DISCUSSION PAPER SERIES

DP16632
(v. 3)

IDENTIFYING STATE DEPENDENCE IN BRAND CHOICE: EVIDENCE FROM HURRICANES

Stephan Seiler and Julia Levine

INDUSTRIAL ORGANIZATION

ISSN 0265-8003

IDENTIFYING STATE DEPENDENCE IN BRAND CHOICE: EVIDENCE FROM HURRICANES

Stephan Seiler and Julia Levine

Discussion Paper DP16632
First Published 14 June 2022
This Revision 23 October 2022

Centre for Economic Policy Research
33 Great Sutton Street, London EC1V 0DX, UK
Tel: +44 (0)20 7183 8801
www.cepr.org

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.

JEL Classification: D12, L81

Keywords: Brand choice, Brand loyalty, State dependence, Preference heterogeneity

Stephan Seiler - stephan.a.seiler@gmail.com
Imperial College London and CEPR

Julia Levine - levinejulia@ucla.edu
Ucla

Acknowledgements
We thank Anand Bodapati, Randolph Bucklin, Brett Hollenbeck, Elisabeth Honka, Sylvia Hristakeva, Peter Rossi, and Adam Smith for helpful comments. 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. 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. Please contact Levine (levinejulia@ucla.edu) or Seiler (stephan.a.seiler@gmail.com) for correspondence. All authors certify that they have no affiliations with or involvement in any organization or entity with any financial interest or non-financial interest in the subject matter or materials discussed in this manuscript. The authors have no funding to report.

Identifying State Dependence in Brand Choice: Evidence from Hurricanes*

Julia Levine
Anderson School of Management
UCLA

Stephan Seiler
Imperial College
& CEPR

This draft: September 1, 2022

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.

Keywords: Brand Choice, Brand Loyalty, State Dependence, Preference Heterogeneity

*We thank Anand Bodapati, Randolph Bucklin, Brett Hollenbeck, Elisabeth Honka, Sylvia Hristakeva, Peter Rossi, and Adam Smith for helpful comments. 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. 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. Please contact Levine (levinejulia@ucla.edu) or Seiler (stephan.a.seiler@gmail.com) for correspondence. Neither of the authors received external funding for this paper.

1 Introduction

A large literature in marketing and economics (e.g., Jones and Landwehr (1988), Keane (1997), Seetharaman et al. (1999), Dubé et al. (2010)) 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 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.

[1] Similar to Dubé et al. (2009), Dubé et al. (2010), 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.

[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.

Our empirical findings differ from most prior papers (e.g., Keane (1997), Seetharaman et al. (1999), Dubé et al. (2010), 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 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. (2010) 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 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

[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 that are also observed in the Nielsen store-level data. We provide more details on sample construction in Section 5 and Appendix E.

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. (2021) 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 2 we present the data and descriptive statistics. In Section 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 4 we present our main empirical analysis and robustness checks. In Section 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 6.

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.

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 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 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. 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

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

[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

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

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 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.

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.

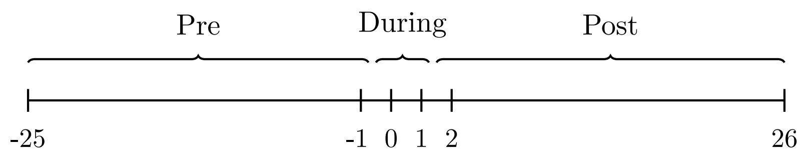

Figure 1: **Timeline of the Estimation Sample.** The graph shows a running counter of weeks. Week 0 is defined as the week that ends in the hurricane. The "during" period comprises week 0 and week 1. The pre and post periods comprise 25 weeks each.

2.2 Choice Persistence

We define choice persistence within a category as

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

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.

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

[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 2: **Descriptive Statistics.**

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 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 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 2 also describes choice persistence for each product category, calculated as shown in equation (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 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 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]

[13]These criteria are used for all of our main empirical analysis in Section 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.

[16]We also observe a slightly higher-than-usual expenditure pattern in week 1 after the hurricane in the treatment

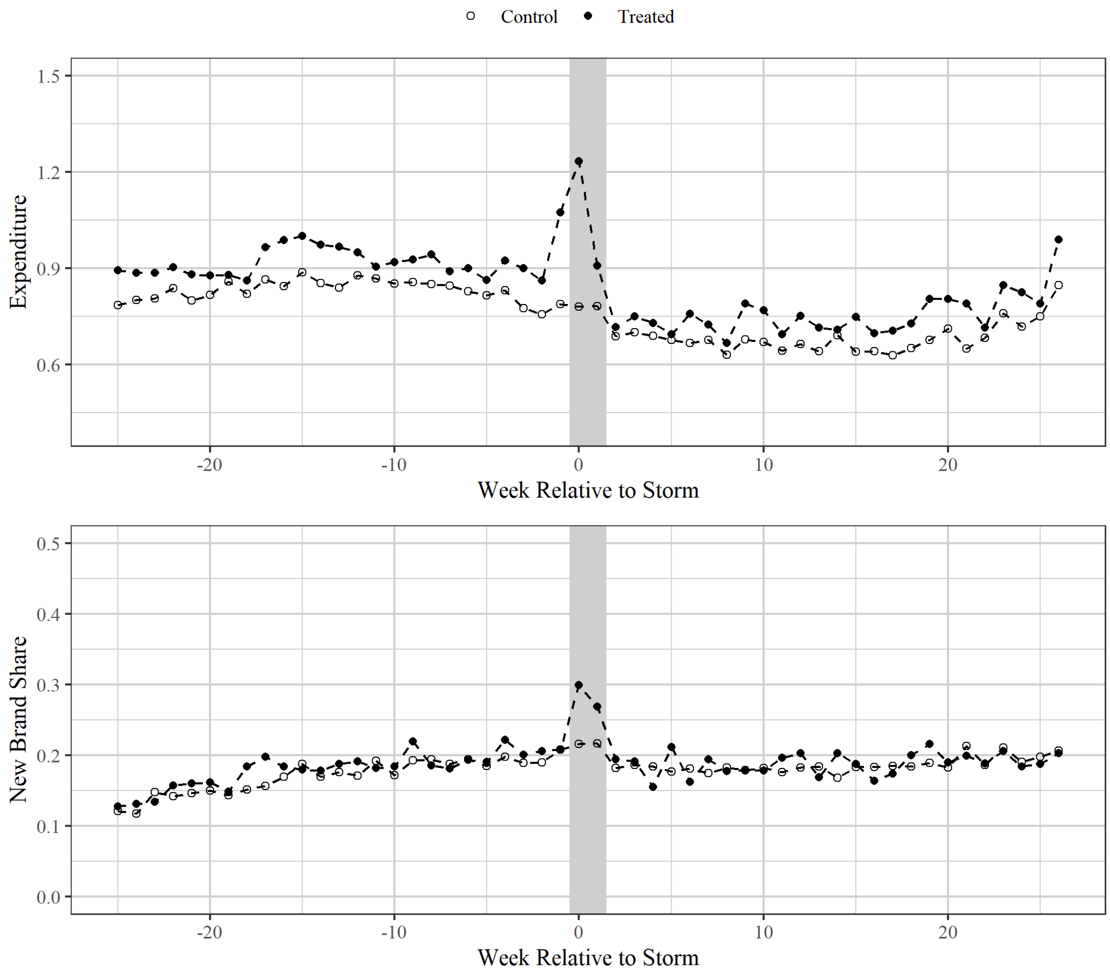

Figure 2: **Bottled Water Expenditure and Purchases of New Brands.** The top graph displays average weekly expenditure (in dollars) per household in the bottled water category for the treatment and control group. New brand share is the share of unique brands purchased on a given shopping trip that were 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.

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.

In the bottom graph of Figure 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

group, possibly due to imperfect data on the exact timing of the hurricane.

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 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.

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 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

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

[17] In Section 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.

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:

$$\widetilde{Persist}_{it} = \begin{cases} \frac{1}{\#J_{it}} \sum_{j \in J_{it}} 1(j \in J_{it}^{pre-hurricane}) & \text{if "first trip after the hurricane"} \\ \frac{1}{\#J_{it}} \sum_{j \in J_{it}} 1(j \in J_{it}^{last}) & \text{otherwise.} \end{cases} \tag{2}$$

This measure of choice persistence is identical to the standard definition of choice persistence in equation (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 3 and leads to an average choice persistence of around 0.65 in the pre-hurricane

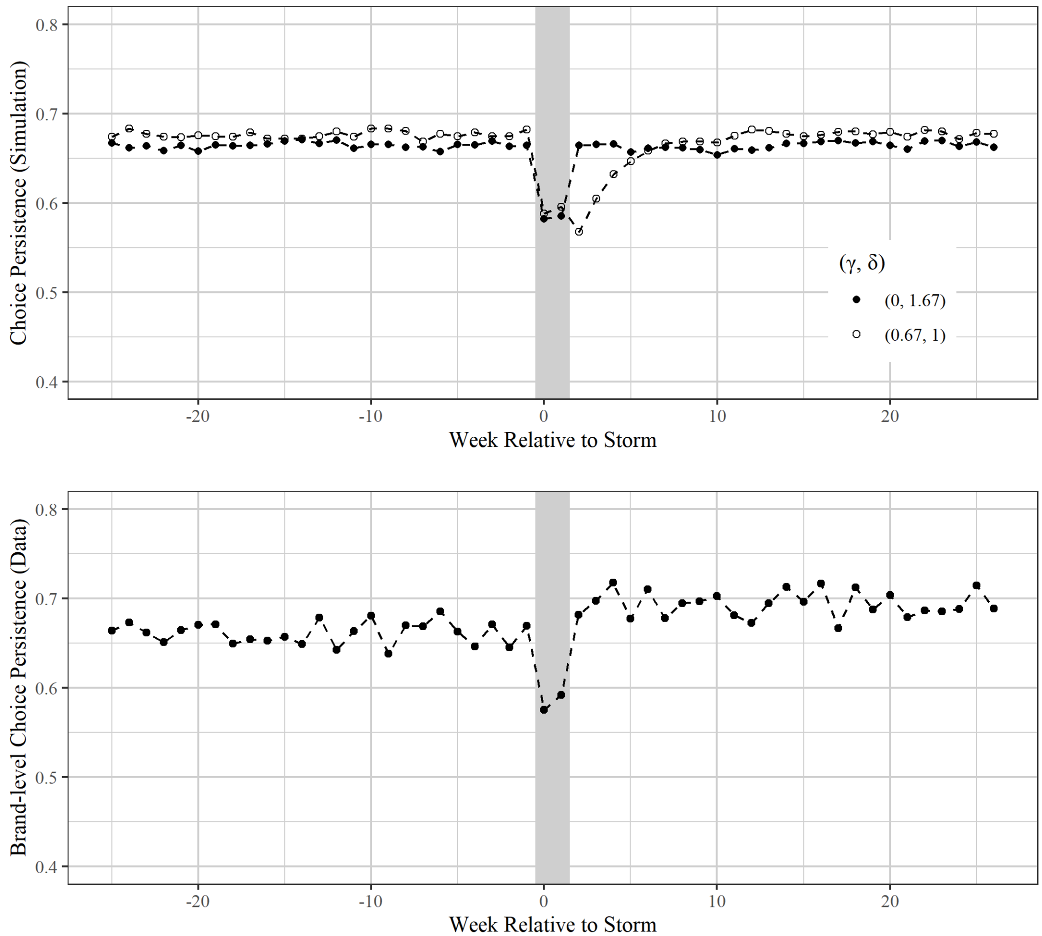

Figure 3: **Average Choice Persistence: Simulated Data and Empirical Patterns.** The top graph shows how average choice persistence evolves in response to a stock-out shock based on simulations with and without structural state dependence. The bottom graph plots average choice persistence in the data before and after a hurricane. In both graphs the vertical gray bars indicate weeks 0 and 1 which are affected by stock-outs.

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 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 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 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 A 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 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 3.

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. (2010), Simonov et al. (2020)). We return to a more detailed comparison to this alternative approach in Section 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.

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 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 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.

[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 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.

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] 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 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 4.2, and an analysis of which products consumers switch to and whether they alter other aspects of their behavior in Section 4.3. Finally, we analyze possible other ways in which hurricanes can impact consumers apart from stock-outs in Section 4.4.

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:

$$\widetilde{Persist}_{it} = \delta_{it} D_{it} + \alpha_t + \boldsymbol{\lambda}_i' \boldsymbol{f}_t + \epsilon_{it} \tag{3}$$

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 (3) assumes that both treated and control units are affected by the same set and number of unobserved

[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, 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 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 B we analyze time trends in the treatment and control group in detail.

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]

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:

$$\widehat{Persist}_{it}(0) = \hat{\alpha}_t + \hat{\boldsymbol{\lambda}}_i{}' \hat{\boldsymbol{f}}_t \tag{4}$$

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 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 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 4 therefore suggest that consumers' choices do not exhibit structural state dependence.

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 5 and Appendix A) suggest that persistence

[21] Note that equation (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.

[23] We re-iterate that our analysis uses the modified persistence measure defined in equation 2, which defines persistence on the first purchase after the hurricane in relation to the last purchase before the hurricane.

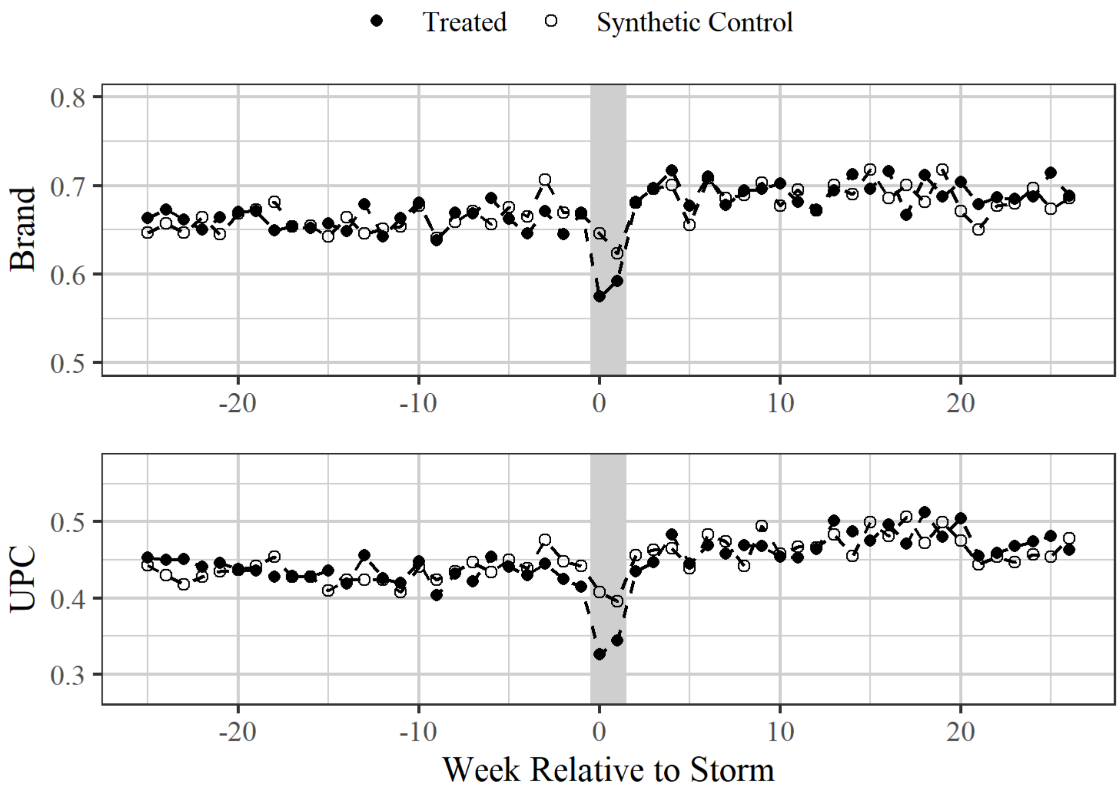

Figure 4: **Average Choice Persistence: Treatment Group and Synthetic Control.** The graphs display average choice persistence at the brand- and UPC-level. Closed and open dots represent choice persistence in the treatment group and the synthetic control value respectively. The vertical gray bars indicate weeks 0 and 1 which are affected by stock-outs.

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 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.

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

	(1)	(2)	(3)	(4)	(5)
			Robustness Checks		
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 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.

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 4 which plots choice persistence in the treatment group and the synthetic control.

In the final two columns of Table 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:

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

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 B 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 3.[24]

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

[24] As we show in Appendix B, 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 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 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.

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 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 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 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 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 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 A1 in Appendix C 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 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.

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 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 5 and find that expenditure increased significantly during the hurricane. We then decompose the expenditure

[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 5.

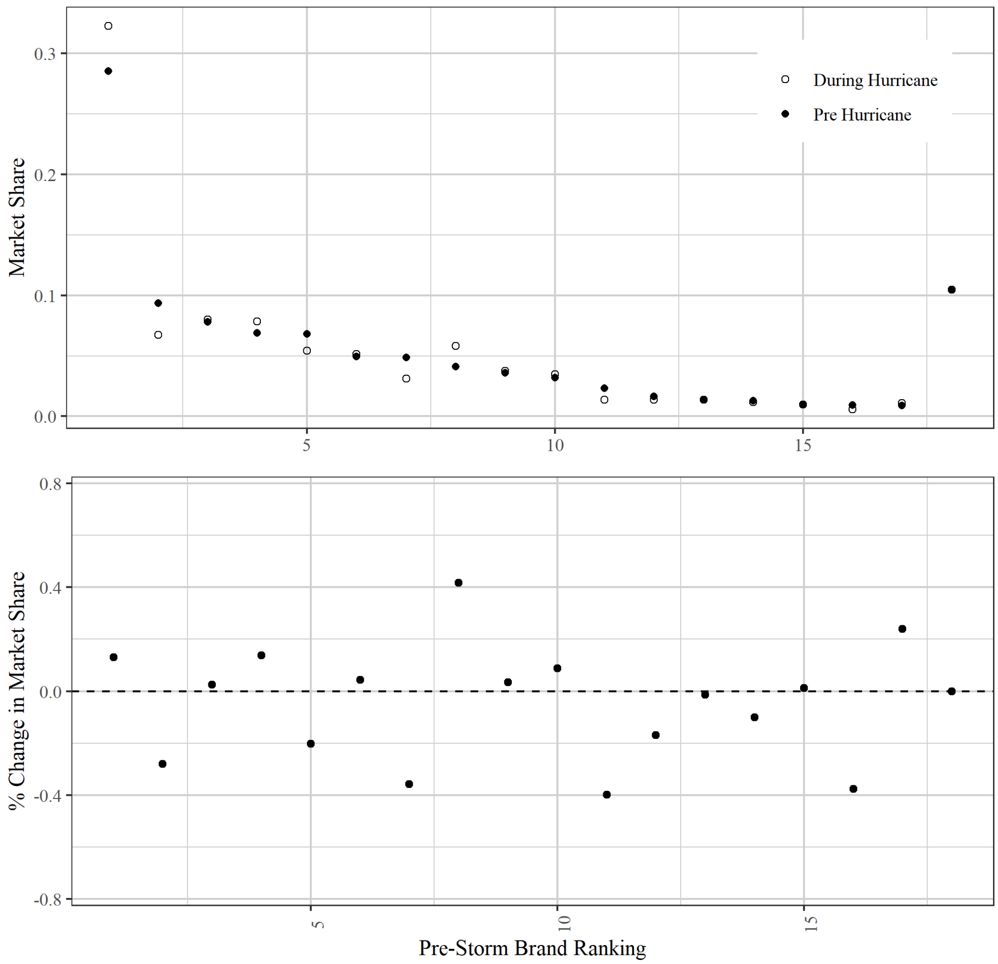

Figure 5: **Market Share of Top Brands Before and During the Hurricane.** The top graphs displays market-shares for the top 17 brands (ranked from largest the smallest) and a residual category of all other brands (the right-most points) before the hurricane (solid dots) and during the hurricane (open 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.

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 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 5, a point that we will return to in 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

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 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.

control approach introduced in Section 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 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 6 we 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 D 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

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

	(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 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.

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 un-

[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.

usual purchase behavior during the hurricane such as purchases of niche products, bulk buying, or purchases of multiple brands on the same shopping trip.

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.

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 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 A2 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 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.

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. (2010), 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 E. We also note that the set of households used to analyze behavior in the bottled water category in our main analysis is different

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

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 3 as the basis for illustrating our identification strategy. Specifically, we assume that the 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 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 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 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 A. In Table 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 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

[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.

[30] In Appendix E, 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 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.

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.

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. (2010)). 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

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 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.

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.

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. (2010), 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.

5.3 Other Differences

We re-iterate that a series of other differences that we discussed in Sections 4.3 and 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 4.3) or disruptive effects of hurricanes that directly impact consumers' purchase behavior (see Section 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 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.

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.

References

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.

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

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.

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

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.

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.

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

_ , _ , _ , and Maria Ana Vitorino, "Category Pricing with State-Dependent Utility," *Marketing Science*, 2008, *27* (3), 417–429.

Figueroa, Cristian, Andrés Musalem, and Carlos Noton, "Earthquakes and Brand Loyalty: Beyond the Short-Term Effects of Stockouts," *Working Paper*, 2021.

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

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

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.

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.

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

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

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.

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

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.

Xu, Yiqing, “Generalized synthetic control method: Causal inference with interactive fixed effects models,” *Political Analysis*, 2017, *25* (1), 57–76.

A 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 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 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]

Figure A1 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 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 8.

[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.

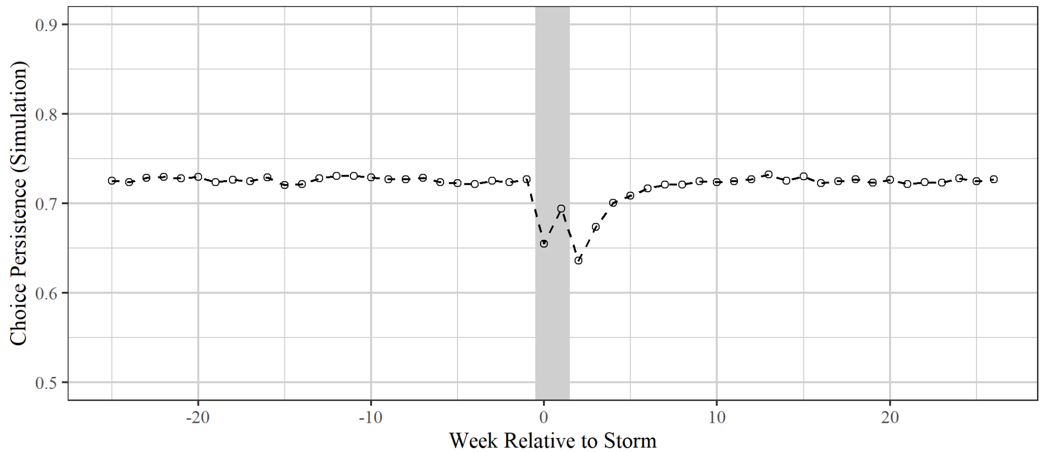

Figure A1: **Average Choice Persistence: Simulated Data with Preferences from Structural Model Estimates.** The vertical gray bar indicates weeks 0 and 1 which are affected by stock-outs.

B 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:

$$\begin{aligned} \widetilde{Persist}_{it} &= \sum_{t=-24}^{t=26} \alpha_t \times \mathbf{1}(Week = t) \times Treat_i \\ &+ \sum_{t=-24}^{t=26} \beta_t \times \mathbf{1}(Week = t) \times (1 - Treat_i) \\ &+ \gamma_i + \varepsilon_{it} \end{aligned} \tag{5}$$

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 A2. 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 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

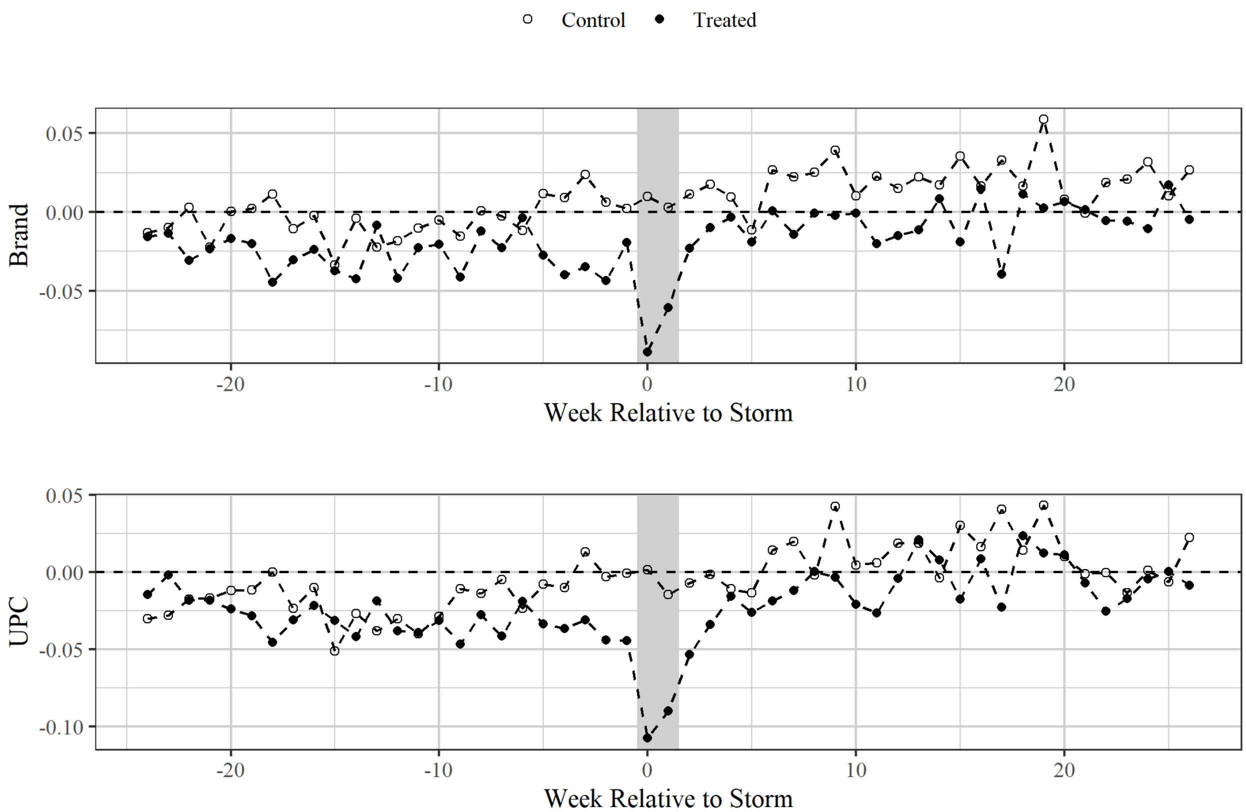

Figure A2: **Choice Persistence Over Time.** The graphs plot estimated week dummies for the treatment and control group from a regression that also includes consumer fixed effects. Weekly effects are estimated for weeks -24 to 26. The first week of the sample (week -25) constitutes the omitted category. The vertical gray bars indicate weeks 0 and 1 which are affected by stock-outs.

time period fixed effects.

C 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 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 A1 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.

		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 A1: **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.

D Additional Robustness Checks

In this section we discuss the results from a series of additional regressions reported in Table A2. 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 3) as a benchmark in column (1) of Table A2. The remaining columns present results from regressions that either change the outcome variable or focus on specific subsets of households.

D.1 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) 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 6 in the main paper, we conclude that our null finding is robust to a variety of ways of tackling multi-brand purchases.

D.2 Disruptions due to Hurricanes

In Section 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.

D.3 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 4 in Section 4.2 deals with this issues most directly, because

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

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 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 4.2, 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 A2 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 4 we conclude that compositional changes due to different purchase frequencies across households are not driving our null result.

E 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 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.

E.1 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 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

	(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 A2: **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.

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

		(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 A3: **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.

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.

E.2 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 A3. 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 7 in the main paper correspond to column (4) in Table A3.

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 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.