-term changes in purchase behavior by analyzing how consumers' choices behave in the long run. Our main estimation results establish that choice persistence does not exhibit any long-run changes. Moreover, the results in columns (1) and (2) of Table 1.5 show that consumers did not alter their level of expenditure in the category nor did they become more price sensitive in the long-run. Finally, we re-run our main analysis based on a sub-sample of less severe hurricanes. In particular, we re-run our analysis excluding the two largest and most disruptive hurricanes and based only on hurricanes that generates less than 10 billion dollars in damage. Results from these regressions are reported in Table 1.10 in the appendix. We do not find that results based on these sub-samples of hurricanes are qualitatively different from our main results based on the full sample of households. We conclude that the hurricanes are unlikely to have lead to a longer-term financial impact on consumers.

Context-dependent Consumption Because our empirical strategy leverages an increase in brand-switching around the time of a hurricane, there are two possible explanations for why consumers switch brands. Either consumers face stock-outs and therefore need to switch to a different brand or consumers might perceive of a pre-storm shopping trip as a different context that leads them to purchase a different brand (even if their preferred brand is available). If the latter channel is driving the observed pattern, the finding of no structural state dependence might be specific to the hurricane shock we study and may not extrapolate to other drivers of brand switching such as price discounts.

For several reasons we believe it is more likely that consumers switch brands due stock-outs rather than due to a change in consumption context. First, we find that expenditure increases a week before we observe brand switching (see Figure 1.2). Therefore, while hurricane preparation occurs already in week -1, we don't observe an increase in brand switching until week 0. If context effects were important, we would instead expect brand switching to coincide with the increase in demand due to hurricane preparations. By contrast, stock-outs likely occur with a slight lag after a demand spike. Therefore, the fact that brand switching occurs one week after the initial demand spike is consistent with consumers switching brands due to stock-outs. Second, the most likely context specific type of brand switching would be to cheaper or lower quality niche products.

However, our findings in the previous section show that purchases during the hurricane are similar in terms of product popularity and price level. Third, we find that the null effect holds for households that did not purchase in bulk and were hence less likely to engage in purchase behavior specific to hurricane preparations. Taken together these data patterns provide evidence that brand switching is likely driven by stock-outs rather than context-specific purchase behavior.

1.5 Comparison to Structural Estimation Approach

Next, we compare our findings to the common approach of estimating a structural model of consumer choice that allows for a lagged-choice term in the utility function that captures structural state dependence. A series of papers (e.g. Dubé et al. (2008), Dubé et al. (2009), Dubé et al. (2010a), Simonov et al. (2020)) takes such an approach and they tend to find evidence for structural state dependence. In order to understand why the findings of these papers deviate from ours, we first estimate a discrete choice model that allows for structural state dependence on data from the bottled water category. For comparison, we also replicate the estimates from a choice model with state dependence based on margarine data in Simonov et al. (2020).

We follow the methodology in Simonov et al. (2020) and estimate a model that allows for a flexible distribution of heterogeneity and accounts for the initial condition. We also closely follow Simonov et al. (2020) in terms of how to construct the estimation samples for both categories.[28] Sample construction is somewhat involved, because we need to find households in the consumer data that visited stores that are present in the store level data-set. This overlap is required because we rely on the store data to construct price series for all available products. We further need to confine the analysis to the top brands in order to reliably construct prices series. We described the details of how we construct the estimation samples in Appendix 1.7.5. We also note that the set of households used to analyze behavior in the bottled water category in our main analysis is different from the households used in this section due to different sample selection criteria.[29]

We estimate a discrete choice model based on a utility function similar to the one used in Section 1.3 as the basis for illustrating our identification strategy. Specifically, we assume that the utility

[28] We thank the authors of Simonov et al. (2020) for sharing their code with us.

[29] For the estimation in this section we do not impose any geographic selection criteria as we do in our main analysis based on hurricane locations. Instead, we select households primarily based on whether they visit stores that are present in the store-level data and whether they purchase the top brands of water. Both criteria are imposed in order to obtain reliable price series.

for consumer i in time period t when purchasing product j is given by:

$$u_{ijt} = \delta_{ij} - \alpha_i p_{jt} + \gamma_i \times \mathbf{1}(j \text{ purchased on last trip}) + \varepsilon_{ijt}$$

where we allow for heterogeneity in brand intercepts δ_{ij}, the price coefficient α_i, and the state dependence term γ_i. The error term ε_{ijt} is extreme value type 1 distributed and independent across consumers, products, and time periods. We report results from this model based on data from the bottled water category as well as the replication of Simonov et al. (2020) using margarine data in Table 1.7. We find that the estimated mean of the state dependence parameter is similar between the two product categories, but slightly larger for the bottled water category. Moreover, the price coefficient is somewhat smaller in the bottled water category relative to margarine. Therefore, the monetized state dependence parameter is larger for bottled water.[30]

Next, to provide a direct comparison to our method of analyzing the time series of average choice persistence, we simulate consumer behavior in reaction to a stock-out based on the estimated parameters from the bottled water category. We follow the same template that we used for the simulations in Section 1.3 and we induce a stock-out shock that leads to a change in choice persistence that is exactly equal to the one observed in our data. We provide additional details on how this simulation is implemented in Appendix 1.7.1. In Table 1.8 we report our main estimation results from the synthetic control method and compare them against the values of choice persistence in the weeks following the stock-out shock that result from the simulation. We find that the simulated effect in weeks 2 to 5 based on the structural estimates is equal to -0.046 whereas the estimated effect is equal to 0.010 with a 95% confidence interval of $(-0.010, 0.033)$. The effect based on the data-generating process from the structural model therefore lies far outside of the confidence interval of our estimate and we can reject that the observed pattern of choice persistence in the bottled water category was generated by the estimates from the structural model. When we split the post stock-out effect into separate weekly effects in columns (3) and (4) we find that the simulated effect lies outside the respective confidence interval for all four weeks.

The results presented above show that our null results are not driven by our choice of category, because a discrete choice model does result in estimates of structural state dependence similar to

[30]In Appendix 1.7.5, we provide additional results for both categories.

		Water	Margarine
Brand 1	μ_{δ_1}	1.823	-2.004
		(1.153, 2.533)	(-2.161, -1.855)
	σ_{δ_1}	4.861	3.250
		(4.037, 5.748)	(3.065, 3.440)
Brand 2	μ_{δ_2}	1.642	0.207
		(0.979, 2.329)	(-0.007, 0.427)
	σ_{δ_2}	4.956	3.443
		(4.142, 5.869)	(3.176, 3.727)
Brand 3	μ_{δ_3}		-1.697
			(-1.823, -1.568)
	σ_{δ_3}		2.961
			(2.791, 3.141)
Brand 4	μ_{δ_4}		-1.088
			(-1.360, -0.809)
	σ_{δ_4}		4.027
			(3.731, 4.317)
Price	μ_{α}	-0.766	-1.146
		(-0.896, -0.645)	(-1.228, -1.063)
	σ_{α}	0.841	1.366
		(0.688, 1.010)	(1.26, 1.473)
State	μ_{γ}	1.233	0.987
Dependence		(0.943, 1.540)	(0.899, 1.075)
	σ_{γ}	1.471	1.113
		(1.173, 1.804)	(1.035, 1.204)
Obsservations		8,661	51,122
Households		272	2,232

Table 1.7: **Structural Estimation of State Dependence.** The estimates in this table are based on the method in Simonov et al. (2020) that corrects for consumers' initial condition and allows for a first-order Markov process in prices. 95% posterior credible intervals are reported in paranthesis.

those found in the prior literature. Moreover, the null effect is not driven by a lack of statistical power and the estimated post stock-out choice persistence patterns allow us to rule out state dependence effects of the magnitude implied by the structural estimates. We therefore conclude that the differences between our approach and the structural choice model approach likely originate from differences in methodology and the variation used to estimate structural state dependence. In the next sub-sections, we explore these differences in more detail.

Dependent Variable	(1) Synthetic Control Estimates $\widetilde{Persist}$	(2) Simulated Values (Structural Model DGP) $\widetilde{Persist}$	(3) Synthetic Control Estimates $\widetilde{Persist}$	(4) Simulated Values (Structural Model DGP) $\widetilde{Persist}$
Week 0	-0.071***	-0.071	-0.071***	-0.071
	(-0.107, -0.039)		(-0.107, -0.039)	
Week 1	-0.032*	-0.032	-0.032*	-0.032
	(-0.068, 0.006)		(-0.068, 0.006)	
Weeks 2 -5	0.010	-0.046		
	(-0.010, 0.033)			
Week 2			0.001	-0.090
			(-0.040, 0.041)	
Week 3			0.001	-0.053
			(-0.032, 0.042)	
Week 4			0.017	-0.026
			(-0.019, 0.055)	
Week 5			0.022	-0.018
			(-0.016, 0.059)	
Weeks 6-26	0.004	-0.002	0.004	-0.002
	(-0.010, 0.018)		(-0.010, 0.018)	
Treated Observations	38,044		38,044	
Treated Households	2,201		2,201	

Table 1.8: **Comparison of Estimates to Simulated Values (Based on Structural Model Estimates).** Columns (1) and (3) report average treatment effects for specific weeks (or groups of weeks) based on the generalized synthetic control method. 95% confidence intervals are reported in paranthesis. Confidence intervals and significance levels are based on 500 bootstrap samples and not based on a normal approximation. Columns (2) and (4) report the simulated values of choice persistence when using the estimates from the structural model as the data-generating process. Significance codes: *p<0.1; **p<0.05; ***p<0.01.

1.5.1 Price Variation vs. Stock-outs

The primary source of identification with regards to structural state dependence in prior work is often price variation over time (e.g., Dubé et al. (2010a)). Intuitively, a discount on a given shopping trip might make a consumer switch to the discounted product. In a world without state dependence, the consumer will revert to her pre-discount behavior on the next shopping trip when the price is back at its regular level. Instead, in the presence of structural state dependence, the consumer is likely to continue purchasing the product she switched to. Therefore, a causal effect of past prices on current behavior identifies structural state dependence (Chamberlain (1985)). In our setting, a stock-out (instead of a price discount) induces consumers to switch to a different product. Similar to the impact of price changes just described, consumers will revert back to their pre-hurricane behavior immediately in the absence of structural state dependence. In the presence of structural state dependence, consumers continue to purchase the product they switched to even after the hurricane.

The identifying assumption in both approaches (price- or hurricane-based) to identifying state dependence is that product switches are uncorrelated with product-specific demand shocks in the next period. For example, if a price discount for a specific product coincides with the start of an advertising campaign that lasts several weeks, then switches to the discounted product will be correlated with higher demand for the same product next period. Such a pattern of correlated choices could spuriously generate patterns that are incorrectly attributed to structural state dependence. This kind of pattern is less likely in the context of switches due to hurricane stock-outs. In particular, it is unlikely that demand for non-stocked-out products, i.e., the products that consumers switch to during the hurricane, is systematically higher or lower in the post-hurricane period for households in our sample. Hurricanes, of course, do not occur in reaction to demand shocks, and advertising and pricing schedules are unlikely to change in response to a hurricane.

1.5.2 Estimation Framework and Identifying Assumptions

Contrary to a structural model of demand with state dependence, the estimation framework presented in this paper requires fewer assumptions. Our approach of analyzing the time series of average choice persistence allows us to derive a test for structural state dependence that does not depend on the distribution of heterogeneity and does not require us to estimate that distribution. Instead, our test only relies on the independence of choices over time when consumers do not exhibit structural state dependence. More generally, our estimation approach requires fewer functional form assumptions and, apart from assumptions regarding the distribution of heterogeneity, we also do not need to specify the distribution of the error terms entering utility (typically assumed to be extreme value type 1 distributed). Being able to avoid functional form assumptions with regards to different components of preferences constitutes an important advantage of our approach because restrictive functional form assumptions can lead to spurious results with regards to structural state dependence (Dubé et al. (2010a), Paulson (2012)). Moreover, our approach does not require us to model a consumer's initial condition which can lead to biased estimates of state dependence if not handled correctly (see Simonov et al. (2020)). Instead, our approach is based on the assumption that average choice persistence prior to a hurricane reflects consumers' steady state behavior. In summary, our approach is less likely to be affected by model mis-specification that arises from the way in which preference heterogeneity and the initial condition are handled in estimation.

1.5.3 Other Differences

We re-iterate that a series of other differences that we discussed in Sections 1.4.3 and 1.4.4 can be reasonably ruled out as drivers behind our null effect. In particular, we rule out that our findings are driven by unusual purchase behavior during a hurricane such as purchases of niche products, bulk buying, and purchases of multiple brands (see Section 1.4.3) or disruptive effects of hurricanes that directly impact consumers' purchase behavior (see Section 1.4.4). A final reason is that hurricanes might trigger context-specific purchase behavior and therefore we do not see a lasting effect of brand switches during the hurricane. As we explain in Section 1.4.4, a series of data patterns such as the timing of brand switches and the absence of switches to lower popularity and cheaper products are at odds with a context-specific interpretation of our results.

1.6 Conclusion

In this paper, we propose a simple test for structural state dependence based on the evolution of the time-series of average choice persistence following an exogenous shock. We apply our framework to panel data in the bottled water category and exploit stock-outs induced by hurricanes as an exogenous shock to consumers' purchase decisions. We are unable to reject the null hypothesis of no structural state dependence using our estimation framework, but find a positive and significant state dependence effect when estimating a choice model with state dependence on data from the same category. We show that our approach does not lack statistical power and provide evidence against any direct impact of hurricanes on purchase behavior (other than through stock-outs). We argue that our approach is better suited to identify the causal impact of past choices because it requires fewer assumptions and is based on more plausibly exogenous variation in brand switching due to stock-outs rather than price discounts.

1.7 Appendix

1.7.1 Additional Simulations

In this section we provide additional details on the simulation of choice persistence that uses estimated preference parameters from the structural choice model presented in Section 1.5.

Utility for consumer i in time period t when purchasing product j is given by

$$u_{ijt} = \delta_{ij} - \alpha_i p_{jt} + \gamma_i \times \mathbf{1}(j \text{ purchased on last trip}) + \varepsilon_{ijt},$$

where ε_{ijt} is extreme value type 1 distributed and preference parameters are distributed according to the estimated distribution of preference parameters in the bottled water category (see Section 1.5):

$$\begin{bmatrix} \delta_{i1} \\ \delta_{i2} \\ \alpha_i \\ \gamma_i \end{bmatrix} \sim N \left(\begin{bmatrix} 1.82 \\ 1.64 \\ -0.77 \\ 1.23 \end{bmatrix} \begin{bmatrix} 23.63 & 0 & 0 & 0 \\ 0 & 24.56 & 0 & 0 \\ 0 & 0 & 0.71 & 0 \\ 0 & 0 & 0 & 2.16 \end{bmatrix} \right)$$

We assume that prices follow a process similar to the one observed in the data. For each product we set the regular price to the modal price during our sample period. In terms of discount frequency and depth we assume that each product is discounted by 20% in 15% of weeks and price discounts are iid across products, weeks, and consumers (because different consumers shop in different stores). We assume that consumers purchase in 43% of weeks (which corresponds to the purchase frequency in our data) and simulate behavior over a 52 week period with a simulated stock-out in the middle of the sample period. We choose the size of the stock-out shock so that the decrease in choice persistence matches the magnitude of the decrease in our data. In particular, we randomly remove one brand for X% of consumers in week 0 and week 1 and choose "X" such that the change in choice persistence in those two weeks matches the one in our data.[31]

[31] We assume an equal number of consumers shop in week 0 and week 1. We note that the first post-hurricane observation in week 2 includes many consumers that purchased in week 0 (but not week 1) and therefore choice persistence in week 2 is lower than the value in week 1, as these consumers were more likely to be exposed to a stock-out.

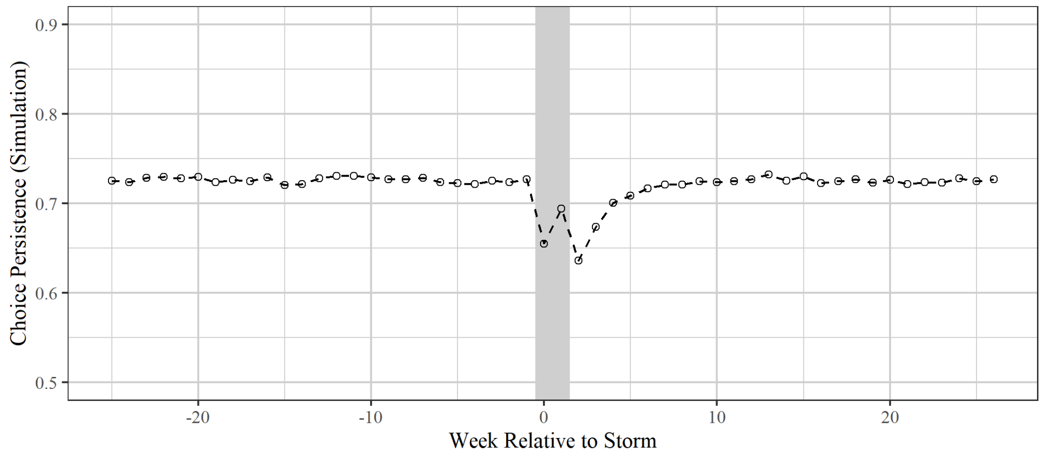


Figure 1.6: **Average Choice Persistence: Simulated Data with Preferences from Structural Model Estimates.** The vertical gray bar indicates weeks 0 and 1 which are affected by stock-outs.

Figure 1.6 shows the resulting pattern of choice persistence based on the utility function and preference distribution specified above. We find that the simulation shows a clear transition pattern in choice persistence after the hurricane, similar to the one in our earlier simulations in Section 1.3. In particular, it takes roughly 4 weeks for choice persistence to revert back to its pre-hurricane level following the stock-out shock in weeks 0 and 1. This simulation is also used to generate the values of choice persistence (conditional on the data-generating process being given by the parameter estimates above) in Table 1.8.

1.7.2 Differential Time Trends in Treatment and Control Group

To explore differential behavior in the treatment and control group over time we run a fixed effect regression with separate weekly coefficients for the treatment and control group and then plot out the estimated time-trends for both groups. Specifically, we implement the following regression:

$$\begin{aligned} \widetilde{Persist}_{it} &= \sum_{t=-24}^{t=26} \alpha_t \times \mathbf{1}(Week = t) \times Treat_i \\ &+ \sum_{t=-24}^{t=26} \beta_t \times \mathbf{1}(Week = t) \times (1 - Treat_i) \\ &+ \gamma_i + \varepsilon_{it} \end{aligned} \tag{1.5}$$

We plot out the estimated treatment group (α_t) and control group (β_t) coefficients across the 52 weeks of our sample (minus the first week which constitutes the omitted category for both time series) in figure 1.7. At both the brand- and the UPC-level we observe a time trend in the treatment group with lower values of choice persistence in the middle of the sample period. Moreover, trends in the treatment and control group do not exactly match each other in the pre-treatment period. Based on this initial analysis of choice persistence in the treatment and control group, we conclude the pre-treatment trends differ between treatment and control group in the bottled water category. This finding is the primary motivation for our use of the generalized synthetic control method in Section 1.4.1. We also note that the difference in trends roughly follows a U-shape where the gap between treatment and control first widens and then closes again towards the end of the sample period. This pattern in the data informs one of our robustness checks where we include a quadratic time trend (interacted with treatment status) in a two-way fixed effect model with household and time period fixed effects.

1.7.3 Persistence Comparison with Varying Time-Window

In this section we report additional results for the analysis that compares persistence between the first purchase after the hurricane relative to the last purchase before the hurricane. In Table 1.4 we reported results based on all households that purchased at least once in the 4 weeks before and the 4 weeks after the hurricane. In Table 1.9 we report additional results when varying the time window between 1 and 10 weeks. All specifications are based on a balanced panel of households and only include the last purchase before the hurricane and the first purchase after the hurricane for each household. Widening the window increases the number of households for which we observe at least one purchase in the pre- and post-hurricane periods.

1.7.4 Additional Robustness Checks

In this section we discuss the results from a series of additional regressions reported in Table 1.10. All of the regressions in the table are based on the generalized synthetic control method we use as our primary specification. We replicate our baseline results for the full sample (column (1) of Table 1.3) as a benchmark in column (1) of Table 1.10. The remaining columns present results from regressions that either change the outcome variable or focus on specific subsets of households.

	Before / During				Before / After				
	Before	During	Diff.	SE	Before	After	Diff.	SE	# HHs
Brand-level									
10 weeks	0.623	0.571	-0.052***	(0.011)	0.623	0.629	0.006	(0.011)	2,252
9 weeks	0.629	0.575	-0.053***	(0.011)	0.629	0.634	0.005	(0.011)	2,178
8 weeks	0.633	0.576	-0.057***	(0.012)	0.633	0.640	0.006	(0.011)	2,063
7 weeks	0.640	0.577	-0.063***	(0.012)	0.640	0.648	0.009	(0.011)	1,945
6 weeks	0.644	0.581	-0.063***	(0.012)	0.644	0.655	0.011	(0.012)	1,818
5 weeks	0.652	0.589	-0.063***	(0.013)	0.652	0.668	0.016	(0.012)	1,653
4 weeks	0.661	0.600	-0.061***	(0.014)	0.661	0.682	0.021	(0.013)	1,430
3 weeks	0.683	0.636	-0.047***	(0.015)	0.683	0.696	0.013	(0.014)	1,113
2 weeks	0.703	0.659	-0.045**	(0.018)	0.703	0.698	-0.005	(0.017)	774
1 week	0.736	0.668	-0.068**	(0.028)	0.736	0.722	-0.015	(0.026)	307
UPC-level									
10 weeks	0.385	0.324	-0.061***	(0.011)	0.385	0.379	-0.006	(0.011)	2,252
9 weeks	0.390	0.329	-0.061***	(0.011)	0.390	0.383	-0.007	(0.011)	2,178
8 weeks	0.398	0.334	-0.064***	(0.012)	0.398	0.390	-0.008	(0.011)	2,063
7 weeks	0.407	0.338	-0.069***	(0.012)	0.407	0.402	-0.005	(0.012)	1,945
6 weeks	0.413	0.340	-0.073***	(0.012)	0.413	0.412	-0.001	(0.012)	1,818
5 weeks	0.418	0.344	-0.074***	(0.013)	0.418	0.421	0.003	(0.013)	1,653
4 weeks	0.424	0.349	-0.075***	(0.014)	0.424	0.435	0.011	(0.014)	1,430
3 weeks	0.445	0.372	-0.074***	(0.016)	0.445	0.446	0.001	(0.015)	1,113
2 weeks	0.457	0.388	-0.069***	(0.019)	0.457	0.457	0.000	(0.019)	774
1 week	0.481	0.414	-0.067**	(0.031)	0.481	0.478	-0.003	(0.029)	307

Table 1.9: **Robustness Check: Choice Persistence Before / After Comparison with Varying Time Windows.** Each row reports results from a balanced panel of consumers that purchased at least once during the hurricane and once during a specific number of weeks before and after the hurricane. The number of weeks used to define choice persistence before and after the hurricane varies across rows. For each consumer we only use the last purchase before the hurricane and the first purchase during and after the hurricane. Significance codes: *p<0.1; **p<0.05; ***p<0.01.

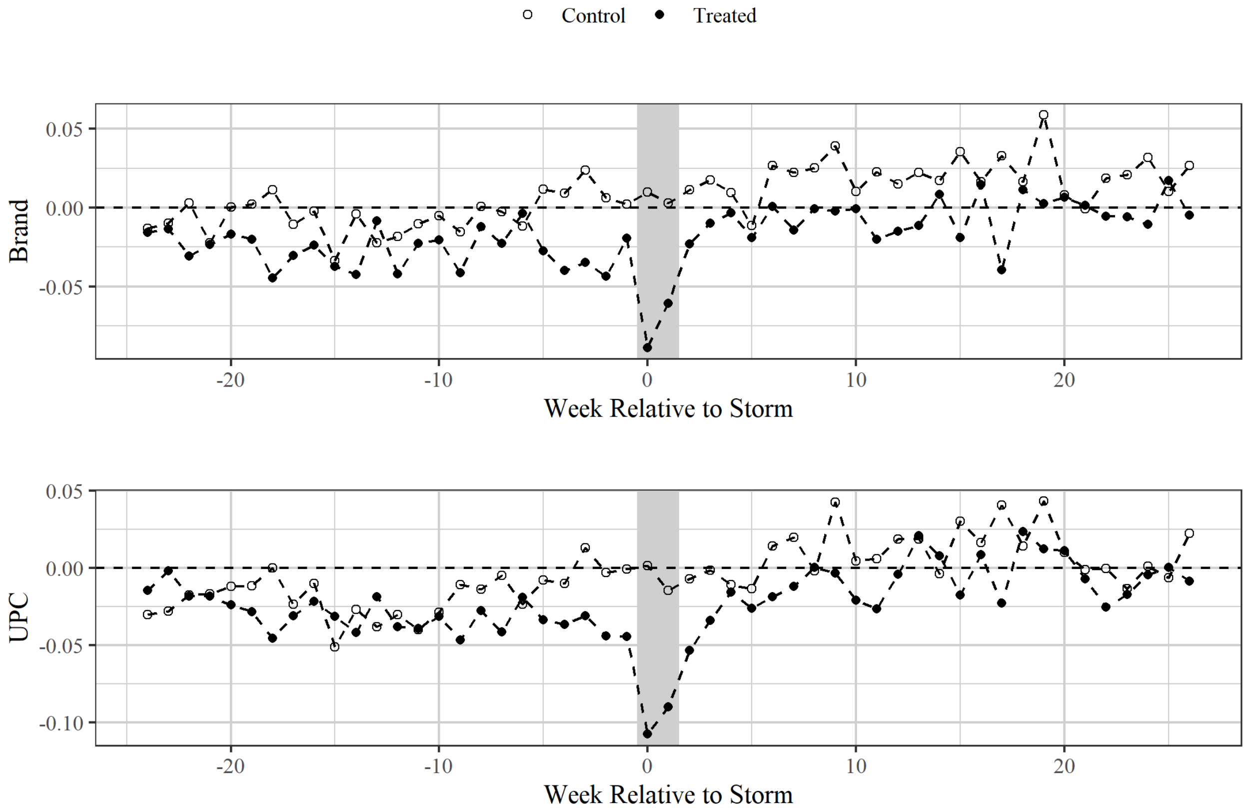


Figure 1.7: **Choice Persistence Over Time.** The graphs plot estimated week dummies for the treatment and control group from a regression that also includes consumer fixed effects. Weekly effects are estimated for weeks -24 to 26. The first week of the sample (week -25) constitutes the omitted category. The vertical gray bars indicate weeks 0 and 1 which are affected by stock-outs.

Multi-brand Purchases In columns (2) to (4) we provide additional robustness checks that deal with multi-brand purchases. Contrary to a structural demand modeling approach which assumes that consumers only purchase one product from the category on each trip, our definition of persistence in equation (1.1) can accommodate consumers purchasing multiple brands. For example, if a consumer purchases two brands on a given trip and one of those brands was also purchased on her previous trip, our persistence variable is equal to 0.5. As a first robustness check, we define a new persistence metric that is equal to 1 if *any* brand purchase on the current trip was also purchased on the previous trip and zero otherwise. Results from a synthetic control regression using this modified outcome variables are reported in column (2) and are similar to our baseline specification in column (1). Next, we switch back to our main measure of persistence and focus on households that only ever purchased one brand on any of their shopping trips. The results in columns (3) and (4) show that the null effect in the weeks immediately after the hurricane continues to hold for single-brand households as well as households that purchased more than one brand on at least one occasion. Together with results presented in columns (4) and (5) of Table 1.6 in the main paper, we conclude that our null finding is robust to a variety of ways of tackling multi-brand purchases.

Disruptions due to Hurricanes In Section 1.4.4 of the paper we discuss the possiblity that hurricanes lead to longer-term disruptions that affect consumers' purchase behavior. We show that consumer expenditure and average purchase price do not change in the long-run, which we interpret as evidence against long-term changes in purchase behavior. As an additional robustness check we re-estimate our main specification based on specific subsets of hurricanes that were relatively less disruptive. In column (5) we report results when we exclude households that were exposed to the two largest and most disruptive hurricanes, Harvey and Sandy. In column (6) we further restrict the sample and exclude households that were exposed to hurricanes that caused more than 10 billion dollars in damages.[32] We find that results look similar when analyzing behavior for those subsets of households.

Purchase Frequency In our main synthetic control specification, the composition of households changes over time because not all household purchase bottled water in every week of the sample. Moreover, the number of households that purchase in the category in weeks 0 and 1 is larger than in other weeks (due to additional purchases in the category that are triggered by the hurricane). Due to the increase in purchase incidence during the hurricane, it is likely that we oversample low frequency households during the hurricane relative to other time periods. Such a compositional change could impact our analysis if the behavior of households with a high or low purchase frequency differs systematically. The robustness check reported in Table 1.4 in Section 1.4.2 deals with this issues most directly, because we compare choice persistence for a given household on the last trip before and the first trip after a hurricane. Contrary to the synthetic control approach, the robustness check in Table 1.4 is based on balanced sample of households and therefore not affected by changes in the composition of households over time. As we discuss in detail in Section 1.4.2,

[32] https://www.ncei.noaa.gov/access/monitoring/billions/dcmi.pdf

this robustness check confirms our null results and yields a similar magnitude with regards to the persistence decrease during the hurricane as the synthetic control approach.

As an additional robustness check, we also re-estimate our synthetic control specification only based on households with a relatively high purchase frequency. In column (7) of Table 1.10 we report results when basing the synthetic control approach only on households with above median purchase frequency. We find that the null effects in weeks 2-5 continues to hold in this sub-sample of households. Based on this regression and the robustness check in Section 1.4 we conclude that compositional changes due to different purchase frequencies across households are not driving our null result.

1.7.5 Demand Model with State Dependence: Additional Details

In this section we provide additional details on the estimates from a choice model with structural state dependence presented in Section 1.5. We first outline how we select our sample and then provide a set of additional results for both the bottled water and the margarine category.

Sample Selection We follow Simonov et al. (2020) closely in terms of how we construct our estimation sample and we refer the interested reader to Appendix A of Simonov et al. (2020) for additional details on their sample construction for the margarine category. We replicate the sample construction outlined in Simonov et al. (2020) for margarine and also build a similar data-set for the bottled water category with slightly modified criteria, which we outline below. To construct both samples we combine the consumer-level Nielsen-Kilts Homescan (HMS) data and the store-level Retail Measurement System (RMS) data sets for the time span between 2006 and 2011. The combination of both data sets is required because we rely on the store-level data to construct price series.

In a first step we select the brand-size combinations with the highest purchase shares such that sales across all brand-size combinations constitute roughly 50% of the market. For margarine, this selection results in 4 brand-size combinations with 38 UPCs and for bottled water it results in 27 brand-size combinations with 78 UPCs. We then restrict the sample to households that made at least 85% of their category purchases at one store that appears in the RMS data set. For each such store, we obtain the weekly prices of the UPCs selected in the first step from the store-level data

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Dependent Variable	$\widetilde{Persist}$	Alt. Persist Metric	$\widetilde{Persist}$	$\widetilde{Persist}$	$\widetilde{Persist}$	$\widetilde{Persist}$	$\widetilde{Persist}$
Sample	Full Sample	Full Sample	Only 1 Brand	>1 Brand	Excl. 2 costliest hurricanes	Excl. hurricane with damage > 10 billion	Above M. Purchase Frequ.
Brand-level							
Week 0	-0.071***	-0.064***	-0.103***	-0.054***	-0.055***	-0.059**	-0.123***
	(0.018)	(0.019)	(0.026)	(0.020)	(0.020)	(0.024)	(0.022)
Week 1	-0.032*	-0.030	-0.098	-0.001	-0.019	-0.029	-0.080
	(0.019)	(0.020)	(0.028)	(0.022)	(0.024)	(0.032)	(0.025)
Weeks 2 -5	0.010	0.011	-0.006	0.017	0.018	0.011	0.000
	(0.012)	(0.012)	(0.017)	(0.014)	(0.013)	(0.020)	(0.015)
UPC-level							
Week 0	-0.081***	-0.070***	-0.117***	-0.062***	-0.085***	-0.093***	-0.124***
	(0.018)	(0.019)	(0.025)	(0.020)	(0.019)	(0.024)	(0.022)
Week 1	-0.051***	-0.053***	-0.113**	-0.024**	-0.033	0.011	-0.079**
	(0.016)	(0.019)	(0.026)	(0.020)	(0.022)	(0.032)	(0.023)
Weeks 2 -5	-0.004	-0.009	-0.026	0.005	-0.004	0.010	-0.007
	(0.011)	(0.013)	(0.018)	(0.014)	(0.014)	(0.019)	(0.016)
Treated Observations	38,044	38,044	10,699	27,345	19,855	10,507	25,454
Treated Households	2,201	2,201	775	1,426	1,197	589	1,055

Table 1.10: **Additional Robustness Checks.** All columns report average treatment effects for specific weeks (or groups of weeks) based on the generalized synthetic control method. Standard errors and significance levels are based on 500 bootstrap samples. Significance levels are calculated based on the distribution of bootstrap estimates and not based on a normal approximation. Significance codes: *p<0.1; **p<0.05; ***p<0.01.

and group products of the same brand and pack size together if their prices are highly correlated. For margarine this reduces 38 UPCs into 6 product groups, whereas for water this reduces 78 UPCs into 42 product groups. In both cases, we only maintain the largest product groups (2 product groups in the case of bottled water and 4 in for the margarine category). We then drop data for households who's primary store did not carry all of the product groups. Lastly, we drop households that made less than three non-outside option purchases.

Our final estimation samples comprises 2,232 households making 51,122 purchases from a set of four products in the case of margarine and 272 households making 8,661 purchases from a set of two products in the case of bottled water. The outside option is defined as the purchase of any other margarine / bottled water product. Shopping trips without a purchase in the category are not included in the sample.

We note that we end up with only 2 products groups for bottled water, relative to 4 product groups for margarine as well as a smaller sample of households in the case of bottled water. There are several reasons for this difference. First, spending on water is spread across more store types, including gas stations and convenience stores in addition to grocery stores. Therefore, limiting the sample to households that do at least 85% of their water purchases at one store leads to a larger decrease in sample size. Second, the market for water is less concentrated and exhibits more variation in the brands and products that different stores carry. Therefore, only very few stores carry the top 3 or 4 product groups of bottled water, which leads us to restrict the sample to only the top 2 product groups. For simplicity we refer to product groups as "brands" when presenting estimation results.

Estimation Results We present additional estimation results for the bottled water category (the primary category used in this paper) and margarine (the category used in Simonov et al. (2020)) in Table 1.11. Our estimates are based on our replication of the code from Simonov et al. (2020) which the authors generously shared with us.

We present results from 4 different specifications for both categories. In particular, we report state dependence estimates when (1) assuming no initial loyalty, (2) assuming the initial condition is exogenous, or when drawing the initial state from the appropriate distribution under the assumption of (3) i.i.d. prices or (4) prices that follow a first-order Markov process. This structure of organizing results mirrors Table 6 in Simonov et al. (2020). The margarine results are based on our replication of their estimates and therefore do not exactly match the numbers in Table 6 of Simonov et al. (2020). For simplicity we focus on the estimates of the state dependence parameters and omit other parameter estimates. The estimates presented in Table 1.7 in the main paper correspond to column (4) in Table 1.11.

Overall we find that the results in the bottled water category are remarkably similar to those based on margarine data. We find that the bias pattern described in Simonov et al. (2020) holds for bottled water as well. In particular, we find that structural state dependence is underestimated

		(1)	(2)	(3)	(4)
		$s_0 = 0$	$P(s_0\|\theta)$ ignored	$P(s_0\|\theta)$ included	
				Prices i.i.d.	Prices Markov
			Water		
State	μ_γ	0.721	2.139	1.005	1.233
Dependence		(0.528, 0.924)	(1.768, 2.555)	(0.74, 1.287)	(0.943, 1.54)
	σ_γ	1.188	2.054	1.195	1.471
		(0.99, 1.406)	(1.652, 2.49)	(0.938, 1.498)	(1.173, 1.804)
			Margarine		
State	μ_γ	0.641	2.508	0.985	0.987
Dependence		(0.581, 0.704)	(2.377, 2.639)	(0.887, 1.08)	(0.899, 1.075)
	σ_γ	0.887	2.509	1.118	1.113
		(0.824, 0.947)	(2.352, 2.668)	(1.022, 1.223)	(1.035, 1.204)

Table 1.11: **Structural Estimation of State Dependence under Different Treatments of the Initial Condition.** Column (1) and (2) either set initial loyalty to zero or treat it as exogenous. Columns (3) and (4) correct for the initial conditions based on different assumptions about the price process (i.i.d. prices versus a first-order Markov process). 95% posterior credible intervals are reported in paranthesis.

when the initial loyalty state is set to zero in column (1) and overestimated when assuming that initial loyalty is exogenous in column (2). The difference in the estimated state dependence parameter when allowing for a first-order Markov process in prices is relatively small in both categories, although there is slightly larger shift in the point estimate of mean state dependence in the water category. More importantly for the main research question of this paper, the estimates of state dependence for our preferred specification in column (4) are similar across the two categories and the positive and significant state dependence estimate for bottled water from the structural model is at odds with the null result we obtain when using our framework based on hurricane-induced stock-outs.

2 Grocery Store Closures and Household Nutritional Choices

Abstract: We analyze the impact of a temporary shock to food supply on households' dietary choices. We use hurricane-induced closures of grocery stores, focusing on temporary closures. Results show that store closures influence households' purchasing patterns even after the grocery store has reopened. We find a decrease in the nutritional value of purchase baskets of treated households for the nine-month period after the store has reopened, despite no change in total expenditures. Our findings support the hypothesis that supply factors play a substantive role in shaping household diets.

2.1 Introduction

Health practitioners and researchers link consumers' dietary choices to health outcomes, highlighting the importance for improved understanding of what drives consumer nutrition. For example, the Centers for Disease Control and Prevention (CDC) links poor nutrition to heart disease, stroke, type 2 diabetes, and some cancers; and, in response, recommends that consumers maintain a healthy diet to lower the risk of such preventable chronic diseases.[33] Even though dietary choices are well-understood to be important determinants of health outcomes, what drives differences in these choices across households is still subject to policy debate. Researchers and policy experts fall into two camps. One argues for supply-side mechanisms and attributes the less healthy diets in a neighborhood to low availability (and/or high prices) of healthy foods in that location (Sharkey et al. (2010); Algert et al. (2006)). In response, policy discussion on food availability often emphasizes proximity to grocery stores, spurring initiatives to assist grocery stores in under-served areas (e.g., Fresh Food Financing Initiative in Pennsylvania). The other camp argues for demand-side explanations where consumer preferences are the main determinants of diets and the lower prevalence of grocery stores in a region is an equilibrium response to lower demand (Cummins et al. (2014)). Most recently, Allcott et al. (2019) look at entries of supermarkets and conclude that these positive shocks to food availability do not meaningfully change households' nutritional intake.

We add to this policy debate by studying how negative shocks to supply may impact the nutritional composition of households' purchased baskets. We ask whether (and how) a household's

[33]See `https://www.cdc.gov/chronicdisease/resources/publications/factsheets/nutrition.htm`

shopping behavior changes after a temporary closure of a grocery store. The key idea is that inertia in households' purchase patterns potentially creates an asymmetry in how negative and positive shocks to supply affect households' purchase patterns. It is well-documented that households repeatedly purchase the same product and visit the same grocery store, which may be attributed to consumer preferences or habit formation (e.g., see Dubé et al. (2010b)).[34] This inertia in households' purchase patterns implies that they are less likely to explore new stores (and products) unless forced to do so. Our analyses leverage temporary closures of grocery stores as negative shocks to supply of healthier foods, which forces households to visit alternative stores in the interim. The potentially differential impacts of marginally increasing or decreasing supply call for a separate analysis of how households' nutritional choices change after a negative shock to supply.

The challenge in empirically evaluating the effects of food accessibility on healthfulness is that observed food supply and demand are jointly determined. For example, the entry or exit of a grocery store is not random and likely reflects changes in local neighborhood conditions, which may be correlated with households' purchased baskets. We overcome this challenge by using novel variation in supply availability—temporary closures of grocery stores, which occur right after a hurricane passes through the geographic location, due to, for example, short term supply-chain disruptions. The empirical strategy compares shopping patterns and dietary choices of households who visited the affected store to those of households who live and shop in the same area, but did not frequent the affected store. Importantly, treated households are forced to consider alternative stores during the month of the closure. The short-term nature of the closures is by design as it supports the assumption that changes in demand for healthy foods is not correlated with the temporary closure. It also allows us to look into households' store choices, by, for example, analyzing whether households return their business back to the affected store once it reopens. Thus, our analyses speak both to the policy debate on the causal relationship between nutrition and availability of healthy foods, as well as to increasing our understanding of consumers' decisions on store-choice and habit formation.

Our main data sources are the Nielsen's Homescan dataset, which tracks a panel of 61,000 households and reports their grocery purchases over the sample period of 2004-2019; and the

[34]In our sample, we see that 60% of a household's expenditure (within grocery stores) is attributed to the store that the household visited in their last trip.

Nielsen's Retail Measurement Services (RMS) panel, which records sales at the store-UPC-level over the same sample period. We match these datasets to information on hurricane location, dates, and wind speed from HURDAT2. We mark a grocery store as closed due to a hurricane if the timing of a hurricane in that area perfectly coincides with a large and temporary drop in reported sales and no recorded visits from Homescan panelists. We mark a household as treated if it meets the following two conditions: (1) the closed store was the most important grocery store for the household (with highest share of food spending in the three months before the storm), and (2) the household reports grocery purchases in the same Zip3 in the months following the storm. We identify 751 households as treated by a temporary closure of a grocery store. We compare their shopping behavior for nine months after the store has reopened to a control group of 7,524 households, who shopped in the same area but did not frequent a grocery store that temporarily closed right after a hurricane.

Shopping baskets are multi-dimensional objects that are hard to summarize. We approximate the USDA's Healthy Eating Index (HEI) by tracking the quantity of fruit, vegetables, and legumes, which correspond to the first four food groups in the HEI.[35] Our measure of nutritional value (NV) captures quantities of fresh, frozen, and canned products. We see the previously documented nutrition-income gap with our NV measure—households in the top income quartile have purchase baskets scoring 0.22 standard deviations higher than households in the bottom income quartile.

Results suggest that these temporary closures of grocery stores impact household shopping behavior, and these effects persist even after the stores have reopened. Looking at the nine months after the store has reopened, we see a drop in NV of 0.058 for treated households relative to the control group. The drop corresponds to 13.6 ounces of spinach (1.7 bags) a month. One may also interpret the magnitude of this result in terms of its size relative to the income-nutrition gap. We find that the drop is economically meaningful as it corresponds to 29% of the estimated nutritional-income gap in our sample. Importantly, we see that households' monthly expenditures on food item and number of trips do not change; instead, we see that the places where (treated) households shop change in the 'short-run.'

The key idea is that during the time of the closure households are potentially forced to explore

[35] This is a useful measure of household diets as health practitioners have raised concerns that Americans consume an "inadequate" quantity of fruit and vegetables in their diets. As such, these food groups are heavily weighted in the construction of the HEI and have large positive impacts on the index.

new stores, which may allow them to discover new food items. To help interpret our results, we analyze potential mechanisms and ask where do households shop after the negative shock to supply? We only consider temporary closures, so we first ask if households return their spending back to the affected store after it has reopened. We look at households' shopping patterns separately for the first three months after the reopening (months 1-3) and a second post-reopening period capturing months 4-9. The closure of a favorite store (with largest share of food expenditure) is likely to encourage households to experiment and shift expenditure to other stores. Results confirm the expected decrease in households' share of expenditure in the treated store in months 1-3. Households' share of expenditures at the affected store returns back to its original levels when we look at the second post-reopening period of months 4-9.

Next, we look at the share of expenditure across store types. Grocery stores are of particular interest to health researchers and policy makers, relative to discount and convenient stores, because they offer a wider selection of food items.[36] Thus, we analyze changes in households' share of expenditure at grocery stores, at the largest discounter in the data (which we infer to be Walmart), and at drug/convenience stores. We find that in the first 3 months after the store has reopened treated households spend a higher proportion of their food expenditure at Walmart and at drug/convenience stores. Different store formats have, on average, different product selections, prices, and locations; hence, one expects that changes in the set of stores visited by a household may lead to changes in its purchase basket. Even though we may not distinguish between the role of prices and assortments in shaping diets, our results contribute to our general understanding of the relationship between store choice and households' diets. We connect these changes in households' diets to forced experimentation, during which households shift expenditures away from grocery stores and towards a large discounter and/or drug stores in the first three months following the reopening of the store.

Results also show that households shift expenditures back to grocery stores by months 4-9. Nevertheless, the changes in the nutritional value of purchased baskets are more persistent and remain economically meaningful for the second post-reopening period. Any changes in diets may become permanent if the forced experimentation resulted in households discovering food items that

[36] We see in RMS data that this is especially the case for healthier product categories, such as fruit and vegetables, compared to less healthy options such as grains and snacks.

they prefer (or due to state dependence in purchase choices).

Our identification strategy adds credibility to the assumption that demand for healthy foods is not correlated with the store closures, because our store closures are temporary and associated with randomly occurring hurricanes. However, one may be concerned that treated households are more heavily affected by the hurricane through mechanisms other than the temporary closure of the grocery store. We argue that this is not likely because we use temporary closures and confirm that the areas "recovered quickly" after the hurricane. In addition, several patterns in the data alleviate such concerns. Even though there is a negative effect on nutrition in the post-reopening period, there is no effect on total expenditures on food items and trips. Reassuringly, during the month of the hurricane, we only see a change in shopping at the treated store, and no decreases in other stores. We also confirm that treated and control households were similarly exposed to the hurricane in terms of the distance between household Zip5 and the center of the hurricane, and both groups similarly prepare for the hurricane. We acknowledge that, if the storm affected treated households more severely in unobserved ways that we may not capture, then our estimates may overstate the effect of the grocery store closure on household diets.

The importance of dietary choices for health outcomes has spurred a wide range of research, which looks at how information, circumstances, and food accessibility may impact nutritional choices (few examples include Algert et al. (2006), Sharkey et al. (2010), Atkin (2016), Allcott et al. (2019), Hut (2020)). While the causes of nutritional inequality are not fully understood, recent works find that households exhibit persistent patterns in their nutritional diets regardless of changes in availability (Allcott et al. (2019)) or health status (Hut and Oster (2019)). For example, Allcott et al. (2019) look at the marginal effect of an additional grocery store and conclude that increased food availability has economically small effects on healthy eating. These results do not necessarily conflict with our findings, because they look at the marginal effect of an additional store, whereas we use shocks that temporarily restrict access to consumers' favorite store, forcing them to shop in other places. We confirm that, absent store closures, consumers often visit the same store; and we speculate that a negative shock to supply is more likely to influence shopping patterns than the addition of a new store, in line with the forced experimentation arguments in Larcom et al. (2017).

Prior research has also looked at the relationship between store format and dietary choices:

Courtemanche and Carden (2011) looks at entry of Walmart, Volpe et al. (2019) supermarkets and supercenters, Ailawadi et al. (2018) club stores. Our analyses of potential mechanisms lend support to the hypothesis that store formats likely influence household shopping baskets. More broadly, this project contributes to the literature connecting availability to consumer choices (e.g., Hut (2020), Bronnenberg et al. (2021)). Last, our identification strategy relates to papers that use natural events for identification, such as Raval et al. (2020), Levine and Seiler (2022), Figueroa et al. (2019). For example, Raval et al. (2020) use hospital closures following natural disasters to evaluate the performance of discrete choice demand models.

The remainder of this paper is organized as follows. Section 1.2 describes the data and construction of the nutritional index. Section 2.3 discusses our empirical strategy. Results are described in section 2.4. Section 3.7 concludes.

2.2 Data

The analyses use three data sources—Nielsen's RMS and Homescan panels, and HURDAT2. The RMS dataset records weekly sales volumes for each UPC sold at the tracked stores from 2004 to 2019. The Homescan data is a longitudinal panel following over 40,000 U.S. households over the same time period. Each year households report demographic information describing their income, education, size of household, zip code, marital status, and race. Table 2.1 describes demographics for the full sample of households. Panelists continually provide information about their purchases using in-home scanners or mobile apps. For each household, the dataset records the stores they visit, the products purchased, and prices paid.

We use hurricane information from HURDAT2, which is collected by the National Hurricane Center. HURDAT2 tracks the timing, intensity, and geographic coordinates of all known hurricanes from 1,851 to today. This detailed information allows us to identify the hurricane's path through geographic areas, the timing of the storm, and wind speed. To identify stores that close right after a hurricane, we match hurricane information to zip codes of RMS stores and Homescan panelists.

Next, we describe how we identify the temporary closures of grocery stores, the households affected by them (which we refer to as the treatment group), and the construction of the control group. Section 2.2.2 details the construction of households' nutritional intake using detailed purchase data. Household shopping patterns are summarized in section 2.2.3.

	mean	st. dev
Household Size	2.25	1.19
Age	56.92	11.54
Household Income ($000s)	56.94	33.35
Years Education	14.59	2.39
Married	0.61	0.49
White	0.85	0.36
Black	0.09	0.28
Employed	0.68	0.47
Weekly Work Hours	19.51	16.76

Table 2.1: **Demographic Information of Nielsen Panelists.** The table describes household demographic variables collected each year by the Nielsen Consumer Panel of over 40,000 U.S households from 2004 to 2019.

2.2.1 Store Closures and Treated Households

To identify hurricane-induced closures of grocery stores, we match the locations of stores visited by Nielsen panelists to the geographic locations affected by hurricanes. Nielsen suppresses the last two digits of store zip codes to maintain retailer confidentiality; thus, we match physical locations of stores to the first three digits of a zipcode (we refer to a region as a Zip3). Next, we identify which of these Zip3s intersect with the path of a hurricane using storm-tracking data. From HURDAT2 we identify the weeks and Zip3s affected by hurricanes during our sample period, from 2004 to 2019.[37] We mark a Zip3 as affected by the hurricane if the winds (at any intensity) intersected with the Zip3. We track 41 hurricanes, passing through 2,150 Zip3-weeks, for a total of 22,685 grocery store-weeks exposed to a hurricane. We mark a grocery store as closed due to a hurricane if we see (1) a large and temporary drop in sales (relative to a 10 week moving average) in one of the two weeks following the hurricane, and (2) no trips to that store from Homescan data. We identify 447 grocery store closures, whose sales decrease by an average of 80% right after the storm.

Households are *treated* by a closure if they regularly visited a *treated* store prior to the hurricane. There are 2,793 households that visited at least one of the identified 447 stores in the three months preceding the hurricane. We mark a household as treated if it meets the following conditions: (1) the closed store was of relative importance to that household, and (2) the household reports grocery

[37]We create Zip3 regions by taking the union of all Zip Code Tabulation Areas (ZCTAs) that correspond to a Zip3 code.

purchases in the affected Zip3 in the month following the storm. We define a store as important if it accounts for the largest share of food spending for the household during the three months prior to the storm.[38] The second condition requires that panelists remain in the same geographic area and continue reporting to Nielsen, which aims to minimize concerns about evacuation and other impacts of the hurricane (separately from the closure). There are 751 households that satisfy these conditions. Table 2.2 shows the number of treated households and store closures used for the analysis for each hurricane in our sample. For example, we identify 42 store closures and 98 treated households affected by Sandy in November 2012. Our analyses compare shopping patterns of treated households to panelists that live and shop in the same area but did not experience a store closure. The control group is comprised of households that conduct their shopping within the Zip3 of the closed store and did not frequent the closed grocery store. We identify 7,524 control households.

2.2.2 Nutritional Value of Purchase Baskets

Our primary outcome summarizes the nutritional value of household diets using reported food purchases in the Homescan dataset. We use nutrition data from USDA's Food and Nutrient Database for Dietary Studies (FNDDS) to describe the nutritional value (NV) of purchase baskets. A common starting point is the Healthy Eating Index (HEI), which is a standard measure of dietary quality. The HEI is developed by the USDA in order to help practitioners, researchers, and policy makers assess how household food consumption aligns with the Dietary Guidelines for Americans (Krebs-Smith et al., 2018). The HEI rewards consumption of 'adequacy' components (fruits, vegetables, legumes, grains, dairy, protein, and fatty acids), and penalizing consumption of 'moderation' components (refined grains, sodium, added sugars, and saturated fats). Using the HEI to score Homescan transaction data is nontrivial because the USDA and Nielsen databases describe food items at different levels of granularity. Carlson et al. (2019) develops a useful crosswalk for mapping scanner data (from IRI) to the HEI. Unfortunately, the crosswalk is not yet available for use with the Nielsen panel data and the authors advise that the crosswalk offers a reliable match only for UPCs and purchases in 2017-2018.

[38]We considered other definitions of store importance, such as capturing at least 30% of food spending, and got similar results.

Hurricane	Month	Grocery-Store Closures	Treated Households
Matthew	2016m10	86	208
Ike	2008m9	38	163
Florence	2018m9	50	124
Sandy	2012m11	42	98
Gustav	2008m9	19	51
Irene	2011m9	22	39
Harvey	2017m9	7	15
Hanna	2008m9	9	12
Irma	2017m9	8	9
Hermine	2016m9	7	8
Michael	2018m10	5	5
Frances	2004m9	3	3
Dorian	2019m9	3	3
Isaac	2012m9	2	2
Arthur	2014m7	1	2
Jeanne	2004m10	2	2
Barry	2019m7	2	2
Noel	2007m11	2	2
Rita	2005m9	1	2
Dennis	2005m7	1	1

Table 2.2: **Store Closures and Treated Households.** The table reports the number of grocery-store closures used for the analysis and the number of treated households per hurricane over the sample period of 2004-2019.

Thus, we approximate the HEI by scoring purchase baskets focusing on quantities of fruit, vegetables, and legumes. The included products span the departments of fresh produce, dry grocery, and frozen foods. We capture protein, sodium, sugars, and fats through their presence canned, frozen or dry products. Reflecting the importance of these food groups for healthy diets, they comprise the first four (most heavily weighted) food components in the HEI. We view our NV variable as a good proxy for healthy diets as practitioners have raised concerns that Americans consume an "inadequate" quantity of fruit and vegetables (Center for Nutrition Policy and Promotion (CNPP), Volpe et al. (2019)). Reassuringly, the strongest correlates of the HEI measure used by Allcott et al. (2019) are purchases of fruits and vegetables (correlated at $\rho = 0.56$ and $\rho = 0.41$ respectively). Data appendix 2.6 describes the extensive work related to mapping Homescan to FNDDS, and how our measure of nutritional intake relates to the HEI.

Next, we confirm that we observe the previously documented nutrition-income gap. We calculate the average monthly NV for each household and normalize NV to be mean zero and have a standard deviation of 1. Figure 2.1 shows the expected positive relationship between household income and nutrition. The x-axis tracks percentiles of household income, the y-axis reports average monthly NV after controlling for household size, age of head of household, and panel year. The figure depicts both a fitted polynomial as well as the average NV by income decile. To put these values in perspective, consider the income gap between households in the top and bottom quartiles of the income distribution. We find that the higher income households on average purchase groceries that are 0.22 standard deviations higher than the lower income households. Similar to Allcott et al. (2019), our measure shows that the income-nutrition gap is increasing during the sample period. For example, the difference between high and low income households increases from 0.18 standard deviations in 2004–2012 to 0.24 standard deviations in 2013-2019.

Homescan presents great detail on household purchases of grocery items, which allows us to evaluate household shopping patterns after a shock to supply. The dataset does not record consumption away from home, e.g. in restaurants. We acknowledge that we may not capture the effect on overall diets, if the shock to supply influences food consumption away from home or purchases do not map into consumption.

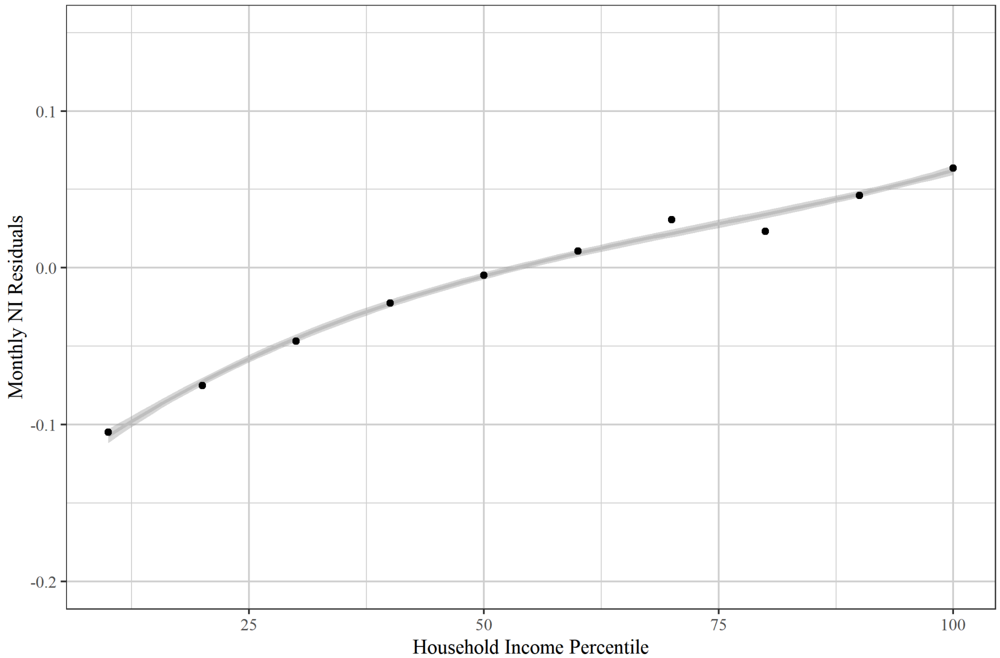


Figure 2.1: **Income-Nutrition Gap.** This graph shows a third degree polynomial of household monthly HEI by income, when controlling for household size, age, and panel year.

2.2.3 Household Shopping Patterns

Table 2.3 reports shopping behavior for the treated and control households. In each row, we summarize average monthly shopping patterns in the pre-treatment period. Column 1 summarizes the variables for treated households, and column 2 for the control sample. Treated households spend, on average, $237.80 and shop on 8.9 occasions in a month. In terms of dietary choices, treated households purchase baskets with a nutritional value of 0.023, on average (monthly NV is normalized to be mean zero and have a standard deviation of 1). We approximate transportation costs as the share of trips outside of home Zip5. The last two columns show the differences between these averages across treated and control households and the p-value of a t-test. Reassuringly, we see that the two groups of households have very similar shopping patterns before the hurricane.

Another dimension of shopping patterns describes where households purchase their groceries in terms of the stores type that they visit. Perhaps unsurprising, households do most of their food shopping at grocery stores, which, on average, capture 60% of food expenditure. The largest discount store in the data, which we infer to be Walmart, accounts for 25% of food spending, while drug and convenience stores account for 5%. The remaining stores span dollar stores, club stores,

	Treated	Control	diff	p-val
spending	237.808	233.397	4.411	0.459
trips	8.996	8.719	0.277	0.211
nutrition (NV)	0.030	-0.003	0.033	0.393
travel distance	0.393	0.387	0.006	0.683
spend at grocery	142.227	141.940	0.287	0.946
spend at most preferred store	104.332	108.278	-3.946	0.284
Observations	751	7524		

Table 2.3: **Shopping Patterns of Treated and Control Households.** The table summarizes household shopping patterns in the 5 months prior to the storm separately for treated and control households. The last two columns report the difference in averages between the two groups and the p-values of a t-test. Summarized variables are: spending on food items, number of trips, NV (approximation of nutritional value=quantity of fruit, vegetable, and legume) normalized to mean 0 and standard deviation of 1, transportation cost is approximated as share of store trips outside of home zip5, spending at grocery stores, spending in store with highest food expenditure.

and stores that are not identified in the Nielsen data.

We also summarize average households' spending in their "most preferred" stores. For each household, the "most preferred" store is defined as the grocery store with highest expenditure in the three months before the storm. For treated households, this is the spending in the store with a temporary closure, which coincides with the passing of a hurricane. We see that the most preferred store accounts for 44% to food expenditure for treated households and 46% for control households. Again, we do not see any statistically significant level differences across treated and control households. We discuss parallel trends in section 2.3.

2.3 Empirical Strategy

The empirical analyses add to our understanding on whether (and how) supply-side conditions affect households' diets. Policy discussions on food availability emphasize the importance of proximity to a grocery store. As a result, we focus on understanding how consumer shopping behavior may change after the closure of a grocery store. The usual caveat of studying the relationship between food availability and household shopping patterns is that food supply is likely correlated with demand. For example, using store exist may overstate the effect of supply-side conditions as the strategic decision by the store is likely correlated with a decrease in demand for its products. To avoid these concerns we use store closures that are plausibly exogenous, as they are temporary and occur immediately after a hurricane passes through an area.

For each household, we define the following time periods: (1) pre-storm period: the 5 months before a hurricane passes through the geographic area; (2) the closure/hurricane month; (3) the post-reopening period: the 9 months after the store has reopened. Figure 2.2 displays the timing and notation for a representative treated unit whose grocery store closes during August (month 0). The pre-closure period (months -5 to -1) is used for verifying that the treated and control groups follow parallel trends. We expect that treated households adjust their shopping behaviors and substitute to other stores while a grocery store they frequented is closed (month 0). These immediate changes in store visits and potential changes in households' purchased baskets during the closure period are not the focus of our analysis. Instead, our analyses focus on households' shopping and dietary choices in months 1 through 9, which we call the post-reopening period.

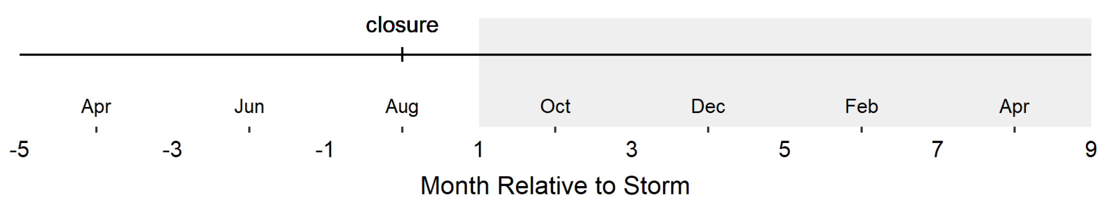
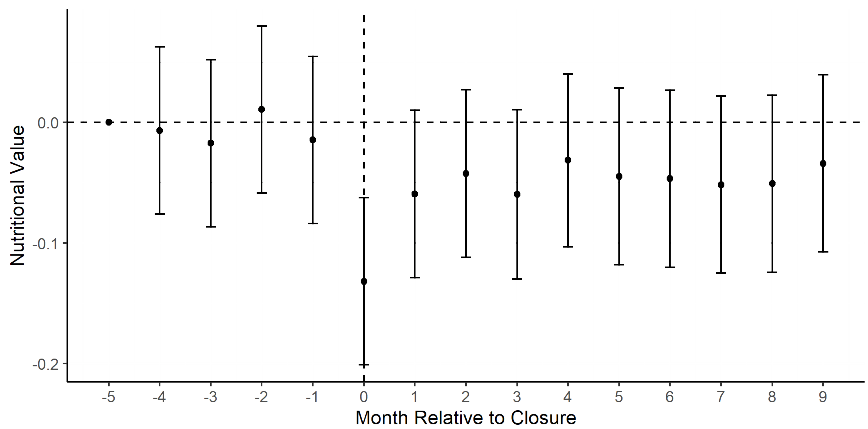


Figure 2.2: **Panel Structure.** This figure displays the timing and notation for a hypothetical treated unit, whose store closed during the month of August. Months -5 to -1 denote the pre-treatment period. The vertical dashed line marks the store closure, at month 0 (August). We define the post-reopening period as months 1-9, represented by the shaded grey region.

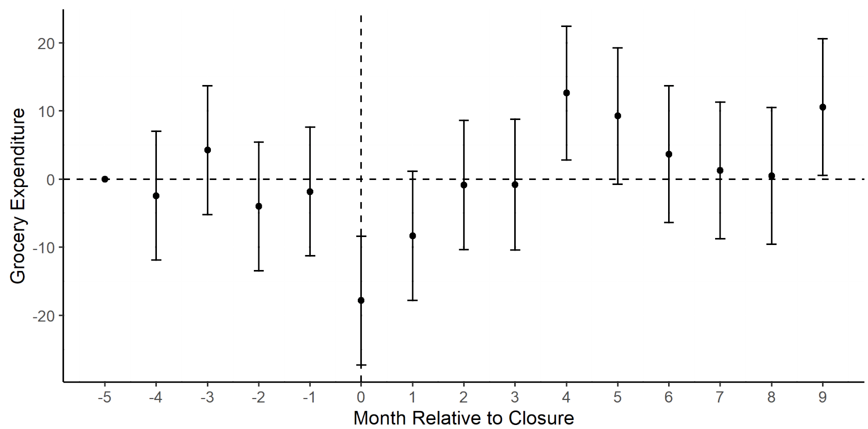
We estimate two-way fixed effect models specified as

$$Y_{it} = \alpha_i + \gamma_t + \beta_0 T_i \mathbf{1}(\text{Closure}) + \beta_1 T_i \mathbf{1}(\text{Post-reopening}) + \epsilon_{it}, \tag{2.1}$$

where T_i is a dummy equal to one for households in the treatment group. Household and month fixed effects are denoted by α_i and γ_t respectively, ϵ_{it} is a random error term. The immediate effect of the closure is captured in β_0; β_1 is our main coefficient of interest capturing the effect in the post-reopening period (months 1-9). Our main variables of interest for Y_{it} are nutritional value of purchase baskets and expenditures.



(a) Nutritional Value



(b) Food Spending

Figure 2.3: **Households' Shopping Patterns over Time.** These graphs show simple differences between treated and control households in shopping patterns over time. In the first panel we plot nutritional value, in the bottom panel we plot monthly expenditure on food items (groceries purchased in any type of store). The vertical dashed line at month 0 represents the timing of the store closure.

Figure 2.3 plots the difference in NV and spending across treated and control households in each month relative to the store closure (marked as month 0). The pre-treatment shopping patterns match closely (months -5 to -1). During the month of the closure, NV and grocery spending fall for treated households (marked by the dashed vertical line). We estimate these effects but do not interpret them, because households may need some time to adjust to other stores after an unexpected closure. We see a persistent drop in NV in the post-reopening period. However, this drop is not matched by a drop in spending. Panel (b) shows that expenditure returns back to the the level of the control group by month 2, with some noisy estimates at month 4 and 9.

There are two main threats to identification: (i) treated and control households are different and have different pre-treatment trends in shopping patterns; (ii) separately identifying the effects of the hurricane from the effect of the store closure. Using pre-treatment periods, we regress each outcome on a set of household and month fixed effects, as well as a linear time trend interacted with treatment dummy. The estimates in table 2.4 confirm our intuition from the graphs: there is no (statistically significant) difference in pre-treatment linear trends between treated and control groups for the set of variables describing shopping behavior. We proceed under the assumption that absent the shock to supply, in the form of a temporary closure of a grocery store, the nutritional value for treated and non-treated households would have evolved similarly. Section 2.4.2 provides supportive evidence that the treated and control households are similarly affected by the hurricane.

	NV	Spending	Trips	Spend in "Treated" Store	Travel Distance	Spend at Grocery Store	Share Spend at Grocery Store
treated × time trend	0.001	-0.088	-0.021	1.173	-0.003	0.182	0.001
	(0.013)	(1.158)	(0.045)	(0.904)	(0.005)	(0.958)	(0.002)
Observations	41,373	41,373	41,373	41,373	41,373	41,373	40,686
Adjusted R^2	0.506	0.681	0.755	0.652	0.602	0.669	0.668

Significance levels: *p< 0.01, **p< 0.05, ***p< 0.01.

Table 2.4: **Analysis of Differential Pre-Closure Trends.** These are results from regressions using only pre-treatment data. An observation is at the household-month level. 'Time trend' is a continuous variable tracking months prior to the hurricane. All regressions include household and month fixed effects. We find no significant difference in pre-treatment trends between treated and control households for any of the outcome variables.

2.4 Results

Visual inspection of Figure 2.3 suggests that the temporary closures of grocery stores lead to a decrease in the nutritional value of household purchase baskets even after the store has reopened (months 1 to 9). This occurs despite no decrease in monthly expenditures on grocery items relative to control households. Table 2.5 presents our main results using regression equation 2.1, where standard errors are clustered at the Zip3 level. Our focus is on the β_1 estimate of the post-reopening period, and we report the estimates for the closure month (β_0) for completeness.

Column 1 shows the results for our measure of household diets, NV, which tracks quantity of

	NV	Spending	Trips	Spend in Other Stores	Travel Distance
treated × closure	-0.132***	-16.171***	-0.746***	-1.884	0.006
	(0.029)	(6.063)	(0.196)	(4.444)	(0.012)
treated × post-reopening	-0.058***	1.112	0.018	2.265	0.012
	(0.019)	(3.747)	(0.151)	(3.595)	(0.011)
Observations	114,088	114,088	114,088	114,088	114,088
Adjusted R^2	0.456	0.634	0.722	0.619	0.544

Significance levels: *p< 0.01, **p< 0.05, ***p< 0.01.

Table 2.5: **Effects of Store Closure on Household Nutrition and Shopping Patterns.** An observation is at the household-month level. We report the change in behavior separately for the hurricane month (treated× hurricane), and for the 9-month period after the hurricane/the store reopens (treated×re-opening). All regressions include household and month fixed effects. Standard errors are clustered at the Zip3 of the treated store.

fruit, vegetables, and legumes. NV of treated households decreases by 0.058 points (with a standard error of 0.019). The drop corresponds to about 13.6 ounces of spinach (about 1.7 bags) a month. As health guidelines emphasize the consumption of fruit, vegetables, and legumes, and refer to them as the most healthy food groups, we interpret a decrease in these food groups as a decrease in the nutritional intake of households. One may also interpret the magnitude of this result in terms of its size relative to the income-nutrition gap. In our sample, households in the bottom quartile of the income distribution, on average, purchase food baskets with 0.20 points lower NV than households in the top quartile of the income distribution. The β_1 estimate corresponds to 29% of the estimated nutritional-income gap.

An important follow-up question is whether there is heterogeneity in how supply shocks are borne by different households. From a policy perspective, one may want to understand whether these effects differ systematically between demographic groups. For example, do more vulnerable populations (e.g., lower income, lower educated) change their diets more after a closure of a grocery store? The benefit of our estimation approach is that it allows for a clear identification strategy. Its cost is that meaningful analysis of heterogeneous treatment effects is difficult because it involves partitioning an already small sample of treated households. We analyze heterogeneous treatment effects on NV across the following dimensions: income, presence of a child, and age. As expected, these analyses are noisy and do not imply differences in adjustments across household groups.[39]

Next, we check whether households' total expenditures on food items or number of trips change across treated and control units. Results in column 2 show no persistent differences in expenditures in the post-reopening period—treated households spend as much as control households on food items after the affected store has reopened. The results are similar for the number of trips involving a grocery purchase. Another relevant dimension for household shopping choices is store locations and travel costs. For this exercise we inferred store locations at the Zip5 level and asked if treated households are more likely to shop in stores located outside of their home Zip5. Using this coarse measure of transportation costs, column 5 shows no change in the post-reopening period.

Even though we do not interpret the immediate changes in shopping patterns and diets during the store closure, it is informative to decompose why spending decreases for the treated households during the month of the closure. The average monthly basket for a household is \$237.80, so the \$16.17 drop accounts for 6.8% of spending. We also look at the change in spending in stores outside of the ("most preferred") treated store, what we call *spending in other stores* in column 4. For each household, the "most preferred" store is defined as the grocery store with highest expenditure in the three months before the storm; for treated households this is the closed store. Our estimates in column 4 show that during the month of the storm, there is no drop in spending in stores other than the closed one. That is, the immediate change in spending is explained by a decrease in spending at the closed store only. These results mitigate concerns that unobserved hurricane effects are affecting treated and control households differentially, separately from the analyzed store closures.

Our results contribute to the discussion on the nutrition-income relationship, showing that supply-side conditions may be important determinants of households' diets. The motivating idea behind our empirical strategy is that it forces households to explore alternative stores. Households exhibit persistence in their shopping patterns and repeatedly visit the same stores. In our data 60% of expenditure (within grocery stores) is attributed to the same store that the household last visited. Therefore, unless forced, households are unlikely to deviate from their usual store. Our empirical strategy does just that—the temporary closure of the grocery store forces households to shift expenditure to alternative stores during the closure period. The change in where households shop may influence their shopping baskets, if, for example, the alternative stores offer different product

[39]Results are available from the authors.

selections or charge different prices. Households may continue shopping in these alternative stores (even after their favorite store reopens in months 1 through 9) if they find a store they like better or due to state dependence in store choices. Similarly, even if households return back to the affected store, they may maintain the changed purchase baskets for analogous reasons (discovering more preferred food items or state dependence). In the next section we look at how shopping behavior changes during the analyzed nine months, which helps us evaluate potential mechanisms behind our results on nutritional value of purchased baskets.

2.4.1 Potential Mechanisms - Purchases across Store Types

Availability of grocery stores is at the heart of policy discussions aiming to "improve" household nutritional intake. In this section we explore whether treated households switch food purchases away from grocery stores (i.e., towards discount or convenience stores). We use temporary closures, thus, we expect that after the initial adjustments during the closure period, households may return back to their most preferred store after it reopens. To describe potential mechanisms, we allow for a time interaction of the treatment effect and look separately at the effect of closure during months 1-3, and 4-9. That is, we adjust our estimation equation to estimate a separate effect of the closure for the first three months (β_1) and the rest of the post-reopening period (β_2)

$$\begin{aligned} Y_{it} &= \alpha_i + \gamma_t + \beta_0 T_i \mathbf{1}(\text{Closure}) \\ &\quad + \beta_1 T_i \mathbf{1}(\text{Post-reopening (1-3)}) + \beta_2 T_i \mathbf{1}(\text{Post-reopening (4-9)}) + \epsilon_{it}. \end{aligned} \tag{2.2}$$

Table 2.6 shows the results from regressions that allow for these interactions. We do not report the estimates during the closure months ($\hat{\beta}_0$) for ease of readability. Columns 1 and 2 repeat the analyses using basket nutritional value and spending as variables of interest. The effect of the temporary shocks to supply on NV remains relatively stable for the two post-reopening periods. Again, we do not see any contemporaneous changes in spending. Column 3 shows that treated households switch away from the affected store in the short term (months 1-3), during which they direct a larger share of their expenditure towards other stores. We interpret these initial changes in household shopping patterns as the driving factor for the change in NV. Next, we ask: where do treated households go?

	NV	Spending	Sp. Share Other Store	Spend Share across Store Types Grocery	Largest Discounter	Drug/Conv.
treated × (m1-3)	-0.060***	-3.162	0.026***	-0.011*	0.013**	0.005**
	(0.019)	(4.309)	(0.008)	(0.006)	(0.005)	(0.002)
treated × (m4-9)	0.057**	3.937	0.018	-0.008	0.003	0.006***
	(0.025)	(3.830)	(0.012)	(0.006)	(0.005)	(0.002)
Observations	114,088	114,088	111,882	111,837	11,837	11,837
Adjusted R^2	0.456	0.634	0.596	0.646	0.707	0.440

Significance levels: *p< 0.01, **p< 0.05, ***p< 0.01.

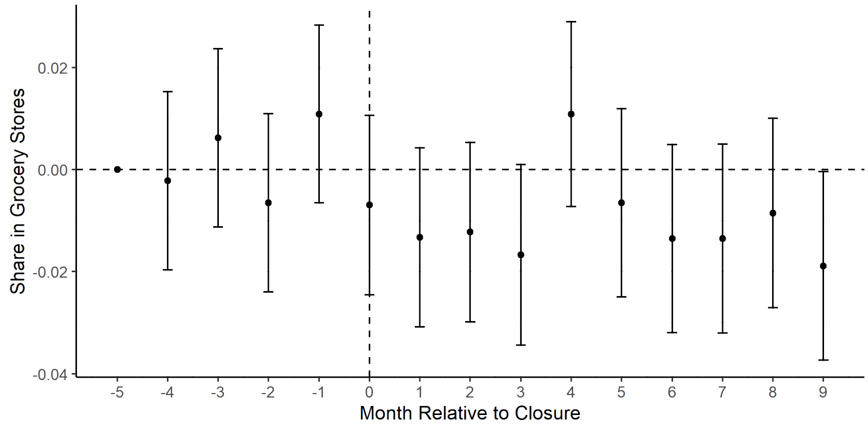
Table 2.6: **Mechanisms.** An observation is at the household-month level. All regressions include household and month fixed effects. Standard errors are clustered at the Zip3 of the treated store. We report the change in choices separately for the months 1-3 and months 4-9 after the store reopens. Column 3-6 have fewer observations, as we only use months in which households report food spending to construct share variables.

We look at the share of expenditure allocated to different channel types: grocery stores, the largest discount store in the Nielsen dataset (which we call Walmart), and drug/convenience stores. Figure 2.4 shows the difference in share of spending across these types of stores in each month relative to the closure. The adjustments are noisy, but some general patterns emerge: treated households adjust their shopping choices across store types in months 1 through 3, which informs our decision to look at the short-term effects of the store closure in these months, separately from the second part of the post-reopening period. Columns 4-6 in table 2.6 show the results. We see a drop in share of spending at grocery stores, which is matched by increases in spending at Walmart and at drug/convenience stores.[40] These analyses show that the changes in the nutritional value of household shopping baskets occur simultaneously with changes in store choices. Mechanisms connecting store format to purchase baskets relate to differences in location, assortments, and prices.[41]

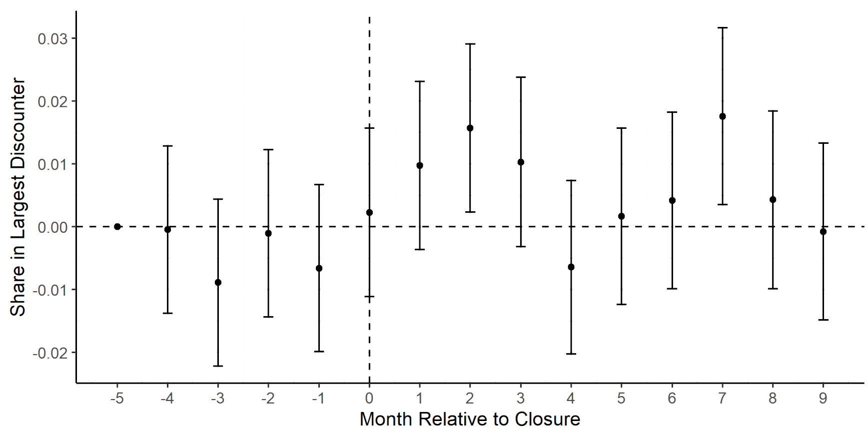
Our results also show that, even though households shift expenditures back to grocery stores (and to their preferred store) by months 4-9, the changes in the nutritional value of purchased baskets are more persistent. We saw in column 1 that treated households have lower NV relative to control group for the second post-reopening period of months 4-9. Intuitively, the initial changes

[40]The dollar spending across store types changes in the same directions, however, the adjustments are only statistically significant for drug/convenience stores.

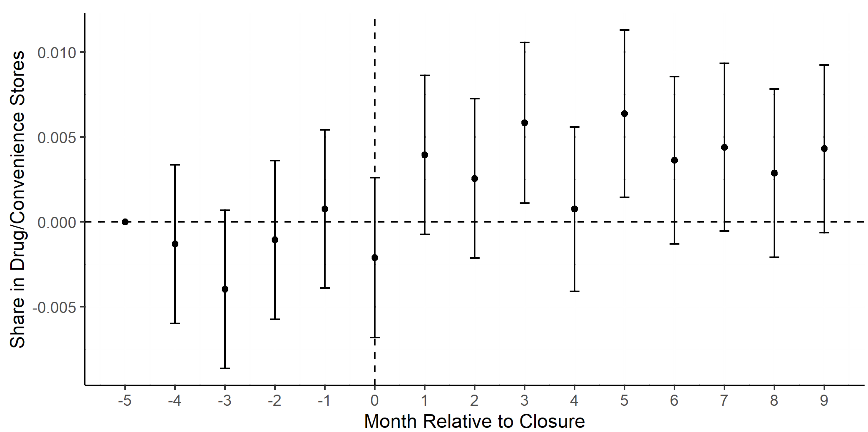
[41]Our data and identification strategy do not allow us to distinguish between these mechanisms.



(a) Spend Share at Grocery Stores



(b) Spend Share at Largest Discounter



(c) Spend Share at Drug/Convenience Stores

Figure 2.4: **Spend Share across Store Types.** These graphs show the average share of grocery expenditure attributed to different store types over time for treated and control households.

in the purchased baskets of treated households may have a longer-term influence on their diets if the forced experimentation resulted in households discovering food items that they prefer or due to state dependence in purchase choices.

Lastly, it is worth highlighting that, due to the nature of our variation, we may only speak to the effect of a short-term decrease in store access. The short-term nature of our supply shocks is a desired feature of the careful identification strategy rather than a shortcoming of the paper. Even though we look at store closures lasting for less than a month, we find that supply shocks may explain an economically important part of the nutrition-income gap. Our estimates may be interpreted as a meaningful lower-bound on the effect of long-term grocery inaccessibility.

2.4.2 Limitations

In our setting, exposure to a store closure corresponds with a hurricane. In an attempt to disentangle the effects of the hurricane from the effects of the store closure, control households are selected from the same geographic regions as treated households. However, treated households may have been more heavily affected by the hurricane through mechanisms other than the store closure. The passing of a hurricane may affect household shopping patterns (separately from shocks to supply) through other damages to property or effects on employment. If this is the case, our estimates may overstate the effect of a grocery store closure on nutritional choices. To evaluate these concerns we explore several patterns in the data that help us evaluate whether treated households were disproportionately affected by the hurricane, separately from the store closure.

Our first two exercises aim to confirm that both treated and control groups anticipated the hurricane and were similarly affected based on hurricane strength. First, we examine household expenditure on bottled water right before the hurricane. Bottled water is a staple for hurricane preparation and we expect that both treated and control households to increase their purchases in this category right before the hurricane. Figure 2.5 shows these patterns, where we summarize weekly expenditures and week 0 marks the week of the hurricane. For this graph, data are aggregated at the week level in order to highlight the stockpiling behavior. We see that both treated and control groups prepared for the hurricane, although the treated sample is much noisier because of its smaller size.

Next, we confirm that treated and control households were similarly exposed to the hurricane.

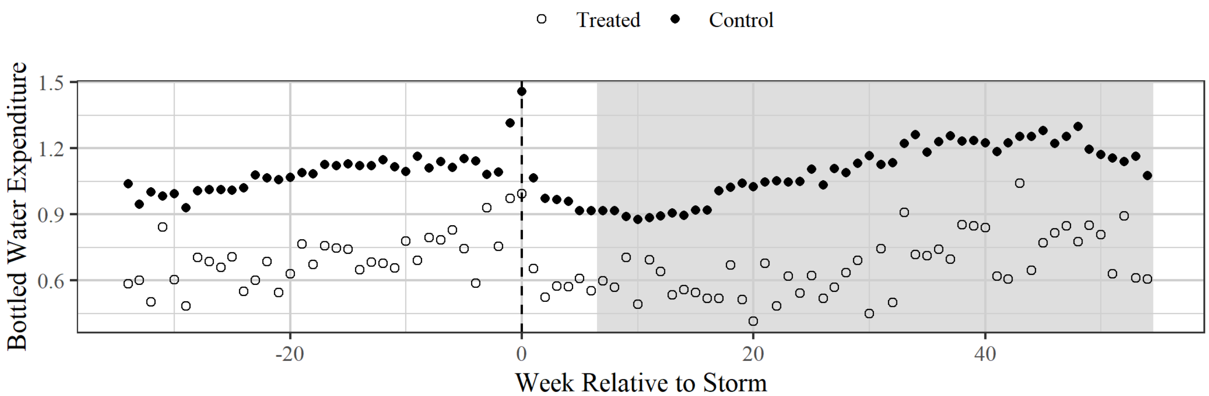


Figure 2.5: **Average Expenditure on Bottled Water over Time.** The graph plots the average weekly expenditure on bottled water, conditional on making a trip in that week. We plot the series separately for treated (not filled dots) and treated households (filled dots). The x-axis tracks the weeks relative to the storm, with week 0 marking the storm and store closure affecting treated households.

HURDAT2 reports the coordinates of the 'center of the hurricane' over time. We look at Zip5-hurricane observations and we calculate the minimum distance between the center of the hurricane (at any given point in time) and each Zip5. Next, mark a Zip5-hurricane as 'treated' if at least one treated household lives in that location. The 'control' Zip5s include locations where we have only control households for that hurricane. We confirm that, conditional on hurricane fixed effects, there is no statistically significant difference in distance from the center of the hurricane across 'treated' and 'control' Zip5s.

In the results section we saw that total food expenditures by treated households remain unchanged relative to control households. Reassuringly, we find that spending on staples such as eggs and dairy remains unchanged after treatment; and spending on private labels remains the same. We also check these patterns at the Zip3 level. For this exercise, we aggregate all food expenditures from any HomeScan panelist to the Zip3 level. We mark a Zip3 as treated if at least one treated household lives there; we mark a Zip3 as control if at least one of our control households lives in that location. For this analysis, we compare expenditures at the Zip3-month (relative to storm) level across treated and control Zip3s. If a Zip3 is more heavily affected by the storm and resulted in an evacuation, we expect to see a decrease in reported spending in those areas. To test this, we compare total expenditures across treated and control Zip3s for the 3 months before and after the hurricane. We do not see any differences across the two groups of locations.

The patterns above suggest that both the treated and control groups are likely similarly affected by the hurricane. As a result, we proceed under the assumption that, apart from the grocery store closure, treated households were not more heavily affected by the storm than the control group. We acknowledge that, if the storm affects treated households more severely in unobserved ways that we may not capture with the above analyses, then our estimates may overstate the effect of the grocery store closure on household diets. Our identification strategy provides a clean way to identify exogenous store closures, but does not allow us to separate potential interaction effects. That is, it is possible that the identified effects only occur when a store closure coincides with a hurricane.

2.5 Conclusion

The analyses in this project increase our understanding on the drivers of households' dietary choices. We analyze how availability may impact changes in the healthfulness of shopping baskets because health practitioner link dietary choices to health outcomes. Combined with the well-documented and increasing nutrition-income gap, one may clearly see the importance for improved understanding of what drives differences in dietary choices across households. Despite the importance of this question, studying the effects of of food accessibility on healthfulness is difficult because food supply is likely correlated with food demand. To circumvent the confounding issues, we use short-term variation in food availability, which occurs right after a hurricane passes through a geographic area. Our supply shocks are defined as temporary closures of grocery stores following a hurricane, which are plausibly exogenous because demand for healthy foods is unlikely to be correlated with the probability of being affected by a grocery store closure right after a hurricane passes through the area.

We identify households as treated if they frequented a closed store in the pre-hurricane period. Our control group consists of households that shop in the same area but did not visit the affected grocery store. We compare total expenditures and the nutritional value of purchase baskets across the two groups, focusing on a 9 month period after the store has reopened. We find that supply-side effects are real in this setting. The temporary closure of a grocery store leads to a meaningful decrease in the nutritional value of household shopping baskets. At the same time, we do not see changes in household expenditures on food items in the post-reopening period.

The finding that a temporary closure of a grocery store leads to persistent changes in household basket composition may have important implications for the relationship between household nutrition and food availability. Thus, we explore whether changes in nutritional value may be related to changes in the types of stores visited by treated households in the post-reopening period. Splitting stores into three groups (grocery, discount, and convenience stores), we ask: do households change where they shop? We see that in the post-reopening period, treated households shift a larger share of their food expenditure away from grocery stores. Store formats differ in their product selections, prices, and locations, which may explain why households change their purchase baskets when shopping at non-grocery stores.

2.6 Data Appendix

2.6.1 Measuring the Nutritional Value of Purchase Baskets

We use nutrition data from USDA's Food and Nutrient Database for Dietary Studies (FNDDS) to describe the nutritional value (NV) of purchase baskets. We use the information to score purchase baskets using an approximation of the Healthy Eating Index, developed by the USDA in order to help practitioners, researchers, and policy makers assess how household food consumption aligns with the Dietary Guidelines for Americans (Krebs-Smith et al., 2018). The HEI scores diets on a scale from 0 to 100, rewarding consumption of 'adequacy' components (fruits, vegetables, legumes, grains, dairy, protein, and fatty acids), and penalizing consumption of 'moderation' components (refined grains, sodium, added sugars, and saturated fats).

Using the HEI to score Homescan transaction data is nontrivial because the USDA and Nielsen databases describe food items at different levels of granularity. For example, nutrient data for raw broccoli can be found in the FNDDS database under the description of 'broccoli, raw.' By our count, raw broccoli purchases in the scanner data are associated with 674 UPCs, representing a large range of products. We, therefore, approximate the HEI using only the fruit, vegetable, and legume food groups. Table 2.7 describes our approximation of the HEI. Columns 1 and 2 describe how the HEI is constructed using different food components. Column 1 shows the units that each component is measured in, with column 2 showing the points awarded per unit. Adequacy components have positive slopes, and moderation components are penalized with negative slopes.

Column 3 indicates whether that component is used in our approximation of the HEI, to which we refer as the nutritional value (NV).

The HEI is nonlinear in its components, for example, consumption of vegetables increases the HEI linearly by around 4.5 points per cup, until quantity consumed reaches a threshold of 1.1 cups per 1,000 kcals. Following Allcott et al. (2019), we construct a linearized version of the HEI, applying the prescribed slope with no maximum cutoff.[42]

We construct our 'partial,' linearized version of the HEI using products identified as fruit, vegetables, or legumes. To do so, we use pattern matching and regular expressions to map the UPC descriptions provided by Nielsen to food descriptions in the FNDDS. For example, UPCs described as "brc f", "brc flrt f", and "brc stalk f" are all matched to the nutrient data for raw broccoli. In the case of multiple matches, UPCs are scored using an average across the nutrient values for each match. For example, a 12 ounce bag described as "brc&cfl flrt f" attributes 6 ounces to broccoli, and 6 ounces to cauliflower. Using this method, we are able to map 88% of expenditure within the relevant product groups to USDA nutrition data. Prior to 2012, items with non-standard UPCs, items such as fruits, vegetables, and in-store baked goods, were coded broadly as "Magnet Data." Following Allcott et al. (2019), we do not attempt to score purchases of these items. However, starting in 2012, Nielsen records these items in distinct product modules with detailed UPC descriptions, which allows us to include them in the construction of our nutritional index. Therefore, measures of the nutritional index may be more complete for households treated during or after 2012.

The included products span the fresh produce, dry grocery, and frozen foods departments. We capture protein, sodium, sugars, and fats through their presence in fruits, vegetables, and legumes, especially for canned, frozen or dry food items.

[42] Allcott et al. (2019) confirm that the linearized and nonlinear versions of the HEI match closely with a correlation of 0.91 at the household year level.

Component	(1) Units	(2) Points per Unit	(3) NI component
Panel A: Adequacy			
Fruits	cups	18.50	Yes
Vegetables[1]	cups	4.55	Yes
Greens and Beans[1]	cups	25.00	Yes
Whole Grains	cups	6.67	No
Dairy[2]	cups	7.69	No
Protein[1,3]	cups	8.25	Yes, partial
Fatty Acids[4]	(PUFAs + MUFAs)/SFAs	7.69	No
Panel B: Moderation			
Refined Grains	ounces	-4.00	No
Sodium	grams	-11.11	Yes, partial
Added Sugars	%kcals	-0.51	Yes, partial
Saturated Fats	%kcals	-1.25	Yes, partial

[1]Includes legumes (beans and peas)

[2]Includes all milk products, such as fluid milk, yogurt, and cheese, and fortified soy beverages.

[3]Includes seafood, nuts, seeds, soy products (other than beverages), and legumes (beans and peas).

[4]Ratio of poly- and mono-unsaturated fatty acids (PUFAs and MUFAs) to saturated fatty acids (SFAs).

Table 2.7: **Nutrition Index Scoring and Descriptives.** This table presents the linearized slopes for each component included in the HEI. Column 1 shows the units that each component is measured in, column 2 shows the points awarded per unit. Column 3 indicates whether the component is included in our nutrition index. Protein, sodium, sugars, and fats are included only if they are found in fruits, vegetables, and legumes. These groups are indicated by ‘partial.’

3 Are Menthol Cigarettes More Addictive? A Cross-Category Comparison of Habit Formation

Abstract: Menthol cigarettes have been banned in parts of the U.S. based on the premise that they are more addictive than non-menthol cigarettes. In this paper, I propose a framework and a novel identification strategy to compare addictiveness across different categories based on consumer panel data. Using variation in the length of temporary breaks in consumption, I compare the effect of past consumption to that of static preferences in driving consumption levels for each cigarette type. I find that demand for menthol cigarettes depends less on past consumption and therefore menthol cigarettes are less addictive. Despite lower addictiveness, menthol cigarettes compare unfavorably to non-menthol cigarettes on other dimensions of addictive behavior: they are harder to quit successfully and more attractive to first-time users.

3.1 Introduction

Smoking is the leading cause of preventable deaths in the U.S., estimated to kill more than 480,000 people annually and to impact 16 million people through smoking-related illnesses.[43] Despite well-known health risks, nearly 13% of Americans currently smoke, a fact which is often attributed to nicotine addiction with roughly half of smokers attempting to quit each year.[44] However, not all tobacco products are necessarily equally addictive. Many believe that flavored cigarettes are more addictive and make smoking more appealing to young people. As a result, Congress passed the Family Smoking Prevention and Tobacco Control Act (TCA) in 2009, which banned flavors in cigarettes with one notable exception: menthol, a mint-flavored additive that alleviates the harshness of smoke. Cigarette consumption has declined by 26% since the passing of the TCA. However, menthol cigarettes, which make up around 30% of the cigarette market, are responsible for less than 10% of that decline (Delnevo et al. (2020a)). The FDA has recently announced its intention to impose a ban on the sale of menthol cigarettes, renewing the debate as to whether menthol cigarettes are more addictive than non-menthol cigarettes (FDA (2022)). In this paper, I propose

[43] https://www.cdc.gov/tobacco/data_statistics/fact_sheets/fast_facts

[44] https://www.cdc.gov/tobacco/data_statistics/fact_sheets/adult_data/cig_smoking

a framework with a novel identification strategy to evaluate the extent to which consumption is different produce categories is driven by addiction. By applying my framework to menthol and non-menthol cigarettes, I can answer the important question of whether menthol cigarette smokers are more addicted than non-menthol cigarette smokers.

I first develop a dynamic model of habit formation that builds on the consumption capital theory (Stigler and Becker (1977)) and is closely related to the framework in Bronnenberg et al. (2012). Following Bronnenberg et al. (2012), consumption levels within a category are modeled as a function of consumers' static preferences and their stock of consumption capital, accumulated from prior consumption. The model allows me to evaluate dimensions of addictive behavior using only two parameters. One parameter, which I refer to as *state dependence*, captures the effect of past consumption on demand, relative to the effect of consumers' static preferences. The second parameter, which I refer to as the *depreciation rate*, governs the speed with which consumption capital accumulates or depreciates. Among other things, these parameters describe how much demand is inflated above consumers' baseline demand (the quantity consumed when consumption capital is equal to zero) and how difficult it is for consumers to decrease consumption.

The key challenge with estimating this model is that state dependence cannot be identified from steady-state consumption levels: high steady-state consumption levels within a category could be consistent with that category being highly addictive (i.e., consumption is highly state-dependent) or with consumers having high baseline demand for that category. To separately identify the parameters of this model, one needs to observe consumer behavior over time outside of steady state. I develop an estimation strategy based on the set of consumers who temporarily pause consumption of a category, hereafter referred to as quitters. Importantly, rather than comparing quitters to non-quitters, who likely differ in their unobserved propensity to decrease consumption, I compare consumers who quit for different lengths of time before returning to the category. The key idea is that consumers who quit for different lengths of time will experience different depletion levels of their consumption capital. When they resume consumption within the category, these differences in their stocks of consumption capital will drive differences in consumption levels. First, I flexibly estimate the effects of an additional period without consumption on post-quit consumption levels using a generalized synthetic control approach. Then, I estimate the dynamic parameters that best rationalize the evolution of these effects over time.

I apply this estimation strategy to Nielsen HomeScan panelists, constructing a sample of menthol cigarette quitters and a sample of non-menthol cigarette quitters. My approach relies on identifying the effects of an additional quit period for each category, which I estimate by comparing post-quit consumption levels of those who quit for one versus two time periods. This provides me with unbiased estimates of the effect of an additional quit period under the assumption that the marginal length of the quit is exogenous. This assumption is supported by various data patterns such as similar pre-quit consumption levels and trends, similar demographic profiles, and similar post-quit purchase behaviors. Interpreting my results as causal, I then estimate the parameters that rationalize these effects over time.[45]

Research in economics and marketing considers a good to be addictive if consumption increases the subsequent demand for that good (Stigler and Becker (1977)). Therefore, the larger the positive effect of consumption on subsequent demand, the more addictive a good is. I find that demand for menthol cigarettes is less state-dependent than demand for non-menthol cigarettes. Therefore, using the above definition, menthol cigarettes are less addictive than non-menthol cigarettes. This result has implications for the difference between baseline demand and steady-state consumption levels: Intuitively, for more addictive categories, the level of steady-state consumption is much higher than the amount consumed upon first entering the category. At steady state, menthol cigarette consumption is only 2.20 times higher than baseline demand, while non-menthol cigarette consumption is 4.18 times higher than baseline demand. However, just focusing on the degree of state dependence and the resulting inflation in demand ignores important intricacies of addictive behavior, such as initiation and cessation rates.

My estimation strategy allows me to compare menthol and non-menthol cigarettes along these other dimensions of addictive behavior. Although menthol cigarette consumption is less state-dependent, i.e., less addictive, I find that, on average, the absolute level of baseline demand is higher for menthol than for non-menthol cigarettes. This result suggests that upon first entering the category, individuals consume menthol cigarettes at higher levels than non-menthol cigarettes. The higher level of baseline demand also suggests that initiation (first-time consumption) through menthol cigarettes is more probable than initiation through non-menthol cigarettes, supporting the

[45]For this application, I define a period as 12 weeks. However, my approach can accommodate different period lengths, so long as purchases plausibly equal consumption at the selected length. The assumption that marginal quit length is exogenous is more plausible for smaller period lengths.

perception that menthol cigarettes are a gateway to smoking.

Higher baseline demand for menthol cigarettes also has implications for smoking cessation rates. There is evidence that menthol cigarette smokers are less likely to successfully quit consumption, despite making more quit attempts and reporting fewer withdrawal symptoms (Foulds et al. (2010); Smith et al. (2014, 2020); Levy et al. (2011)). Guided by this phenomenon, I consider the act of quitting cigarettes in two parts: 1) the likelihood that a consumer will make a quit attempt and sustain it in the short term, and 2) the likelihood that a consumer will relapse (i.e., resume consumption) after a long-term quit attempt. The level of baseline demand influences the likelihood of relapse after a long-term quit attempt, which I define as a quit attempt that has lasted long enough for the stock of consumption capital to depreciate near zero. Using the same logic as discussed in the previous paragraph on smoking initiation, menthol cigarette quitters are more likely than non-menthol cigarette quitters to relapse after long-term quit attempts: When the stock of consumption capital approaches zero, the higher level of baseline demand for menthol compared to non-menthol cigarettes tempts menthol cigarette quitters more to resume consumption than non-menthol cigarette quitters. Therefore, traditional quitting techniques which target the stock of consumption capital, such as sustained abstinence and nicotine replacement therapy, may be less effective for menthol cigarette quitters because the higher likelihood of eventual relapse is driven by higher levels of baseline demand rather than a larger stock of consumption capital.

My model of consumption relies on two parameters: state dependence and the depreciation rate. Thus far, the findings discussed follow from just the estimated degree of state dependence. The depreciation rate of consumption capital has implications for both the likelihood of sustaining a quit attempt in the short term and for the long-term effects of a temporary break in consumption. I estimate a higher depreciation rate for menthol than for non-menthol cigarettes.[46] Therefore, during a quit attempt, consumption capital depreciates more quickly for menthol cigarette quitters. This result, combined with the lower degree of state dependence, implies that sustaining a quit attempt, in the short term, is easier for menthol cigarette quitters than for non-menthol cigarette quitters. This is because withdrawal symptoms, which can be thought of as an increasing function of

[46] I am agnostic as to whether the difference in depreciation rates is driven by psychological or physiological differences in category consumption. Regarding the latter, there is some evidence within the biological nicotine literature that suggests that menthol can change the neurological effects of nicotine which could effect the importance placed on recent vs. past consumption (Wickham (2021)).

consumption capital's contribution to demand, should be less extreme for menthol cigarette quitters than for non-menthol cigarette quitters. However, once a relapse occurs, the higher depreciation rate and lower degree of state dependence work against menthol cigarette quitters by causing consumption to revert to the steady-state level more quickly: I find that a temporary quit attempt has smaller effects on long-term consumption for menthol cigarette smokers than for non-menthol cigarette smokers.

A ban on menthol cigarettes in the U.S. is controversial because the policy is likely to have the largest impact on Black smokers, nearly 85% of whom smoke menthol cigarettes compared to just 29% of white smokers (Delnevo et al. (2020b)). Public health experts and civil rights groups argue that the disparate impact of the ban would be positive, addressing the disproportionate harms of menthol cigarettes, which are believed to be more addictive, on Black Americans. However, opponents of the ban argue that the disparate impact of the ban would be negative, purporting that it would be discriminatory to ban products that Black Americans simply prefer. In short, the proponents of the ban evoke addiction, and the opponents of the ban evoke preferences, as the predominant driver of menthol cigarette consumption among Black Americans. My results contribute to this debate by showing that menthol cigarette consumption is not driven by addiction any more than non-menthol cigarette consumption is.

This paper draws upon the literature on state dependence in marketing and economics. Broadly speaking, state dependence refers to any context in which an individual's current choice or circumstance depends on some state variable (Heckman (1981)). In marketing, state dependence is typically studied in the context of brand choice, with a focus on estimating the extent to which persistent brand choices are driven by state dependence as opposed to time-invariant preferences (Dubé et al. (2010a); Levine and Seiler (2022)). More generally, we can think about addiction and habit formation as forms of state dependence, operating at the category, rather than brand, level and affecting the intensity of consumption rather than the identity of what is consumed (Stigler and Becker (1977); Iannaccone (1986); Tuchman (2019); Gordon and Sun (2015)).

Addiction research in economics largely stems from the theory of rational addiction, developed by Becker and Murphy (1988). The implications of rational addiction theory have been tested in contexts such as cigarettes, alcohol, and caffeine (Becker et al. (1994); Olekalns and Bardsley (1996); Baltagi and Griffin (2002); Chaloupka (1991)). However, these reduced-form tests of rational addiction theory do not provide researchers with the tools to evaluate different policies regulating consumption of addictive goods. Recent research in addiction addresses this gap by building dynamic structural models of consumption and considering the implications of addiction within a focal category for different tax policies (Gordon and Sun (2015); Kim and Ishihara (2021)). These methods require the researcher to reduce consumption into a set of discrete choices over quantities, to observe or impute price data, and to remove or account for stockpiled purchases. All of these requirements make cross-category comparisons difficult because stockpiling behaviors and preferences for different quantity options may vary across product types. Furthermore, prices are not observed in commonly used databases (e.g., Nielsen HomeScan) and imputing prices introduces measurement error and often requires a restriction of the sample size. To the best of my knowledge, there does not exist a scalable framework to study the relative addictiveness of different product categories – this paper fills that gap.

Lastly, this paper contributes to the empirical literature that uses consumption capital theory to explain heterogeneity in consumption choices (e.g., Bronnenberg et al. (2009, 2012, 2021); Sudhir and Tewari (2015)). Bronnenberg et al. (2012) show that a large share of geographic variation in brand market shares can be explained by past experiences. Bronnenberg et al. (2021) demonstrate that a large share of the generational differences in consumption of craft beer can be explained by differences in the historical availability of craft beer. To the best of my knowledge, this paper is the first one to apply consumption capital theory at the category level to model the intensive margin of how much people consume–rather than the extensive margin of whether they purchase a brand or product type. This paper also utilizes a new source of identifying variation: rather than geographical or historical variation in supply-side factors, I use variation in the length of temporary breaks in consumption. This variation is broadly available in most product categories, making my approach easily generalizable.

The remainder of the paper is organized as follows. In Section 3.2, I provide a brief summary of the history of menthol cigarettes in the U.S. and discuss some key findings from the literature. In Section 3.3, I outline my empirical framework and identification strategy. In Section 3.4, I describe the data and sample construction before providing preliminary evidence of state dependence in consumption for both menthol and non-menthol cigarettes. I employ my estimation strategy and present results in Section 3.5. In Section 3.6, I discuss the implications of these results for different aspects of addictive behavior and the proposed ban on the sale of menthol cigarettes. I provide concluding remarks in Section 3.7.

3.2 Empirical Setting: Menthol and Non-Menthol Cigarettes

The U.S. tobacco industry earned $4.53 billion in revenue in 2019, the lowest amount since 2000.[47] Cigarette consumption has been declining steadily since the 1980s in the U.S. (Hoffman et al. (2019)). However, a disproportionately large share of recent declines in consumption is attributed to non-menthol cigarettes (Delnevo et al. (2020b)).

The 2009 passing of the TCA, banning all flavors except for menthol, sparked an abundance of research into the properties of menthol (Delnevo et al. (2020b)). Much of this research suggests that menthol cigarettes are more addictive than non-menthol cigarettes: Studies show that young adults are more likely to try a menthol cigarette as their first cigarette, and that those who start by smoking menthol cigarettes are more likely to continue smoking (Kreslake et al. (2008); Villanti et al. (2016)). Furthermore, several studies find that quit rates are lower among menthol cigarette smokers than non-menthol cigarette smokers, despite more quit attempts (Foulds et al. (2010); Smith et al. (2014, 2020); Levy et al. (2011)). Laboratory testing on rats has shown that menthol can even enhance the neurological effects of nicotine (Biswas et al. (2016)).

Based on the premise that menthol cigarettes are more addictive, menthol cigarettes have been banned in several countries, including Brazil, Canada, and the UK, as well as in some U.S. states and cities. Attempts to ban menthol cigarettes at the federal level in the U.S. have sparked debate due to their disproportionately high rates of use among Black Americans, with opponents fearing that a ban would increase interactions between Black Americans and law enforcement if a black market emerged.[48] The difference in the popularity of menthol cigarettes among white and Black Americans is often attributed to the history of tobacco marketing specifically targeting the latter. Following immense public pressure, the tobacco industry developed The Cigarette Advertising Code in 1964, marking the end of cigarette advertising directed towards the youth demographic

[47] https://www.ftc.gov/news-events/news/press-releases/2021/10/ftc-report-finds-annual-cigarette-sales-increased-first-time-20-years

[48] Although the FDA emphasized that the prohibition is not meant to be enforced against individual smokers, many remain concerned about the possibility of a black market. The 2014 fatal arrest of Eric Garner, a Black man suspected of illegally selling "loosies" (single cigarettes from a pack), was used to defeat a menthol ban in NYC in 2019 (Goodman (2019)).

and a pivot towards Black communities. The tobacco industry increased marketing efforts through discounting menthol products in Black neighborhoods, abundant advertising, sponsorship, and financial support of Black leaders, including the NAACP (Wailoo (2021)). A recent review on the biological impact of menthol on dependence emphasizes the role of these marketing tactics in creating racial disparities in smoking-related outcomes: "selectively marketing a more dangerous product to select populations highlights that menthol cigarettes pose not just a public health problem, but a social justice problem as well" (Wickham (2021)).

While the injustice of selectively marketing unhealthy products to marginalized groups is clear, it is not obvious how to remedy this. The targeted marketing of menthol cigarettes to Black communities could have merely affected the extensive margin of menthol cigarette consumption. These marketing tactics, which included giving out free menthol cigarettes, almost certainly contributed to higher initiation of menthol cigarette consumption within Black communities. However, the intensive margin is of particular interest when it comes to evaluating the implications of different policy interventions – upon entering the category, the level of consumption could be determined by static preferences for the category, or the accumulation of consumption capital, i.e., addiction. If consumption is driven by addiction, then a ban that fully removes access would likely benefit menthol cigarette smokers. However, if consumption is driven by static preferences, a ban that only targets the products preferred by Black people could be viewed as discriminatory.

3.3 Model and Identification Strategy

I model consumers' choices of consumption quantity within a category as dependent on past consumption as well as static preferences. I assume the following data generating process for current category consumption:

$$c_t = (1 - \alpha)\mu + \alpha k_t - \epsilon_t, \tag{3.1}$$

where μ represents the consumer's static preferences for how much to consume within a product category. These static preferences capture the influence of all time-invariant demand factors, including consumer characteristics like income and race. Also included in these static preferences are marketing factors, like price and advertising, which are considered to be time-invariant at the category-level. The consumer's stock of consumption capital within the category at time t is denoted by k_t. The state dependence parameter, $\alpha \in [0, 1]$, governs the relative importance of the stock of consumption capital in current consumption quantity.[49] Lastly, ϵ_t is a mean-zero i.i.d. random utility shock drawn by each consumer at each consumption period.[50]

Consumption capital stock is unobserved, but assumed to evolve deterministically as a weighted average of past consumption:

$$k_t = (1 - \delta)k_{t-1} + \delta c_{t-1}, \tag{3.2}$$

where the depreciation rate, δ, represents the importance of last period's consumption as compared to that of last period's stock of consumption capital. Recent consumption contributes more to consumption capital for higher values of δ. Consumption quantity, capital stock, and static preferences are allowed to be consumer-specific, but I omit an i-subscript to simplify notation. Similar to the brand capital model in Bronnenberg et al. (2012), I assume that consumer behavior is characterized by a common set of dynamic parameters α and δ.

For categories in which consumption is state-dependent, this model predicts that consumption will approach a steady-state level that is some factor greater than baseline demand. I define baseline demand as the expected amount consumed when consumption capital equals zero, i.e., $c_{baseline} = (1 - \alpha)\mu$. I can easily solve for the expected level of steady-state demand, $\bar{c}$, by noting that consumption reaches steady state when consumption capital stock reaches steady state. Thus $\bar{k} = \bar{c}$ after rearranging Equation 3.2. Substituting $\bar{c}$ for k_t in Equation 3.1, it follows that the expected steady-state consumption is simply equal to the consumer's static preferences for the category, μ.

$$\bar{c} = \bar{k} = \mu \tag{3.3}$$

Therefore, $1 - \alpha$ represents the ratio between expected baseline demand and steady-state consumption. Intuitively, the higher the degree of state dependence, the greater the ratio between steady-state consumption and baseline demand. For categories for which consumption is not state-dependent, the amount consumed in steady state is equal to the amount consumed without any experience with the category.

Note that expected steady-state consumption, $\bar{c} = \mu$, is independent of both α and δ. This is

[49] Note that state dependence is consistent with both the processes of addiction and habit formation.

[50] Consumption is required to be non-negative, i.e., for the case that $\epsilon_t > (1-\alpha)\mu + \alpha k_t$, consumption is equal zero.

important because there exist infinite combinations of these parameters that could generate the observed steady-state consumption levels. In other words, the observed steady-state consumption levels do not allow for estimates of baseline demand nor predictions about how consumers will react to shocks that nudge them out of steady state. To illustrate this identification issue, I consider two hypothetical categories for which consumers have the same level of steady-state demand, but which vary in their degree of state dependence: Consumption within category a is driven more by past consumption than consumption within category b, i.e., consumption of category a is more state-dependent. I simulate consumption choices for consumers in each category over 50 time periods. Figure 3.1 plots average consumption over time for the hypothetical categories, starting from a consumer's first experience.[51] Starting at first consumption, average consumption for these two groups looks very different. Because consumption in category b, $\theta_b = (0.25, 0.25)$, is less state-dependent, the expected baseline demand (represented by the solid horizontal line) is closer to the steady-state consumption level. Consumption in the category a, $\theta_a = (0.5, 0.25)$, is more state-dependent, so the expected baseline demand (represented by the dashed horizontal line) is further from the steady-state consumption level.[52]

Given panel data on consumers' choices upon first entering a category, one could distinguish between these two hypothetical categories. However, we more frequently have data on consumers' choices after they have had many years of experience with a category. Starting at around $t = 40$, these two categories become indistinguishable, making identification impossible without some assumption about the state of category capital, k_t, at the time of first observed purchase. Alternatively, variation that nudges consumers out of steady state could be used to identify θ. When considering state dependence at the brand level, price promotions are typically used as plausibly exogenous variation that can drive consumers to switch brands. However, at the category level, price variation is more limited and more likely to be correlated with seasonal changes in demand. The following section describes the variation that I use for identification, which is available at the category level.

[51]Technically, the initial state may be exogenous for categories that we first consume as children. Given that the focus of this paper is on cigarettes, I assume that the initial state is endogenously determined and therefore equals expected baseline demand.

[52]Although an example is not graphed, it is worth noting that changing the depreciation rate, δ, also alters consumption behavior prior to steady state. The depreciation rate is independent of baseline demand and steady state, but it affects the amount of time that it takes to travel between the two levels: Higher values of δ mean a faster convergence.

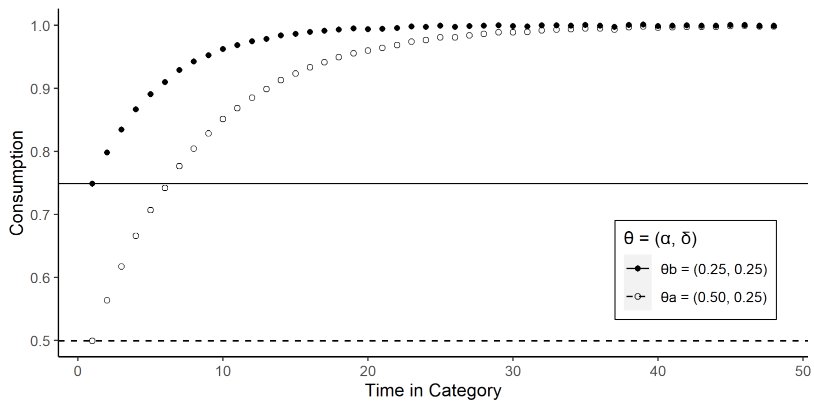


Figure 3.1: **Simulated Choices for Two Hypothetical Categories:** This graph plots average consumption levels for two hypothetical categories, starting at the first consumption experience. Consumption for the category represented by the solid points is less state-dependent (lower α). The horizontal lines represent expected baseline demand, the amount consumed when the stock of consumption capital is equal to 0. Consumption converges to μ for all categories.

3.3.1 Identification Strategy

Consider two identical consumers at steady state, who are randomly assigned to either a treatment or control condition. The consumer in the control condition quits the category for one time period, and the consumer in the treated condition quits the category for two time periods: The "treatment" is an additional period without consumption. Let k_q represent the stock of consumption capital at the time of the quit. When the two-period quitter returns to the category, the value of k_q has depreciated over two periods, as compared with the one-period quitter, for whom the value of k_q has only depreciated over one period. Figure 3.2 visualizes this experiment, with consumption by the two-period quitters marked by the large empty circles, and consumption by the one-period quitters marked by the small black points. The two consumers consume at the same levels pre-quit, followed by a drop to zero for one or two periods, ending at period 1. The grey shaded region represents the period of interest, beginning in the first period that the consumers return to the category. In the first period in which consumption resumes, the two-period quitter should consume less than the one-period quitter if consumption is state-dependent. As consumption capital re-accumulates, the difference between the one and two-period quitters dissolves. Identifying the effect of one additional quit period, i.e., the difference between a one and two-period quitter in the grey region, will allow me to identify the parameters that best rationalize these dynamics.

We can easily formalize this intuition. Recall the expression for current consumption quantity

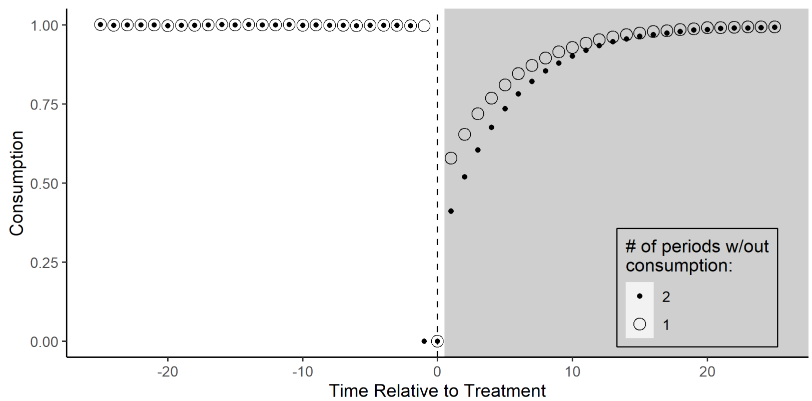


Figure 3.2: **Visualization of Identification Strategy:** This graph plots consumption over time for two identical consumers, assigned to quit for one or two periods. The time without consumption lasts from period -1 to 0 for the two-period quitter, and for period 0 for the one-period quitter. The grey shaded region highlights the period of interest, once both consumers have resumed consumption.

from Equation 3.1. I add notation to indicate whether the consumer was assigned to one or two periods without consumption: the expression $c_t(x)$ (and $k_t(x)$) represents the quantity consumed (and stock of category capital) t periods following the end of the quit, with x indicating how long the individual was assigned to quit for. Quantity consumed for the one and two-period quitters when they return to the category at $t = 1$ is as follows:

$$c_1(1) = (1 - \alpha)\mu + \alpha k_1(1) \tag{3.4}$$

$$c_1(2) = (1 - \alpha)\mu + \alpha k_1(2) \tag{3.5}$$

Let Δc_t (Δk_t) represent the difference in consumption (the stock of consumption capital) between two-period quitters and one-period quitters in the t^{th} period after the end of the quit, which I will hereafter refer to as the t^{th} period post-quit. By subtracting Equation 3.4 from Equation 3.5 we can express the treatment effect on quantity in terms of the latent variable k_t.

$$\begin{aligned} \Delta c_t &:= c_t(2) - c_t(1) \qquad (3.6) \\ &= \alpha \Delta k_t \end{aligned}$$

Intuitively, this suggests that if consumption is highly state-dependent, large α, there should be a large difference between the quantity consumed by one and two-period quitters. Furthermore, it implies that the larger the difference in the stock of consumption capital, Δk_t, the larger the difference in consumption. Below, I show this difference in the stock of consumption capital for the simple case of $t = 1$.

$$\begin{aligned} \Delta k_1 &:= k_1(2) - k_1(1) \\ &= \Big[(1-\delta)[(1-\delta)k_q(2) + \delta c_{-1}(2)] + \delta c_0(2)\Big] - \Big[(1-\delta)k_q(1) + \delta c_0(1)\Big] \\ &= -k_q(1-\delta)\delta \end{aligned} \tag{3.7}$$

By definition, consumption is equal to zero for $t \in \{0\}$ for one-period quitters, and for $t \in \{0, 1\}$ for two-period quitters, i.e. $c_0(1) = c_0(2) = c_{-1}(2) = 0$. Given that both of the identical consumers were at steady-state pre-quit, the stock of category capital at the time of the quit is the same regardless of quit length, i.e., $k_q(1) = k_q(2)$. Therefore, $\Delta k_1 = -k_q(1-\delta)\delta$, and $\Delta c_1 = -\alpha k_q(1-\delta)\delta$. If there is a difference between one and two-period quitters' consumption levels in the first post-quit period, it follows that 1) $\alpha > 0$, meaning that category demand is state-dependent, and 2) $\delta < 1$, meaning that more than just the last period contributes to consumption capital. For the case where $\delta = 1$, the stock of consumption capital would have depreciated to zero for both the one and two-period quitters, resulting in no difference between their stocks of consumption capital.[53]

We can derive a moment condition for each of the T periods observed post-quit, where the difference in consumption is a function of the difference in consumption capital stock at that time. This can be expressed most easily in the first period post-quit, shown above, because the difference in consumption capital stock is driven only by the number of periods that have passed without consumption, i.e., Δk_1 can be expressed in terms of only δ. For $t > 1$, Δk_t is a function of that initial difference, Δk_1, and the amounts consumed post-quit as a result of that initial difference. Using Equation 3.1, we can express the difference in consumption as follows:

$$\Delta c_t = -\alpha(1-\delta)\delta(1 + (\alpha - 1)\delta)^{t-1} k_q. \tag{3.8}$$

[53]Strictly speaking, $\Delta c_1 < 0$ also implies that $\delta > 0$. For the case where $\delta = 0$, only the initial state of consumption contributes to consumption capital, and therefore temporary breaks in consumption would not generate a difference in the stock of consumption capital for either one or two-period quitters. This case is not relevant to the analysis of cigarettes, as we expect some increase in cigarette use over time, and $\delta = 0$ predicts steady-state consumption equal to baseline demand.

3.3.2 Explanations for State-dependent Consumption

Dubé et al. (2010a) consider three explanations for state-dependent brand choice. The baseline explanation, referred to as *loyalty*, is that past choices alter the current utility derived from the current choice, introducing a psychological switching cost that consumers would incur from choosing a different brand (Farrell and Klemperer, 2007). Alternatively, state dependence could arise if consumers face search costs and are thus incentivised to repurchase brands that they've purchased before. Lastly, the authors consider whether consumer learning could explain the observed state dependence.

In the context of category consumption quantity choices, only two of these mechanisms are relevant: loyalty and learning. We can consider consumption capital stock as consisting of addiction and knowledge stock. The loyalty explanation is analogous to the story of addiction: past consumption contributes to addiction stock which alters the utility accrued from current consumption within a category. Alternatively, repeated consumption can contribute to knowledge stock as consumers learn about their preferences for how much to consume within a category. The identification strategy outlined above relies on the depreciation of consumption capital over time. If we assume that addiction stock depreciates and that knowledge stock does not depreciate (i.e., there is no "forgetting" about one's preferences for a category) then this depreciation in consumption capital stock is entirely driven by the change in addiction stock. Therefore, this approach isolates the addiction effect of consumption capital.

Because I am neglecting the learning effect of consumption capital, it is necessary to introduce a more nuanced definition of baseline demand. Baseline demand was previously defined as the expected amount consumed when consumption capital stock is equal to zero. Baseline demand, as identified by this approach, instead represents the expected amount consumed with addiction stock is equal to zero. In other words, given perfect information about preferences for how much to consume within a category, baseline demand is the amount consumed without prior consumption.

3.3.3 Estimation

Estimation proceeds in two steps. I first estimate individual treatment effects in each period post-quit using the generalized synthetic control approach, proposed by Xu (2017). Then, taking those effects as given, I use generalize method of moments (GMM) to estimate the dynamic parameters, $\theta = (\alpha, \delta)$.

In contrast to a standard two-way fixed effects approach, which would estimate one average treatment effect for each post-quit period, the generalized synthetic control approach allows me to estimate a counterfactual quantity consumed for each treated unit in each post-quit period. I consider two-period quitters to be "treated" by the one additional quit period, and use the one-period quitters to impute what they would have consumed if they had only quit for one period. I then subtract observed consumption from the imputed counterfactual consumption to estimate Δc_t for each of the N two-period quitters.

The relationship between the treatment effects, Δc_t, the stock of consumption capital at the time of the quit, k_0, and the dynamic parameters, θ, in each of the T periods is given by Equation 3.8. I assume that consumers are at steady state pre-quit, and therefore k_q is treated as observed, using the pre-quit consumption level for each consumer in accordance with Equation 3.3. With consumer-level treatment effects, i.e., consumption level differences Δc_t, and pre-quit consumption capital stocks, k_0, in hand, I can use Equation 3.8 to estimate the dynamic parameters. Equation 3.9 represents the moment conditions used for GMM, where Δc_t and k_0 are $N \times 1$ vectors. I estimate $\theta = (\alpha, \delta)$, such that the $(T \times N)$-vector $g(\theta, x)$ is equal to zero in expectation.

$$E[g(\theta, x)] \equiv E \begin{bmatrix} \Delta c_1 - \alpha(1-\delta)\delta k_0 \\ \dots \\ \Delta c_T - \alpha(1-\delta)\delta(1+(\alpha-1)\delta)^{T-1} k_0 \end{bmatrix} = 0 \tag{3.9}$$

The moment conditions derived above hinge on the ability to identify Δc_t. To do this, we need to assume that treatment (quitting for an additional period) is not correlated with post-quit consumption levels. In other words, the factors that influence two-period quitters to forgo consumption for one additional period are considered to be exogenous. In a later section, I will show evidence that this assumption holds using standard tests such as comparing the pre-quit trends in consumption and sample compositions. I will also show evidence that depleted consumption capital is driving the post-quit consumption effects (rather than unobservable differences in quit commitment) using post-quit patterns in purchases.

3.3.4 Comparison with Existing Approaches

This framework is closely related to that developed by Bronnenberg et al. (2012), diverging in two ways: the outcome of interest and the identifying variation. Bronnenberg et al. (2012) (Bronnenberg et al. (2021)) apply their framework to study state dependence in the choice between brands (product types). While their approach would be appropriate for understanding why Black Americans prefer menthol cigarettes more than white Americans, it ignores the choice of how much to consume. In light of a potential ban on menthol cigarettes, the intensive margin is of particular interest. My approach also differs in the source of identifying variation: Rather than using geographical or historical differences in supply-side variables, I use variation in quit-lengths to identify the parameters of interest. This identification strategy is appealing because it is highly generalizable, allowing for the application of my framework to any category in which consumers take temporary breaks in consumption.

My approach also diverges from recent research on addiction in economics and marketing (e.g., Gordon and Sun (2015); Kim and Ishihara (2021)). These papers typically estimate dynamic structural models of consumption within a focal category, utilizing purchase-level data, and accounting for stockpiling and expectations over prices. My approach does not require the researcher to observe prices, which I view as an advantage because the commonly used Nielsen HomeScan database does not include them.[54] Furthermore, my approach is less affected by stockpiling behavior: as detailed in a later section, I aggregate over purchases made within relatively large windows of time, assuming that consumption is equal to purchases at this level. While this aggregation allows me to abstract away from concerns of stockpiling, I do so at the cost of meaningful price variation. Therefore, consumers that temporarily quit a category provide one of the few sources of meaningful variation at this level.

3.4 Data

This paper uses the Nielsen HomeScan Panel (HMS) to measure household cigarette consumption over time. I start by describing the database and patterns in cigarette consumption. I then

[54] This limitation is often circumvented by imputing prices using the Nielsen Retail Scanner data; however, such imputation can introduce measurement error and the researcher may need to restrict their sample to only include households that shop at retailers participating in the panel.

outline how the samples of menthol and non-menthol cigarette quitters are constructed and detail limitations of the data that are relevant for the empirical application. Specifically, I discuss the data limitation that panelists report purchases rather than consumption. The section concludes with evidence of the identifying variation and a discussion of the plausibility of the assumption that quit lengths are exogenous.

3.4.1 Cigarette Purchase Data

The HMS database contains information from 194,551 unique households between the years of 2004 and 2019 with the average (median) household participating in the panel for four (two) years. Households are demographically representative and geographically dispersed across the continental United States. For each household, I observe the date of each shopping trip made across all retail outlets. Nielsen assigns retailers to one of 66 mutually exclusive channel types. The broad range of included channels is particularly important for my empirical application because survey data suggests that convenience stores capture a large share of the cigarette market.[55] For Nielsen panelists, convenience stores have been the primary cigarette purchase channel since 2006.

I observe quantities and prices paid for all products purchased during each shopping trip, identified by their unique UPC. Nielsen provides data users with a products file, which details the numeric size and the unit of measurement for each UPC, as well as a description of the product's department, group, and module. This file includes information on over 5.5 million UPCs, with 35,643 UPCs belonging to product group code 4510, described as "Tobacco and Accessories." Within this group, the focal module is "Cigarettes" (product module code 7460).[56]

I classify cigarette UPCs as either menthol or non-menthol cigarettes using the flavor descriptions provided within a supplemental product attributes file.[57] Between the years of 2004 and 2019, menthol cigarettes accounted for 32% of cigarette sales in the data, slightly increasing over time. Figure 3.3 shows the shares of Nielsen panelists who report cigarette purchases in each category over time. The dashed vertical line in 2009 marks the passing of the TCA. Prior to the passing, the

[55] https://www.npr.org/sections/health-shots/2014/02/05/272105414/most-smokers-dont-buy-their-cigarettes-at-cvs

[56] Notably, this product group also includes anti-smoking products, which I will use to control for unobservable differences between consumers with different quit lengths in future analyses.

[57] I classify a small share of UPCs (<0.1% of all sales) as flavored-non-menthol cigarettes if they are associated with a non-menthol flavor. However, since these products have been banned since 2009 ban, these UPCs are likely misclassified and are not included in any analysis.

share of smokers, overall and within both categories, was trending upwards. Following the passing, the share of smokers decreased for both menthol and non-menthol cigarettes. However, it decreased more rapidly for non-menthol cigarettes.

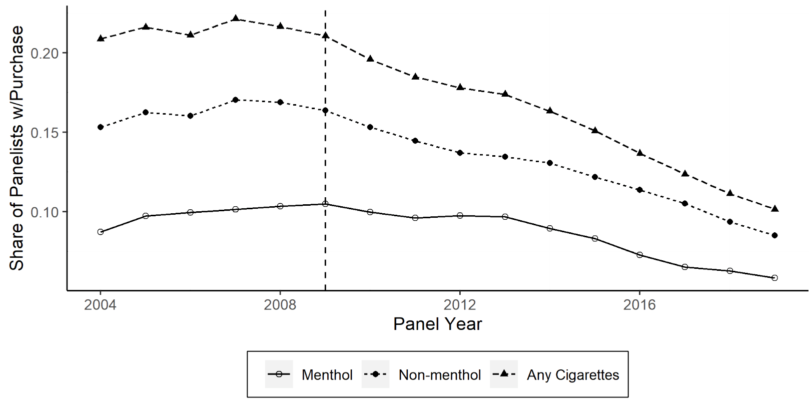


Figure 3.3: **Smoker Shares over Time by Category**: This graph plots the share of unique households observed making a purchase within each category in each year. The dashed vertical line marks the passing of the TCA, after which sales of all cigarettes begin to trend downwards.

For all shopping trips made by households who purchase within the cigarette module, I calculate the total number of menthol and non-menthol cigarettes purchased. Panelists report the quantity of each purchased UPC, which can be matched with the products file to calculate the total number of units purchased, measured in "counts."[58] In the U.S., cigarettes are most commonly sold in packs of 20, accounting for 99.9% of observed transactions.[59]

Table 3.1 summarizes cigarette purchases for Nielsen panelists, classified into three groups: non-smokers, menthol cigarette smokers, and non-menthol cigarette smokers.[60] Descriptives for all Nielsen panelists are provided for comparison. Among the 194,551 households in the sample, 13% smoke menthol and 20% smoke non-menthol cigarettes. Column 3 in Table 3.1 shows that menthol cigarette smokers purchase around two packs per month less than non-menthol cigarette smokers. In column 4, I consider how much time passes between the first and last observed cigarette purchase for each group. Consumers who smoke menthol cigarettes are observed making purchases within

[58]One UPC (out of the 6,738 cigarette UPCs that I observed purchase of) is measured in ounces. This is likely an error, and I assume that quantity for this UPC is measured in counts.

[59]The TCA, which banned non-menthol flavors in cigarettes, also banned packs that contain fewer than 20 cigarettes. However, I observe sales for several UPCs containing less than 20 cigarettes in a year after the TCA was passed. Such UPCs account for <0.1% of all observed sales. I assume that these sizes are in error, and round the values up to 20.

[60]These are not mutually exclusive groups: Some smokers purchase both menthol and non-menthol cigarettes. For the time-being, I do not consider complementarities in consumption between the categories.

the category for longer time periods, suggesting lower quit rates.[61]

	(1) # HHs	(2) HH Share	(3) Avg. Units per Wk	(4) Avg. Years Smoking	(5) Avg. Age	(6) % Black
All panelists	194,551				48.39	10.01
Non-smoker	141,538	0.68			47.97	9.95
Menthol smoker	26,312	0.13	43.05	3.08	49.91	15.47
Non-menthol smoker	40,832	0.20	52.39	2.81	49.50	7.36

Table 3.1: **Summary Statistics by Smoker Status:** This table presents summary statistics for three types of households, with statistics for all panelists provided as a benchmark. Average age is computed using the first reported value for each household.

I match the cigarette transaction data with Nielsen's annual demographic survey of panelists, which provides variables such as income, employment, age, marital status, and race. Columns 5-6 in Table 3.1 describe select demographic variables for each group. Surprisingly, menthol cigarette smokers do not appear to be younger than non-menthol smokers (column 5), despite the perception of menthol cigarettes as being a popular gateway to smoking.

It is well-documented that menthol cigarette use is more prevalent among Black Americans (Delnevo et al. (2020a)). Column 6 in Table 3.1 shows the share of households within each group who are Black. Black households make up over 15% of menthol cigarette smokers, despite representing only 10% of the panel. As I briefly discussed in Section 3.2, the disproportionate popularity of menthol cigarettes among Black Americans can be explained by consumption capital theory even if menthol cigarettes are not addictive. The marketing of menthol cigarettes to Black communities could have increased initiation rates, with the levels of consumption driven by static preferences.

To provide evidence that the targeted marketing of menthol cigarettes towards Black communities increased initiation rates, I consider the racial gap in menthol cigarette consumption by age. Adults who were exposed to such marketing tactics, which were prevalent in the 1960s and observed as late as the early 2000s (Luke et al. (2000)), would be around 37-77 years old in the most recent year of the data (2019). Figure 3.4 plots the probability of smoking menthol cigarettes conditional on age for Black and non-Black households, using a logistic regression. The difference in menthol

[61]This result is not driven by a difference in the panel participation – there is no significant difference in the average years recorded in the panel between menthol and non-menthol cigarette smokers.

cigarette consumption is larger for older panelists, who were potentially exposed to differential marketing. There is no significant difference in menthol cigarette consumption between Black and non-Black households under the age of 25. This finding suggests that, although menthol cigarettes are more popular among Black Americans, they are not much more popular, if at all, among young Black Americans.

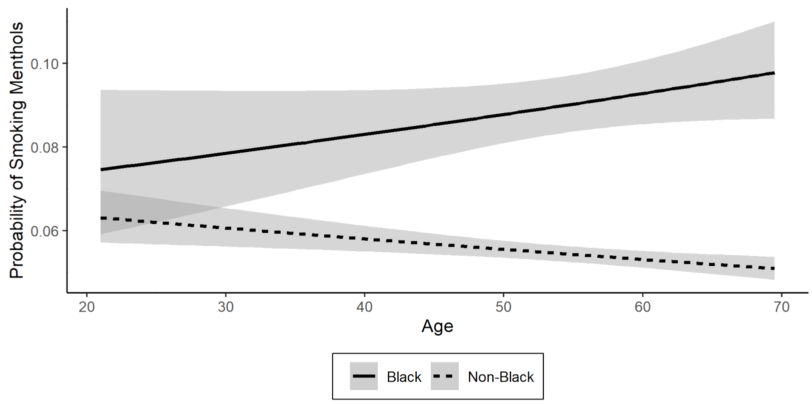


Figure 3.4: **Racial Gap in Menthol Cigarette Consumption by Age:** This graph plots the difference in the probability of menthol cigarette consumption for Black and non-Black households, conditional on age, using data from 2019.

3.4.2 Quitter Sample Construction

To estimate the model described in Section 3.3, I would ideally observe consumed quantities for smokers before and after a quit attempt. This is empirically challenging because panelists only report purchased quantities as opposed to consumed quantities. For the case of perishable food categories, this would likely lead to an overestimation of consumption because I cannot account for food waste. In non-perishable categories like cigarettes, this issue is problematic because of the potential for stockpiling. Panelists may purchase large quantities in preparation for future consumption, which would manifest itself as a form of negative state dependence. To avoid this issue, I aggregate the data to the 12-week (interchangeably referred to as a quarter) level. Consumption rates calculated within a given household-quarter are similar to the rate of consumption using the full panel of that household's reported purchases, with a correlation of $\rho = 0.76$. I therefore consider consumption to equal reported purchases within a 12-week period, and hereafter refer to consumed and purchased quantities interchangeably.

I next set about selecting two groups of quitters: a sample of quitters for menthol cigarettes, and a sample of quitters for non-menthol cigarettes. I define a household to be a potential quitter if there is at least one period (12 weeks) with no reported purchases following six consecutive periods of non-zero purchases. The full sample consists of 13 periods, meaning that potential quitters are required to be observed making a purchase in 46% of periods. I focus on such high-frequency buyers because that is where I can distinguish quitters from individuals that consume erratically. However, this only has a relatively small effect on the generalizeability of my results, as 52% (54%) of expenditure on menthol cigarettes (non-menthol cigarettes) comes from households observed at or above this frequency.

With just the criteria described above, the sample of quitters exhibit a troubling pattern, increasing consumption drastically preceding the supposed quit attempt. I assume that these households are not forgoing consumption during the period without purchases, but rather, they stockpiled and are consuming from their inventory. To exclude these households from the sample, I require that the sum of reported purchases during a window of time including the supposed quit attempt is less than that across a window of the same length preceding it. In other words, I drop households with pre-quit purchase behavior that would allow them to consume at the same rate during the supposed quit as they were pre-quit. For example, I would drop a household with the following purchases: 200 cigarettes at $t \in [-3, -2]$, 400 cigarettes at $t = -1$, and 0 cigarettes at $t = 0$.

I further restrict the sample to households that participated in the panel for the full sample period, six periods before the quit and six periods following resumed consumption. The length of this sample greatly limits the number of selected quitters. For example, a two-period quitter is included if it participates in the 1.5 years before and after the 0.5 year quit, requiring at least 3.5 years of participation in the panel. The nature of the Nielsen panel limits the number of pre and post-quit periods that we can observe. The median panelist participates in the panel for 2 years. Therefore, this restriction drops around 50% of potential quitters.[62]

Lastly, I group quitters based off of how long they quit the category for, retaining one and two-period quitters. Two-period quitters are considered to be “treated” by the additional quit period, and one-period quitters are used as control units to estimate the effect of one additional

[62]In the future, I plan to replicate my current analyses with a shorter sample period and/or an unbalanced panel.

	(1)	(2)	
	Total	Quitters	
	HHs	1	2
Menthol	26,312	651	140
Non-menthol	40,832	1,258	279

Table 3.2: **Number of One and Two-Period Quitters for Menthol and Non-menthol Cigarettes**: This table shows the number of households without purchases for one or two periods following six consecutive periods with purchases. Quitters are dropped from the sample if they appeared to have stockpiled before the quit attempt. Households are required to participate in the panel for six periods before and after the quit.

period without consumption. I drop all quitters that successfully quit (were never observed making another purchase in the category) or that quit for longer than two periods.[63] Column 2 in Table 3.2 shows the number of one and two-period quitters that satisfy these requirements for menthol and non-menthol cigarettes. Column 1 shows the total number of households within each category as a benchmark. I create a panel of consumed quantities in the six periods before the initiation of the quit and after the end of the quit.

3.4.3 Descriptives

I start by showing evidence that two-period quitters have different post-quit consumption levels than one-period quitters using a two-way fixed effects model, specified as follows:

$$c_{it} = \gamma_t + \beta_i + \mathbf{1}(Treated_i = 1) \sum_{t=1}^{T} \phi_t \times \mathbf{1}(TimePost = t)] + \epsilon_{it}, \tag{3.10}$$

where $\mathbf{1}(Treated_i = 1)$ denotes a dummy that is equal to one for a two-period quitter, and zero for a one-period quitter. Consumer and period fixed effects are denoted by β_i and γ_t, respectively. The coefficients of interest, ϕ_t, capture the difference in one and two-period quitters' consumption levels in each post-quit period. The error term is denoted by ϵ_{it}. I cluster standard errors at the household level.

Figure 3.5 plots $\hat{\phi}_t$ in each post-quit period for menthol and non-menthol cigarettes. Note that ϕ_t is the two-way fixed effects model analogue to the treatment effects derived in Section 3.3.1, with

[63]Quitters that return to the category at a later time can be used in future analyses to add power.

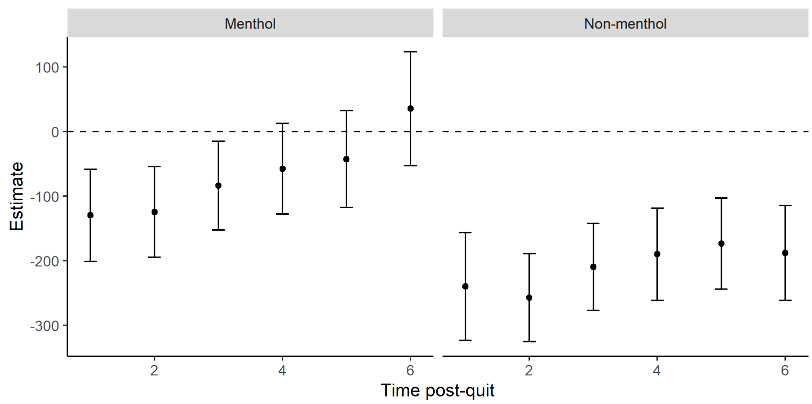


Figure 3.5: **Coefficients from Two-way FE Model:** This graph plots the coefficient on an interaction between treatment (two-period quitter) and each period post-quit, ϕ_t, for menthol and non-menthol cigarettes. Standard errors are clustered at the household level. Each treatment effect is a function of the dynamic parameters, θ, and the stock of consumption capital at the time of the quit attempt, k_q.

Equation 3.8 defining each post-quit effect as a function of the dynamic parameters, θ, and the stock of consumption capital at the time of the quit attempt, k_q. Recall that the first treatment effect, $\phi_1 = -\alpha k_q(1-\delta)\delta$, tells us two things: if $\phi_1 < 0$, this means that $\alpha > 0$ and $\delta < 1$. I find that consumption in the first post-quit period by two-period quitters is significantly lower than consumption by one-period quitters for both menthol and non-menthol cigarette quitters. It follows that $\alpha > 0$ for both categories, meaning consumption in both categories is state-dependent. Importantly, $\delta < 1$ also holds, meaning that more than the last period contributes to consumption capital, which is not trivial given that each period is 12 weeks long. Therefore, I have sufficient variation to identify these dynamic parameters using the estimation method laid out in Section 3.3.3.

Exogeneity of Quit Length Consumers of addictive goods are likely making deliberate and persistent efforts to decrease consumption. Therefore it is possible that there are unobservables correlated with treatment (whether a consumer quits for two periods) and post-quit consumption levels. For example, if a two-period quitter was able to quit for that additional period through the support of nicotine gum, they may still be using nicotine gum after they return to the category. This would result in the magnitudes of the estimated treatment effects being biased upwards, as two-period quitters not only have a lower stock of consumption capital, but also have unobservable factors depressing demand. In this section, I provide evidence that the variation in post-quit consumption, shown above, is coming from a difference in the stock of consumption capital rather than a difference in time-varying consumption trends. I start with a comparison of sample compositions, looking at demographic and purchase behavior variables in the pre-quit period. Then, although the main analysis uses the generalized synthetic control method to match the pre-trend consumption levels and trends, I show evidence supporting the parallel trends assumption using the raw data. Lastly, I discuss post-quit purchase behavior that supports variation in the stock of consumption capital as the mechanism behind the difference in consumption rather than unobservable differences in the commitment to decrease consumption.

Table 3.3 compares the one and two-period quitters on several demographics, as well as on metrics of pre-quit purchase behavior. Columns 1 and 3 (2 and 4) present the average (standard deviations of) pre-quit values for two and one-period quitters, respectively. Column 5 shows the difference in means between two and one-period quitters, with column 6 presenting p-values indicating whether this difference is significantly different from zero. For both menthol and non-menthol cigarette quitters, there are no significant differences in demographics or in pre-quit purchase behaviors between one and two-period quitters: Notably, the total quantity and expenditure in the category is not significantly different between groups.

I conduct an additional set of tests comparing sample compositions by regressing a treatment indicator on three sets of variables: a set of demographic variables, a set of purchase behavior variables, and the union of demographic and purchase behavior variables. The demographic and purchase behavior variables are those presented in Table 3.3. For each specification, I conduct a joint hypothesis test and am unable to reject the null hypothesis that those variables are jointly significantly different from zero. Therefore, I conclude that one and two-period quitters do not differ based on demographics or pre-quit behaviors.

I next show evidence supporting the parallel trends assumption. Although my analysis employs the synthetic control approach, which creates a control group with similar pre-trends to the treated group, differences in pre-quit consumption trends may reflect important differences in post-quit consumption trends. Table 3.4 shows the results of a regression using only pre-quit data. For each category, I regress consumption on an indicator for whether the household quit for two periods, that indicator interacted with a linear time trend, and period fixed effects. Standard errors are

Menthol Cigarettes

	(1)	(2)	(3)	(4)	(5)	(6)
	μ_{q2}	σ_{q2}	μ_{q1}	σ_{q1}	$\mu_{q2} - \mu_{q2}$	p.val
Household Size	2.24	1.21	2.17	1.12	0.07	0.52
Age	55.94	8.66	55.56	8.93	0.38	0.66
Income (000s)	50.17	33.06	50.51	29.39	-0.34	0.91
Years of Education	13.81	1.97	13.65	2.06	0.16	0.40
Marital Status	0.49	0.50	0.47	0.50	0.02	0.72
Black	0.20	0.40	0.27	0.45	-0.07	0.08
Wkly. Work hours	20.85	16.23	22.09	15.82	-1.24	0.41
Packs in category	40.69	36.41	44.28	35.70	-3.59	0.29
Expenditure in category	166.73	161.89	180.82	161.35	-14.09	0.36
Total expenditure (000s)	2.31	1.38	2.52	2.19	-0.21	0.29
Total trips	55.41	35.42	53.69	28.65	1.72	0.55

Non-menthol Cigarettes

	(1)	(2)	(3)	(4)	(5)	(6)
	μ_{q2}	σ_{q2}	μ_{q1}	σ_{q1}	$\mu_{q2} - \mu_{q2}$	p.val
Household Size	2.31	1.24	2.30	1.15	0.02	0.85
Age	55.48	9.38	55.81	9.62	-0.33	0.63
Income (000s)	49.62	31.60	50.00	34.26	-0.38	0.88
Years of Education	13.60	2.04	13.48	2.18	0.12	0.44
Marital Status	0.59	0.49	0.64	0.48	-0.05	0.16
Black	0.05	0.22	0.04	0.19	0.01	0.31
Wkly. Work hours	20.74	16.00	18.68	16.55	2.06	0.08
Packs in category	51.60	53.29	49.79	45.97	1.81	0.59
Expenditure in category	192.31	226.39	182.17	174.69	10.14	0.44
Total expenditure (000s)	2.37	1.38	2.30	1.26	0.07	0.44
Total trips	53.61	31.66	51.76	28.94	1.85	0.38

Table 3.3: **Pre-Quit Comparison of One and Two-period Quitters in the Cigarettes Category**: One and two-period quitters are compared based on their reported demographic variables in the year of the quit, and the average values of purchase behavior variables pre-quit.

	Menthol	Non-menthol
Treated	-98.077	24.863
	(95.553)	(100.575)
Treated x	-3.084	-1.335
Time trend	(10.467)	(12.325)
# Obs	4,746	9,222
R^2	0.012	0.011

$^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$

Table 3.4: **Testing for Parallel Trends:** "Treated" represents the difference in pre-quit levels of consumption between one and two-period quitters. "Treated x Time trend" represents the difference between one and two-period quitters in pre-quit linear consumption trends.

clustered at the household level. For both categories, there is no significant difference in pre-quit consumption levels between one and two-period quitters, indicated by the null result on the treatment indicator. Furthermore, I find no significant difference between one and two-period quitters in pre-quit consumption trends, indicated by the null result on the treatment indicator interacted with the time trend.

The analyses discussed above use pre-quit behavior and time-invariant household variables to show that one and two-period quitters are similar. However, the main concern is that the nature of the quit attempt varies by group. Even if one and two-period quitters are identical in the pre-period, its possible that two-period quitters quit for longer because of unobservable differences in their commitment to quitting, such as a better support system or the use of nicotine gum. Such quitting aids likely persist into the post-quit period, depressing consumption levels beyond the effect of forgoing consumption for an additional period. Therefore, estimates of the effect of one period without consumption would be biased upwards, ultimately biasing the estimates of θ.

As shown in Figure 3.5, the effect of an additional quit period on non-menthol cigarette consumption persists for the entire sample period. To investigate whether a mechanism other than depletion of consumption capital is driving this decrease, I consider purchased quantities after the quit. In contrast to consumed quantities, where I have aggregated purchases over all shopping trips within a quarter, purchased quantities may be more likely to reflect deliberate attempts to decrease consumption. I define purchased quantities as the total quantity purchased during each shopping trip relative to the quit, conditional on a purchase within the category (menthol or non-menthol

	(1)	(2)	(3)	(4)
	Non-menthol		Menthol	
	Quantity	IPT	Quantity	IPT
P1×Treated	-16.596	2.300***	-32.901**	2.833**
	(23.925)	(0.778)	(13.718)	(1.213)
P2×Treated	0.936	1.561**	-8.867	1.714
	(31.351)	(0.789)	(11.354)	(1.167)
P3×Treated	-17.456	0.514	-10.171	1.462
	(19.041)	(0.595)	(16.471)	(0.999)
# Obs	21,299	19,618	10,927	10,058
R^2	0.632	0.216	0.640	0.236

***$p < 0.01$, **$p < 0.05$, *$p < 0.1$

Table 3.5: **Post-quit Purchase Behavior**: This table shows the results of a two-way FE model, presenting coefficients on the interaction between treatment and trip relative to the quit. I present results for both purchased quantity (time to next purchase) during each of the first three trips after the quit with a purchase within the category. P1 represents the first trip with a purchase within the category post-quit, and so on.

cigarettes). There is a vast literature in behavioral economics that shows that people will restrict their choices as a way to commit to achieve their goals; smokers attempting to reduce consumption may do so by purchasing less at the store (Bryan et al. (2010)). I test whether two-period quitters are purchasing in significantly lower quantities that one-period quitters using a two-way fixed effects model, regressing purchased quantities on individual and time fixed effects, and an interaction between a treatment indicator (being a two-period quitter) and time. In this case, "time" represents the n^{th} shopping trip with a purchase within the focal category relative to the quit. Column 1 of Table 3.5 presents the coefficients on the interaction between treatment and time for the first three purchase occasions after the quit. I find that there is no significant difference between one and two-period non-menthol cigarette quitters in the quantity purchased after the quit. Column 2 of Table 3.5 presents the results from the same specification, using time between purchase as the dependent variable. Two-period quitters, for both categories, take more time between purchases. This suggests that the overall decrease in consumption for non-menthol cigarette two-period quitters is driven by an increase in the time between purchases. These results support that the decrease in consumption, for the non-menthol two-period quitters, is driven by a depletion in consumption capital stock rather than a greater commitment to decrease consumption.

3.5 Estimation and Results

As the first step in estimating the model described in Section 3.3, I estimate the effects of an additional quit period using a generalized synthetic control approach. Table 3.6 summarizes the results for each period relative to the quit, presenting the average treatment effects on menthol and non-menthol cigarette consumption. In both categories, the largest effect is in the first period after a consumer resumes consumption, followed by subsequently smaller differences.

This first step serves as a non-parametric test for the presence of state dependence in consumption, and provides an upper bound for the depreciation rate of consumption capital. Recall from Section 3.3.1 that $\Delta c_1 = -\alpha(1-\delta)\delta k_q$. Therefore, the result that $\Delta c_1 < 0$ for both menthol and non-menthol cigarettes tells us that $\alpha > 0$ and $\delta < 1$. Consumption within both categories is state-dependent, with more than the last period of consumption contributing to the stock of consumption capital.

Period post-treatment	Menthol	Non-menthol
1	-112.378*	-229.452***
	(64.134)	(81.013)
2	-108.122*	-217.016***
	(59.236)	(58.803)
3	-67.297	-172.054***
	(58.119)	(58.745)
4	-41.657	-149.049**
	(60.838)	(64.176)
5	-26.489	-137.291**
	(66.998)	(67.351)
6	50.258	-170.342**
	(80.998)	(78.363)
observations	1,960	3,906
units	140	279

$^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$

Table 3.6: **ATTs from Synthetic Control**: This table shows the average difference in consumption between a two-period quitter and their synthetic control unit in each period post-treatment. Standard errors come from 250 bootstrap draws.

To gain further insight into the values of α and δ, I now impose some structure. The moment conditions defined in Equation 3.9 relate the treatment effects, Δc_t, to the stock of consumption capital at the time of the quit, k_q, and the dynamic parameters, $\theta = (\alpha, \delta)$. The first step, described above, provides me with estimates of Δc_t. To estimate k_q, I assume that all quitters are at steady state during the six periods preceding the quit attempt: rather than k_q being equal to $(1-\delta)k_{q-1}+\delta c_{q-1}$ (following Equation 3.2), I assume that $k_q = \bar{k}$. In accordance with Equation 3.3, this steady state value of consumption capital stock is equal to expected consumption. Therefore, I take the average value of pre-quit consumption for each quitter as an estimate of k_q.

Using the estimates of Δc_t from the synthetic control comparison and taking quitters' average pre-quit consumption quantities as estimates of k_q, I estimate the dynamic parameters, $\theta = (\alpha, \delta)$, with GMM applied to the moments defined in Equation 3.9. Recall that, for each quitter, I have six post-treatment periods. Therefore, for menthol cigarettes, $g(\theta, x)$ is 6×140, and for non-menthol cigarettes $g(\theta, x)$ is 6×279.

The estimates of the dynamic parameters are presented in the top panel of Table 3.7. There are two main results: $\hat{\alpha}_{menthols} < \hat{\alpha}_{nonmenthols}$ and $\hat{\delta}_{menthols} > \hat{\delta}_{nonmenthols}$. These results suggest that the demand for menthol cigarettes depends less on the stock of consumption capital, which accumulates and depreciates more quickly than that for non-menthol cigarettes. In the following section, I discuss implications of these results with the aid of one additional estimate. Under the assumption that the quitters are at steady state in the periods preceding the quit attempt, their average pre-quit consumption level serves as an estimate for μ (Equation 3.3), which reflects their static preferences for the category. The bottom panel of Table 3.7, marked by "Static Preferences" presents the average value of μ for each category, with $\hat{\mu}_{menthols} < \hat{\mu}_{nonmenthols}$. This simply reflects that, on average, consumption levels are lower for menthol cigarettes than for non-menthol cigarettes.

3.6 Discussion

Proponents of a ban on the sale of menthol cigarettes argue that menthol cigarettes are more addictive than non-menthol cigarettes. I first consider whether my results support this argument, using the economic definition of an addictive good as one for which past consumption increases subsequent demand (Stigler and Becker (1977)). I then compare menthol and non-menthol cigarettes along three additional dimensions of behavior – comparing the likelihood that consumers begin smoking, attempt to quit smoking, and resume smoking after a quit attempt. I conclude by considering the implications of my results for the proposed ban on the sale of menthol cigarettes.

Dynamic Parameters		
	Menthol	Non-menthol
α	0.542***	0.762***
	(0.169)	(0.118)
δ	0.431**	0.341*
	(0.213)	(0.186)
Treated units	140	279

***$p < 0.01$, **$p < 0.05$, *$p < 0.1$

Static Preferences		
	Menthol	Non-menthol
μ	387.96	572.40

Table 3.7: **Results**: The top panel of this table presents the results from the second stage of estimation, taking treatment effects from the synthetic control as given, and estimating θ using GMM. The bottom panel of this table presents the average pre-quit levels of consumption, which I take as estimates of μ.

3.6.1 Comparisons of Addictiveness

I first discuss the implications of menthol cigarette demand being less state-dependent than non-menthol cigarette demand, i.e., $\hat{\alpha}_{menthols} < \hat{\alpha}_{nonmenthols}$. Recall from Section 3.3 that expected baseline demand, the amount consumed when consumption capital stock is equal to zero, is equal to $(1-\alpha)\mu$. In contrast, the expected steady state level of consumption is μ. I use the ratio of these two values, $\frac{1}{1-\alpha}$, to express how much higher steady-state demand is than baseline demand. Plugging in my parameter estimates, I find that consumption for menthol cigarettes is 2.18 times greater than baseline demand, whereas consumption for non-menthol cigarettes is 4.20 times greater than baseline demand. In short, demand for menthol cigarettes is less state-dependent (less addictive by the economic definition) and therefore less inflated above baseline demand.

3.6.2 Comparisons of Initiation, Quit, and Relapse Rates

The level of state dependence does not show the full story. In this section, I describe the model's predictions regarding the likelihood of three behaviors: initiation, quit attempts, and relapses. Menthol cigarettes are often referred to as a gateway to smoking, suggesting that individuals are more likely to enter the cigarette category through menthol cigarettes than through non-menthol cigarettes. To investigate this, I compare the likelihood of beginning consumption, hereafter referred to as initiation, for each category. For current smokers, a prevalent result from the menthol literature is that menthol cigarette smokers are less likely to successfully quit smoking, despite initiating more quit attempts. I investigate this phenomenon in two steps, first considering the likelihood of making a quit attempt, and then considering the likelihood of resuming consumption (relapsing) after different lengths of time without consumption.

I start by deriving a simple expression for the probability of consumption. Recall the data generating process for consumption, defined in Equation 3.1: Consumption is positive if $(1-\alpha)\mu + \alpha k_t \geq \epsilon_t$. For the purpose of easily comparing initiation, quit, and relapse rates, I assume that $\epsilon \sim \text{Uniform}(a, b)$ for both categories. Therefore the probability of consumption is given by

$$Pr\Big[c_t \geq 0\Big] = \frac{1}{b-a}[(1-\alpha)\mu + \alpha k_t - a], \tag{3.11}$$

where a and b govern the minimum and maximum values of the i.i.d. mean zero shocks to demand, which I assume to be equal for menthol and non-menthol cigarette smokers.[64]

Initiation describes first-time consumption within a category. The stock of consumption capital for a first-time consumer is necessarily equal to zero. Using Equation 3.11, it is simple to show that the difference in the probability of initiation between menthol and non-menthol cigarettes is given by the difference in their respective baseline demands, $(1-\alpha)\mu$, scaled by $b-a$. I use the estimated values of α and μ, presented in Table 3.7, to compare menthol and non-menthol cigarettes in terms of baseline demand. I find that baseline demand is 177.69 cigarettes for menthol cigarette smokers, and 136.23 cigarettes for non-menthol cigarette smokers. In other words, in the first quarter with consumption, a new menthol cigarette smoker would consume around nine packs, in contrast to a new non-menthol cigarette smoker, who would consume around seven packs. It follows that the probability of initiating consumption is greater for menthol cigarettes than for non-menthol cigarettes.

I now turn to the probability of making a quit attempt. Consider a consumer at steady state,

[64] For the comparisons in this section, I assume that my estimates represent the values of the parameters in the full population, rather than just the populations of menthol and non-menthol cigarette quitters. Furthermore, I assume that static preferences, which are allowed to vary by consumer, do not vary systematically between those that enter a category and those that do not, nor between consumers that quit the category and those that do not. As a future robustness check, I plan to compare steady-state consumption levels between households that I classify as quitters and those that I never classify as quitters.

i.e., $k_t = \mu$. For such a consumer, the probability of quitting is simply

$$1 - Pr\Big[c_t \geq 0 | k_t = \mu\Big] = 1 - \frac{\mu - a}{b - a}.$$

It follows that the difference in quit probabilities between menthol and non-menthol smokers is $\mu_{nonmenthol} - \mu_{menthol}$, scaled by $b - a$. Using the estimates of μ from Table 3.7, this difference is positive, meaning menthol smokers are more likely to make a quit attempt than non-menthol smokers.

Lastly, I consider the likelihood of relapse after an n-period quit attempt. Using equation 3.2, I can show that after an n-period break in consumption, a quitter's stock of consumption capital is equal to $(1-\delta)^n k_q$, where k_q again represents the stock of consumption capital at the time of the quit attempt. Assuming that consumers were at steady state pre-quit, I substitute the steady-state value of consumption capital stock, μ, for k_q. Let $\Omega(n)$ represent the expected demand after n periods without consumption, i.e., $\Omega(n) = (1-\alpha)\mu + \alpha(1-\delta)^n \mu$. The probability of consuming at period $n+1$, hereafter referred to as probability of relapse, is given by

$$Pr\Big[c_t \geq 0 | k_0 = \mu, \{c_0 ... c_n\} = 0\Big] = \frac{\Omega(n) - a}{b - a}.$$

The difference in the probability of relapse between a menthol and non-menthol cigarette quitter is the difference in their respective values of $\Omega(n)$, scaled by $b - a$.

I separate quit attempts into "short-term" and "long-term" quit attempts. I define a long-term quit attempt as one in which the consumer has forgone consumption for long enough that consumption capital stock approaches zero, and therefore, demand approaches baseline demand.[65] A short-term quit is any quit attempt before demand approaches baseline demand. Figure 3.6 plots expected demand, $\Omega(n)$, using my estimated parameters, where the x-axis tracks the number of periods without consumption. The vertical lines mark the beginning of a long-term quit attempt for each category. I begin by discussing the relatively static case, when both menthol and non-menthol cigarette quitters are in a long-term quit. When both categories' quitters are in a long-term quit (when $t \geq 21$, or after around five years) menthol cigarette quitters are necessarily more likely to

[65] I define a long-term quit attempt as a quit long enough such that demand is less than 0.1 units from baseline demand.

relapse. This is true through the same logic by which I established that initiation is more likely for menthol cigarettes: The higher level of baseline demand for menthol cigarettes tempts consumption even after a long-term quit.

In the short-term, the likelihood of relapse is driven by the steady-state level of consumption, μ, the degree of state dependence, α, and the depreciation rate, δ. The steady-state level of consumption determines the initial state of consumption capital stock upon first quitting the category. The degree of state dependence determines the ratio between steady-state and baseline demand, and therefore determines the amount that consumption capital needs to depreciate by, before it is near zero. Lastly, the depreciation rate determines how long it takes for consumption capital to approach zero from its steady-state level. It is useful to think of the initial probability of relapse as the complement to the probability of making a quit attempt: Because it is more likely for menthol cigarette smokers to make a quit attempt, it is less likely for them to immediately relapse. Recall that menthol consumption capital stock has a higher depreciation rate, $\hat{\delta}_{menthols} > \hat{\delta}_{nonmenthols}$, meaning that it depends more on recent consumption, and therefore depreciates more quickly during a quit attempt. This means that beyond having less distance to travel, lower μ and higher $\bar{c}_{baseline}$, menthol cigarette consumption capital also depreciates faster. All of these factors work in the same direction, to ensure that a relapse is less likely for menthol cigarette quitters during the first few periods of a quit attempt.[66] However, as demand approaches baseline demand for each category, we see the roles reverse, and relapses become more likely for menthol cigarette quitters. Therefore, menthol cigarette smokers are less likely to successfully quit cigarettes because even after a long-term quit attempt, higher levels of baseline demand tempt them to resume consumption.

In sum, menthol cigarettes have a higher initiation rate, a higher quit attempt rate, and a higher relapse rate.

3.6.3 Comparisons of the Long-term Effects of an Intervention

In this section, I use my parameter estimates to evaluate the effects of a hypothetical intervention that causes smokers of menthol and non-menthol cigarettes to quit for one period. I consider an

[66] Specifically, the probability of relapse for menthol cigarette quitters surpasses that for non-menthol cigarette quitters when the difference in their baseline demands is less than the difference in their stocks of consumption capital, weighted by their respective state-dependence parameters: $(1-\alpha_M)\mu_M - (1-\alpha_{NM})\mu_{NM} \geq \alpha_{NM}(1-\delta_{NM})^n\mu_{NM} - \alpha_M(1-\delta_M)^n\mu_M$

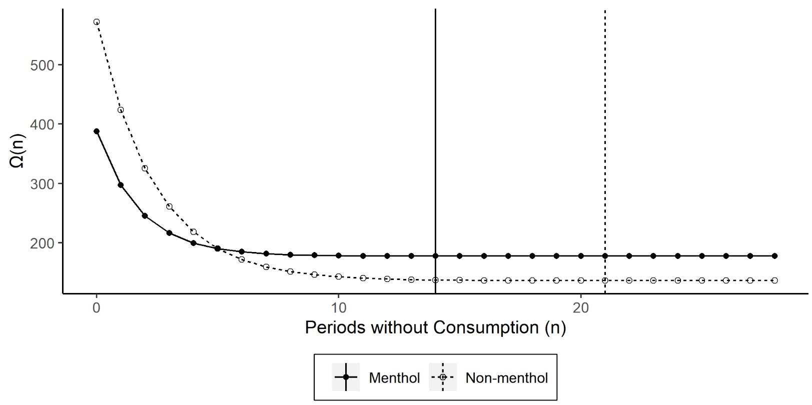


Figure 3.6: **Likelihood or Relapse by Quit-Length** : This graph plots the likelihood of relapse for menthol and non-menthol cigarette quitters by the duration of the quit attempt. Menthol cigarette quitters are more likely to relapse for quit attempts under a year, and more likely to relapse for longer-term quit attempts. The vertical lines indicate the time at which expected demand is within 0.1 units of expected baseline demand.

intervention that occurs at age 50, the average age of a household in my sample. I simulate the effects of this intervention through the age of 80, the average life expectancy in the U.S.

Using Equation 3.8, I can describe the rate at which treatment effects dissipate in terms of θ as follows:

$$\frac{\Delta c_t}{\Delta c_{t-1}} = \frac{\alpha(1-\delta)\delta(1+(\alpha-1)\delta)^{t-1}k_q}{\alpha(1-\delta)\delta(1+(\alpha-1)\delta)^{t-2}k_q} = 1+(\alpha-1)\delta.$$

Intuitively, this metric is decreasing in δ: For higher levels of δ, recent consumption contributes more to the stock of consumption capital causing consumption to revert more quickly to steady state. This metric is increasing in α: If consumption is less state-dependent, a temporary break in consumption will have a short-lived effect because baseline demand is closer to the steady-state level of consumption. Plugging in my parameter estimates, the ratio between the effects at t and $t-1$ is 0.80 for menthol cigarettes and 0.92 for non-menthol cigarettes, meaning that the effect of a one-period quit dissipates more quickly for menthol cigarettes. In Figure 3.7, I plot the simulated effects of a one-period quit over time. The vertical lines mark the time at which the effects of the intervention on current consumption have worn off, occurring after around six years for menthol cigarettes smokers, and after over fifteen years for non-menthol cigarette smokers.

I use the average consumption level, $\hat{\mu}$, for each category to calculate the cumulative effect of this one-time intervention over time. Table 3.8 shows the results of this simulation. Columns 1 and 2 present the cumulative effect of the intervention in terms of packs of cigarettes, and columns

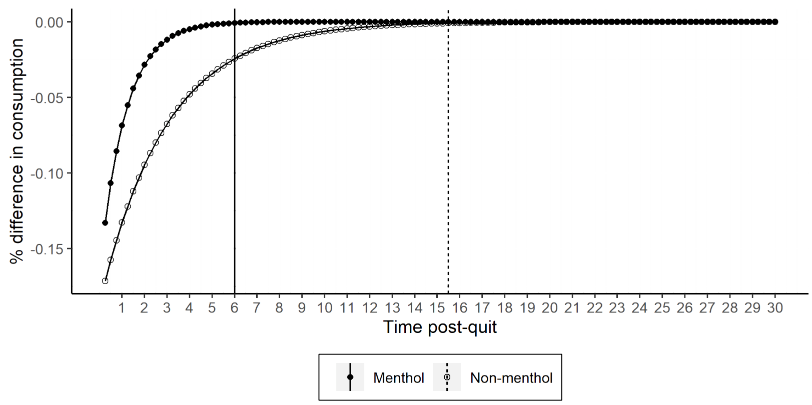


Figure 3.7: **Effects over Time**: This graph plots the effects of a one-period quit on current consumption for thirty years following the quit. The vertical lines indicate when the effects of the quit have worn off. The x-axis labels indicate the number of years since the intervention.

3 and 4 express the cumulative effect in terms of percentages. Columns 5 and 6 uses the average consumption rate to translate this treatment effect into the equivalent number of days without consumption. In the first year following the intervention, menthol (non-menthol) cigarette smokers are predicted to consume 7.64 (17.34) fewer packs than they would have in the absence of treatment, representing around a 10% (15%) decrease. This treatment effect in the first year is equivalent to menthol (non-menthol) cigarette smokers forgoing consumption for 12 (18) additional days post-intervention. Over the course of 30 years, menthol (non-menthol) cigarette smokers are predicted to consume 0.56% (1.76%) fewer cigarettes as a result of an one-period quit. This suggests that a temporary intervention on non-menthol cigarette smokers would be more effective at reducing consumption.

3.6.4 Implications for a Ban on Menthol

Although I do not compare welfare with and without a ban, I use my findings to briefly speculate about its potential impact. The argument that the FDA should ban menthol cigarettes because they are more addictive is not supported by my results: I find that menthol cigarettes are no more addictive than non-menthol cigarettes. However, I find the menthol cigarettes do exhibit properties colloquially associated with addiction, such as low cessation rates, and high initiation rates. These properties in themselves could be grounds for a ban if the goal is to reduce inequities in smoking-related health outcomes for Black Americans. Whether a ban achieves this goal depends on two

	(1)	(2)	(3)	(4)	(5)	(6)
	Difference in packs		Difference in %		Difference in days	
	M	NM	M	NM	M	NM
1 year	7.64	17.34	9.85	15.15	11.82	18.18
5 years	12.90	49.27	3.33	8.61	19.95	51.65
30 years	13.06	60.38	0.56	1.76	20.20	63.29

Table 3.8: **Cumulative Effects of One-Period Intervention**: This table presents cumulative treatment effects from a simulated one-period quit. Results for menthol-cigarette smokers are denoted by an "M", and those for non-menthol cigarette smokers are denoted by a "NM". The difference in packs shows the absolute difference between steady-state consumption and post-intervention consumption levels, scaled by 20. The difference in percentages shows the effect of the intervention relative to the level of steady-state consumption. The difference in days shows the effect in terms of the number of days without consumption at the steady-state rate.

factors that are outside of the scope of this paper: the ability to fully remove access to menthol cigarettes, and the rate at which Black and white youth are picking up menthol cigarettes.

If menthol cigarettes remain accessible, through a black market or gaps in the coverage of the ban, a ban may have modest effects on the current population of menthol smokers because of the lower probability of successfully quitting the category. For consumers that continued to smoke menthol cigarettes, this could have the result of increasing access costs and interactions with law enforcement.

These negative effects on welfare could potentially be offset by the benefit to future generations who do not enter the cigarette category due to the increased cost of menthol cigarettes. Because menthol cigarettes are more attractive to first-time users, a ban on the sale of menthol cigarettes (incomplete or not) could decrease the number of future smokers. However, I find that the difference in menthol consumption rates between white and Black smokers is increasing in age. In Section 3.4 I provide preliminary evidence that there is no difference in menthol cigarette consumption between Black and non-Black households under the age of 25. If there remains access to menthol cigarettes, this is reason to believe that the negative effects of a ban for the current population of Black smokers may not be offset by positive effects for future generations of Black Americans.[67]

[67]There is mixed evidence as to whether menthol cigarettes are more harmful than non-menthol cigarettes. For the purposes of this discussion, I assume that the harm-per-use is equal between menthol and non-menthol cigarettes.

3.7 Conclusion

In this paper, I develop a dynamic model of habit formation that allows current consumption levels to depend on past consumption. The model characterizes consumption dynamics based on a set of easily interpretable parameters, one of which directly measures the degree of addictiveness thus allowing me to assess relative addictiveness across different product categories. I estimate the model based on variation in consumers' consumption capital stock that is driven by consumers who temporarily pause consumption for different lengths of time. By applying this framework to the case of menthol and non-menthol cigarettes, I contribute to the ongoing debate discussing whether menthol cigarettes are more addictive than non-menthol cigarettes. Contrary to common belief, I find that menthol cigarettes are no more addictive than non-menthol cigarettes: Demand for menthol cigarettes is driven less by past consumption than demand for non-menthol cigarettes. Despite this finding, my results support the notion that menthol cigarettes are more attractive to first-time users and more difficult to quit successfully.

The approach I use in this paper is easily scalable, posits relatively few data requirements, and uses broadly available identifying variation. These characteristics make the approach a useful tool for cross-category comparisons of habit formation. There are many contexts in which such comparisons would be interesting and important, both from a public policy and marketing perspective. For example, the model of rational addiction is supported in some categories that are not considered addictive, such as milk, eggs, and oranges (Auld and Grootendorst (2004)). Although my model does not speak directly to rational addiction, this finding suggests that better understanding the role of state dependence across a broad range of categories is important. This information could help retailers arrange their stores such that categories with highly state-dependent demand are prominently featured. Another interesting application of my framework involves comparing the role of state dependence across healthy and unhealthy categories. It is well-documented that lower income individuals have less healthy diets, but there is ongoing debate over whether demand or supply-side differences are the dominant driver of this phenomenon (Allcott et al. (2019); Hristakeva and Levine (2022)). If categories that are less healthy are simultaneously more addictive, then demand-side differences could be the result of supply-side differences. For example, if consumption of fresh produce depends on past consumption (through habit formation), then someone who lived in a food desert should consume less fresh produce than someone who did not live in a food desert, even if present supply-side conditions are equal.

References

Ailawadi, Kusum, Yu Ma, and Dhruv Grewal, "The club store effect: Impact of shopping in warehouse club stores on consumers' packaged food purchases," *Journal of Marketing Research*, 2018, *55* (2), 193–207.

Algert, Susan J., Aditya Agrawal, and Douglas S. Lewis, "Disparities in Access to Fresh Produce in Low-Income Neighborhoods in Los Angeles," *American Journal of Preventive Medicine*, 2006, *30* (5), 365–370.

Allcott, Hunt, Rebecca Diamond, Jean-Pierre Dubé, Jessie Handbury, Ilya Rahkovsky, and Molly Schnell, "Food Deserts and the Causes of Nutritional Inequality," *Quarterly Journal of Economics*, 2019, *134* (4), 1793–1844.

Anupindi, Ravi, Maqbool Dada, and Sachin Gupta, "Estimation of Consumer Demand with Stock-Out Based Substitution: An Application to Vending Machine Products," *Marketing Science*, 1998, *17* (4), 406–423.

Atkin, David, "The caloric costs of culture: Evidence from Indian migrants," *American Economic Review*, 2016, *106* (4), 1144–1181.

Auld, M. Christopher and Paul Grootendorst, "An Empirical Analysis of Milk Addiction," *Journal of Health Economics*, 2004, *23* (6), 1117–1133.

Bai, Jushan, "Panel Data Models With Interactive Fixed Effects," *Econometrica*, 2009, *77* (4), 1229–1279.

Baltagi, Badi H. and James M. Griffin, "Rational Addiction to Alcohol: Panel Data Analysis of Liquor Consumption," *Health Economics*, 2002, *11* (6), 485–491.

Bass, Frank M., Moshe M. Givon, Manohar U. Kalwani, David Reibstein, and Gordon P. Wright, "An Investigation into the Order of the Brand Choice Process," *Marketing Science*, 1984, *3* (4), 267–287.

Becker, Gary and Kevin Murphy, "A Theory of Rational Addiction," *Journal of Political Economy*, 1988, *96* (4), 675–700.

Becker, Gary S., Michael Grossman, and Kevin M. Murphy, "An Empirical Analysis of Cigarette Addiction," *American Economic Review*, 1994, *84* (3), 396–418.

Biswas, Lisa, Erin Harrison, Yongzhen Gong, Ramachandram Avusula, Jonathan Lee, Meiyu Zhang, Thomas Rousselle, Janice Lage, and Xiu Liu, "Enhancing Effect of Menthol on Nicotine Self-Administration in Rats," *Psychopharmacology*, 2016, *233* (18), 3417–3427.

Bronnenberg, Bart J., Jean-Pierre H. Dube, and Joonhwi Joo, "Millennials and the Take-Off of Craft Brands: Preference Formation in the U.S. Beer Industry," *forthcoming at Marketing Science*, 2021.

_, Jean-Pierre H. Dubé, and Matthew Gentzkow, "The Evolution of Brand Preferences: Evidence from Consumer Migration," *American Economic Review*, 2012, *102* (6), 2472–2508.

_, Sanjay K. Dhar, and Jean Pierre H. Dube, "Brand History, Geography, and the Persistence of Brand Shares," *Journal of Political Economy*, 2009, *117* (1), 87–115.

Bruno, Hernán A and Naufel J Vilcassim, "Research Note - Structural Demand Estimation with Varying Product Availability," *Marketing Science*, 2008, *27* (6), 1126–1131.

Bryan, Gharad, Dean Karlan, and Scott Nelson, "Commitment Devices," *Annual Review of Economics*, 2010, *2*, 671–698.

Carlson, Andrea C, Elina Tselepidakis Page, Thea Palmer Zimmerman, Carina E Tornow, and Sigurd Hermansen, "Linking USDA Nutrition Databases to IRI Household-Based and Store- Based Scanner Data," Technical Report Technical Bulletin TB-1952, U.S. Department of Agriculture, Economic Research Service 2019.

Center for Nutrition Policy and Promotion (CNPP), *Dietary Guidelines for Americans 2010.*

Chaloupka, Frank J., "Rational Addictive Behavior and Cigarette Smoking," *Journal of Political Economy*, 1991, *99* (4), 722–742.

Chamberlain, Gary, "Heterogeneity, Omitted Variable Bias, and Duration Dependence," in James J. Heckman and Burton S. Singer, eds., *Longitudinal Analysis of Labor Market Data*, Cambridge University Press, 1985.

Conlon, Christopher T. and Julie Holland Mortimer, "Demand Estimation Under Incomplete Product Availability," *American Economic Journal: Microeconomics*, 2013, *5* (4), 1–30.

Courtemanche, Charles and Art Carden, "Supersizing supercenters? The impact of Walmart Supercenters on body mass index and obesity," *Journal of Urban Economics*, 2011, *69* (2), 165–181.

Cummins, Steven, Ellen Flint, and Stephen A. Matthews, "New neighborhood grocery store increased awareness of food access but did not alter dietary habits or obesity," *Health Affairs*, 2014, *33* (2), 283–291.

Delnevo, Cristine D., Daniel P. Giovenco, and Andrea C. Villanti, "Assessment of Menthol and Nonmenthol Cigarette Consumption in the US, 2000 to 2018," *JAMA Network Open*, 2020, *3* (8), 2020–2023.

_ , Ollie Ganz, and Renee D. Goodwin, "Banning Menthol Cigarettes: A Social Justice Issue Long Overdue," *Nicotine and Tobacco Research*, 2020, *22* (10), 1673–1675.

Dubé, Jean-Pierre, Günter J. Hitsch, and Peter E. Rossi, "Do Switching Costs Make Markets Less Competitive?," *Journal of Marketing Research*, 2009, *46* (4), 435–445.

_ , Günter J. Hitsch, and Peter E. Rossi, "State Dependence and Alternative Explanations for Consumer Inertia," *The RAND Journal of Economics*, 2010, *41* (3), 417–445.

Dubé, Jean-Pierre, Gunter J. Hitsch, and Peter E. Rossi, "State dependence and alternative explanations for consumer inertia," *The RAND Journal of Economics*, 2010, *41* (3), 417–445.

Dubé, Jean-Pierre, Günter J. Hitsch, Peter E. Rossi, and Maria Ana Vitorino, "Category Pricing with State-Dependent Utility," *Marketing Science*, 2008, *27* (3), 417–429.

FDA, "FDA Proposes Rules Prohibiting Menthol Cigarettes and Flavored Cigars to Prevent Youth Initiation, Significantly Reduce Tobacco-Related Disease and Death," 2022.

Figueroa, Cristian, Andres Musalem, and Carlos Noton, "Earthquakes and Brand Loyalty: Beyond the Short-Term Effects of Stockouts," 2019.

Foulds, Jonathan, Monica Webb Hooper, Mark J. Pletcher, and Kolawole S. Okuyemi, "Do Smokers of Menthol Cigarettes Find it Harder to Quit Smoking?," *Nicotine and Tobacco Research*, 2010, *12* (SUPPL. 2), 102–109.

Frank, Ronald E., "Brand Choice as a Probability Process," *The Journal of Business*, 1962, *35*, 10–11.

Goodman, J. David, "When Big Tobacco Invoked Eric Garner to Fight a Menthol Cigarette Ban," *New York Times*, 2019.

Gordon, Brett R. and Baohong Sun, "A Dynamic Model of Rational Addiction: Evaluating Cigarette Taxes," *Marketing Science*, 2015, *34* (3), 452–470.

Heckman, J. J., "Heterogeneity and State Dependence," in ed. Sherwin Rose, ed., *Studies in Labor Markets*, University of Chicago Press, 1981.

Hoffman, Steven J, Jessica Mammone, Susan Rogers Van Katwyk, Lathika Sritharan, Maxwell Tran, Safa Al-Khateeb, Andrej Grjibovski, Elliot Gunn, Sara Kamali-Anaraki, Ben Li, Mathura Mahendren, Yasmeen Mansoor, Navneet Natt, Ejike Nwokoro, Harkanwal Randhawa, Melodie Yunju Song, Kelsey Vercammen, Carolyne Wang, Julia Woo, and Mathieu JP Poirier, "Cigarette Consumption Estimates for 71 Countries from 1970 to 2015: Systematic Collection of Comparable Data to Facilitate Quasi-experimental Evaluations of National and Global Tobacco Control Interventions," *BMJ*, 2019, *365.*

Hristakeva, Sylvia and Julia Levine, "The Effects of Grocery Store Closures on Demand for Healthy Food," *Working Paper*, 2022.

Hut, Stefan, "Determinants of Dietary Choice in the US: Evidence from Consumer Migration," *Journal of Health Economics*, 2020.

_ **and Emily Oster**, "Changes in Household Diet: Determinants and Predictability," *Working*, 2019, pp. 1–55.

Iannaccone, Laurence R., "Addiction and Satiation," *Economics Letters*, 1986, *21* (1), 95–99.

Jones, J. Morgan and Jane T. Landwehr, "Removing Heterogeneity Bias from Logit Model Estimation," *Marketing Science*, 1988, *7* (1), 41–59.

Keane, Michael, "Modeling Heterogeneity and State Dependence in Consumer Choice Behavior," *Journal of Businessand Economic Statistics*, 1997, *15* (3), 310–327.

Kim, Jong Yeob and Masakazu Ishihara, "Evaluating the Effect of Soda Taxes Using a Dynamic Model of Rational Addiction," 2021.

Krebs-Smith, Susan M., Tusa Rebecca E. Pannucci, Amy F. Subar, Sharon I. Kirkpatrick, Jennifer L. Lerman, Janet A. Tooze, Magdalena M. Wilson, and Jill Reedy, "Update of the Healthy Eating Index: HEI-2015," *Journal of the Academy of Nutrition and Dietetics*, 2018, *118* (9), 1591–1602.

Kreslake, Jennifer M., Geoffrey Ferris Wayne, Hillel R. Alpert, Howard K. Koh, and Gregory N. Connolly, "Tobacco Industry Control of Menthol in Cigarettes and Targeting of Adolescents and Young Adults," *American Journal of Public Health*, 2008, *98* (9), 1685–1692.

Larcom, Shaun, Ferdinand Rauch, and Tim Willems, "The Benefits of Forced Experimentation: Striking Evidence from the London Underground Network*," *The Quarterly Journal of Economics*, 05 2017, *132* (4), 2019–2055.

Levine, Julia and Stephan Seiler, "Identifying State Dependence in Brand Choice: Evidence from Hurricanes," *Working Paper*, 2022.

Levy, David T., Kenneth Blackman, John Tauras, Frank J. Chaloupka, Andrea C. Villanti, Raymond S. Niaura, Donna M. Vallone, and David B. Abrams, "Quit Attempts and Quit Rates Among Menthol and Nonmenthol Smokers in the United States," *American Journal of Public Health*, 2011, *101* (7), 1241–1247.

Luke, Douglas, Emily Esmundo, and Yael Bloom, "Smoke Signs: Patterns of Tobacco Billboard Advertising in a Metropolitan Region," *Tobacco Control*, 2000, *9* (1), 16–23.

Massy, William F., "Order and Homogeneity of Family Specific Brand-switching Processes," *Journal of Marketing Research*, 1966, *3* (1), 48–54.

Musalem, Andrés, Marcelo Olivares, Eric T. Bradlow, Christian Terwiesch, and Daniel Corsten, "Structural Estimation of the Effect of Out-of-Stocks," *Management Science*, 2010, *56* (7), 1180–1197.

Olekalns, Nilss and Peter Bardsley, "Rational Addiction to Caffeine: An Analysis of Coffee," *Journal of Political Economy*, 1996, *104* (5), 1100–1104.

Paulson, Kelly, "Identification of Structural State Dependence in Brand Choice Panel Data," *working paper*, 2012.

Raval, Devesh, Ted Rosenbaum, and Nathan Wilson, "Using Disaster Induced Closures to Evaluate Discrete Choice Models of Hospital Demand," *The RAND Journal of Economics*, 2020, *R&R*, 1–92.

Seetharaman, P.B., Andrew Ainslie, and Pradeep Chintagunta, "Investigating Household State Dependence Effects Across Categories," *Journal of Marketing Research*, 1999, *36*, 488–500.

Sharkey, Joseph R., Scott Horel, and Wesley R. Dean, "Neighborhood Deprivation, Vehicle Ownership, and Potential Spatial Access to a Variety of Fruits and Vegetables in a Large Rural Area in Texas," *International Journal of Health Geographics*, 2010, pp. 1–27.

Simonov, Andrey, Jean-Pierre Dubé, Günter Hitsch, and Peter Rossi, "State-Dependent Demand Estimation with Initial Conditions Correction," *Journal of Marketing Research*, 2020, *57* (5), 789–809.

Smith, Philip H., Biruktawit Assefa, Simranpreet Kainth, Kaliris Y. Salas-Ramirez, Sherry A. McKee, and Gary A. Giovino, "Use of Mentholated Cigarettes and Likelihood of Smoking Cessation in the United States: A Meta-analysis," *Nicotine and Tobacco Research*, 2020, *22* (3), 307–316.

Smith, Stevens S., Michael C. Fiore, and Timothy B. Baker, "Smoking Cessation in Smokers Who Smoke Menthol and Non-menthol Cigarettes," *Addiction*, 2014, *109* (12), 2107–2117.

Stigler, George J. and Gary S. Becker, "De Gustibus Non Est Disputandum," *American Economic Review*, 1977, *67* (2), 76–90.

Sudhir, K and Ishani Tewari, "Long Term Effects of 'Prosperity in Youth' on Consumption: Evidence from China," *Working Paper*, 2015, *2466.*

Sudhir, K. and Nathan Yang, "Exploiting the Choice-Consumption Mismatch: A New Approach to Disentangle State Dependence and Heterogeneity," *Working Paper*, 2014.

Tuchman, Anna E., "Advertising and Demand for Addictive Goods: The Effects of E-cigarette Advertising," *Marketing Science*, 2019, *38* (6), 994–1022.

Villanti, Andrea C., Paul D. Mowery, Cristine D. Delnevo, Raymond S. Niaura, David B. Abrams, and Gary A. Giovino, "Changes in the Prevalence and Correlates of Menthol Cigarette Use in the USA, 2004–2014," *Tobacco Control*, 2016, *25*, ii14–ii20.

Volpe, Richard, Xiaowei Cai, Christiane Schroeter, and Lisa Mancino, "The Effect of Produce Purchasing Behavior on Adult Obesity and Body Weight: Evidence from Household-Level Scanner Data," *Agricultural and Resource Economics Review*, 2019, *48* (2), 253–273.

Vulcano, Gustavo, Garrett van Ryzin, and Richard Ratliff, "Estimating Primary Demand for Substitutable Products from Sales Transaction Data," *Operations Research*, 2012, *60* (2), 313–334.

Wailoo, Keith, *Pushing Cool: Big Tobacco, Racial Marketing, and the Untold Story of the Menthol Cigarette*, University of Chicago Press, 2021.

Wickham, Robert J., "The Biological Impact of Menthol on Tobacco Dependence," *Nicotine and Tobacco Research*, 2021, *22* (10), 1676–1684.

Xu, Yiqing, "Generalized Synthetic Control Method: Causal Inference with Interactive Fixed Effects Models," *Political Analysis*, 2017, *25* (1), 57–76.