Title
Essays on the Industrial Organization of Food Retailing

Author
Shih, Tiffany

Publication Date
2013

Peer reviewed|Thesis/dissertation

Essays on the Industrial Organization of Food Retailing

by

Tiffany Ming Shih

A dissertation submitted in partial satisfaction of the
requirements for the degree of
Doctor of Philosophy

in

Agricultural and Resource Economics

in the

Graduate Division
of the
University of California, Berkeley

Committee in charge:
Professor Jeff Perloff, Co-chair
Professor Sofia Villas-Boas, Co-chair
Professor J. Miguel Villas-Boas

Fall 2013

Essays on the Industrial Organization of Food Retailing

Abstract

Essays on the Industrial Organization of Food Retailing

by

Tiffany Ming Shih
Doctor of Philosophy in Agricultural and Resource Economics

University of California, Berkeley

Professor Jeff Perloff, Co-chair
Professor Sofia Villas-Boas, Co-chair

This thesis consists of two empirical studies of the industrial organization of food retailing.

Chapter 1 investigates consumer behavior in response to changes in variety in grocery store offerings. Using a rich panel dataset of consumer purchases from one national retailer, we measure the effects of changes in product variety on sales of ice cream, where the assortment changes were gradually introduced to stores. We use a difference in differences approach to measure treatment effects against trends in stores prior to treatment. Our results show an adjustment to product variety causes consumers to alter their purchasing behavior, at least in the subsequent eight weeks of the change. When the rearrangement consisted of adding additional products, consumers purchased more on average in the immediate term, but sales increases mitigated over time, with smaller magnitude increases in variety experiencing sales reductions in later weeks. Product groups treated with reduced variety experienced decreased sales both in the immediate and longer term. We find that extending and reducing assortment have asymmetric effects, with a marginal change to increased variety boosting sales less than a marginal change in decreased variety. Larger magnitude reductions perform poorly in the short run, but actually perform better than both smaller variety reductions and small variety increases in later weeks.

In Chapter 2, I investigate the following question: do firms respond to quality disclosure by nearby competitors? This paper utilizes exogenous variation provided by Los Angeles County's introduction of restaurant hygiene grade cards to explore inter-firm responses in markets for differentiated goods. Under this program, a subset of cities adopted mandatory grade card posting, requiring restaurants to disclose hygiene. This study demonstrates that restaurants under a voluntary hygiene disclosure policy improved hygiene more when located closer to restaurants in mandatory disclosure areas. The proximity effect implies that mandatory disclosure policies influenced restaurant hygiene in nearby voluntary disclosure areas, thus providing direct evidence that quality-disclosure regulations affect non-targeted firms.

I dedicate this dissertation to my beloved son Robert, who fills my life with meaning. To my husband Barney, for his faithful companionship and love. Finally, to my mother Kai, for her humbling generosity.

Contents

Acknowledgments

For chapter 1, we thank the retailer for sharing his data, helping us understand the data, and allowing us to conduct the analysis. Lydia Ashton and Barney Hartman-Glaser provided particularly useful input. We are also grateful for the help of our research assistants and the Giannini Foundation for support.

For chapter 2, I thank Michael Anderson, Lydia Ashton, Maximillian Auffhaumer, Andrew Dustan, Barney Hartman-Glaser, Guo-Jun He, Robert Helsley, Jeremy Magruder, Elisabeth Sadoulet, Kiran Torani for helpful comments. I am particularly grateful to Jeffrey Perloff and Sofia Villas-Boas for valuable advice and guidance. I also thank Anhdao Truong and the Los Angeles County Department of Public Health, Ginger Jin, and Philip Leslie, for allowing me to access data. I finally thank the excellent work provided by the research assistants. All errors are my own.

Chapter 1

Measuring Consumer Responses to Changes in Variety

1.1 Introduction

How does a change in product variety affect sales of a product category? In recent years, the number of products offered by grocery retailers has grown tremendously. In 2010 the average supermarket occupied 46,000 square feet (a 31% increase from 1994) and offered nearly 39,000 products).[1] Retail shelf space is active real estate in retail markets, as manufacturers compete for shelf schematics that will catch consumer attention. As with other shelving variables, decisions on products offerings in a given store result from negotiations between the manufacturer and the retailer.

Because larger assortments may provide manufacturers with more shelf space and better meet the changing desires of consumers over time, manufacturers may extend a brands line in an attempt to stimulate sales in a competitive environment (Draganska and Jain, 2005; Kahn, 1998; Scheibehenne, 2008). Similarly, retailers may offer large varieties to retain customer loyalty. Indeed, traditional economic theory suggests that larger menus should weakly increase consumption, since offering more choices will increase the probability of fit both between consumers and for the same consumer over time (Lancaster, 1990). Additionally, consumers may prefer larger menus due to a preference for flexibility, or uncertainty regarding future tastes (Kreps 1997).

A reexamination of the effects of product variety has been stimulated by recent studies in psychology demonstrating that sufficiently large assortments may be associated with reduced choice (Shafir, Simonson, and Tversky, 1993; Chernev, 2003a,b; Reutskaja and Hogarth, 2009). Termed "choice overload," this phenomenon has been demonstrated in a number of laboratory experiments and surveys, for a variety of consumer products (see Scheibehenne (2008) for a review). In a well-known experiment, Iyengar and Lepper (2000) found that a larger percentage of shoppers approached a sample table with a large variety of jam samples,

[1] http://www.fmi.org/research-resources/supermarket-facts/median-total-store-size-square-feet

but a smaller percent of these shoppers subsequently purchased jam when compared to those who sampled from a table with less variety.

As reviewed in Scheibehenne, Greifeneder, and Todd (2010), similar findings emerged in experiments using a wide range of consumer goods, including choices of chocolates, coffee, pens, and electronics (Chernev, 2003a; Dhar, 1997; Iyengar and Lepper, 2000; Mogilner, Rudnick, and Iyengar, 2008; Shah and Wolford, 2007). However, in a metanalysis covering experiments on increased choice, Scheibehenne, Greifeneder, and Todd (2010) found the effect on consumption to be highly variable, with a mean of nearly zero. Despite these conflicting results, several determinants seem to precipitate choice overload. When a large variety of choices has no clearly dominant option or when consumers are less familiar with the options provided, individuals are more likely to defer choice or feel less satisfied about any particular option (Chernev, 2003a; Dhar, 1997; Mogilner, Rudnick, and Iyengar, 2008; Scheibehenne, Greifeneder, and Todd, 2010; Shafir, Simonson, and Tversky, 1993). Because large assortments thus obscure the satisfaction from any particular option, individuals may avoid choice in order to avoid regretful decisions (Sarver, 2008; Zeelenberg, Van Dijk, Manstead, and vanr de Pligt, 2000). In addition, search costs may discourage choice when faced with too many options (Kuksov and Villas-Boas, 2010).

Several recent studies have investigated the effect of product removals at the retail level. Boatwright and Nunes (2001) analyzed purchases after an online grocer reduced product offerings to a subset of 292 consumers. They found that for the top-selling product categories, a one-time 56% reduction in products led to an 11% overall increase in sales. Using the same online grocer, Borle, Boatwright, Kadane, Nunes, and Galit (2005) found that product variety reductions, ranging from 21% to 91% in different categories, reduced both shopping frequency and purchase quantity, thereby reducing overall sales, with the strongest effects in less-frequently purchased categories. The authors note their results contrast from Boatwright and Nunes (2001) because the latter focuses on top-selling categories. They venture that infrequent categories may be entered with very specific needs, or inexperience in these categories may lead shoppers to require more choice. Finally, Sloot, Fok, and Verhoef (2006) analyze a 25% item reduction in laundry detergent using a natural experiment in two treatment and two control stores of a major Dutch retailer. Consistent with the results of this paper, they find that the reduction led to large short-term sales losses in the product category but weaker long-term losses.

Our work advances this line of inquiry by utilizing a natural experiment to examine the effect of both increases and decreases in product assortment in a large-scale retail setting. We take advantage of a major change in shelf space allocation that was gradually implemented to measure how changes in product variety affected sales, controlling for changes to product shelf schematics. Strategic alterations to product offerings are endogenously decided by parties involved in the distribution of products, from the manufacturer to the final consumer. However, we use a rich panel dataset of consumer ice cream purchases from one national retailer, covering a period of phased-in shelf schematics changes in stores. By analyzing comparable treatment stores and control stores (stores prior to treatment), we implement a difference in differences strategy to measure the causal effects of changes in product variety. In our empirical research design, we benefit from the fact that stores share most marketing variables, such as price and promotional activities, during the pre-treatment and post-treatment periods. We separate the effects of assortment increases versus assortment decreases, analyze the role of the magnitude of a variety change, and further assess how consumers respond over time.

We find that in the first two to four weeks after the variety change, product groups with increased options experienced an average increase in sales by 4.7% to 5.8% for all ice cream products, with a more moderate effect of up to 2.7% on ice cream staples. However, after this initial period, product groups with increased variety experienced decreased sales, resulting in an average null effect for all ice cream and a negative effect for staples over the eight weeks after the change in assortment. We find that product groups treated with a decrease in variety had reduced sales by 12.2%-15.4% in the first two to four weeks, with the extent of sales reductions increasing over time. Sales of staples also experienced average decreases, but consistent with the pattern between Boatwright and Nunes (2001) and Borle, Boatwright, Kadane, Nunes, and Galit (2005), staples tended to perform better, with sales decreases at only 9.0%-13.4% in the first two to four weeks.

In addition, the magnitude of a variety change had an important effect on sales. In the two months after a change in variety, larger variety increases boosted sales more than smaller variety increases, and larger variety decreases reduced sales more than small decreases. We find an asymmetric effect between adding versus removing options, with a marginal decrease impairing sales more than a marginal increase improves sales. Exploring magnitude effects over time reveals that while larger increases in variety sustained sales growth for at least six weeks after the treatment, smaller variety increases actually led to reduced sales one to two months later. Interestingly, in later weeks, large magnitude decreases hindered sales less than small magnitude decreases, when they perform even better than small magnitude increases in variety, in terms of the effect on percent changes in sales.

The rest of the paper proceeds as follows. Section (1.2) describes the empirical setting and summarizes the data, while section (1.3) outlines the research design and empirical strategy. Section (1.4) presents the results and section (1.5) concludes.

1.2 Empirical Setting and Data

From a major grocery retailer, we obtained a panel dataset of U.S. store sales and shelf schematics at the product version (UPC) level for the entire ice cream category. The dataset covers the period from August 13, 2008 to December 30, 2009, with 1358 retail locations offering a total of 884 unique ice cream products (UPCs).[2] Specifically, for each ice cream UPC at each store, our data include the time series of when the UPC was introduced or removed, how much shelf space was allocated to the UPC, how many units were purchased, and the purchase revenues. The date system in the raw shelf schematics data do not perfectly

[2] We only include UPCs which are confirmed ice cream products, either through the retailer or through online searchers of the UPC code

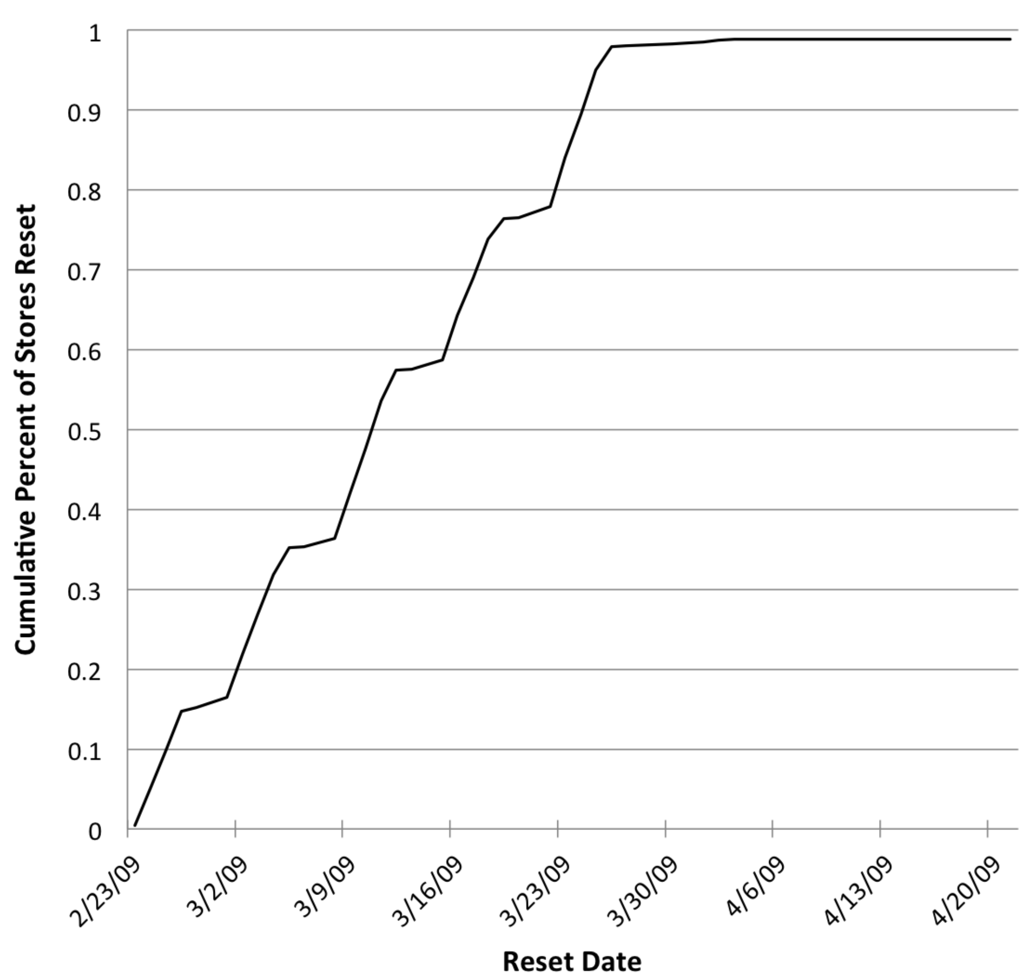
capture the times when products were on the shelf at stores. We worked with the retailer to best determine actual shelf appearances of products (for example, identifying "voided" products which are taken off the shelf due to low sales volumes).

Stores typically undergo a major reorganization of products on the shelf, known as a merchandising reset, once a year. Resets occur during store closing and generally take one night, but no more than two nights. In 2009, resets were completed between February 23 and April 21. Figure (1.1) displays the cumulative percent of stores reset over the February to April 2009 period. Stores implement major introductions of new products during resets. Because they require participation from third parties and may involve major physical modifications to shelves, resets are coordinated at the division level and scheduled to minimize interference with other store activities (such as a store inventory). Importantly, resets are not scheduled based on store sales or type, and thus the timing of the reset in a given store is exogenous to these factors. We utilize a difference in s approach with a staggered entry into the shelving reset treatment. In our analysis, we identify the effects of changes in variety by comparing sales trends for stores under treatment to control trends comprised of stores prior to treatment.

In our analyses, we aggregate ice cream products in order to overcome a large percentage of observations for UPCs that are not purchased in particular weeks. The retailer categorizes ice cream products into 62 subsubclasses (examples of which include "regular traditional premium packaged ice cream," "light/low fat healthy alternative premium packaged ice cream," and "ice cream cups kids novelties"), which is the retailer's finest level of ice cream classification. Thus, for our analysis, we categorize each UPC into one of 195 aggregated groups (AGs), where AGs are defined as a unique combination of brand, product size, and subsubclass. When sizes were not available from the retailer, we performed online searches of the UPC code to determine the size. In our analysis, a single observation therefore includes total sales in a particular AG, in a given store in a given two week period. The only meaningful variation AGs do not capture are flavor varieties within a specific class of ice cream. All the relevant variables are summed to form the aggregation, with shelving schematics then modified as the percent relative to the total schematics for ice cream in the store, during the given two-week period.

After aggregating in the manner described above, approximately 2.4% of our shelf observations have no purchase in the given store, in the given two-week period. Simply dropping these observations may introduce bias in our results. Therefore, in two separate samples, we systematically remove AGs that have the highest frequency of observations without purchase in the dataset. Specifically, we removed AGs whose frequencies of no purchase constituted the worst 10% in terms of observations to create sample 1, and the worst 50% to create sample 2. Tables (1.1) and (1.2) compare the resulting samples to the original dataset. As shown in Table (1.1), to create sample 1, we systematically remove the worst 51 AGs (26.2%), which account for about 104 UPCs (11.8%) and 10.1% of observations. Table (1.2) shows the resulting dataset has approximately 1.5% zeros, which are dropped, so sample 1 excludes about 11.4% of observations in total. Sample 2 uses a more aggressive approach in removing AGs, eliminating the worst 119 AGs (61%), which account for 269 UPCs (30.4%)

Figure 1.1: Percent of Stores Reset over Time.



and 50% of the observations, resulting in 0.3% zeros, which are dropped. Sample 2 therefore excludes a total of about 49.9% of observations from the original dataset.

Table 1.1: Comparison of Samples 1 and 2 from Original Data

	All Data		Sample 1		Sample 2	
	Count	Percent	Count	Percent	Count	Percent
UPCs	884	100%	780	88.2%	615	69.6%
Product Groups (AGs)	195	100%	144	73.8%	76	39.0%
Observations	2154734	100%	1937602	89.9%	1078015	50.0%
Zero Purchase Obs. (Dropped)	50897	100%	28883	56.7%	3384	6.6%

Note while a relatively large number of AGs are removed in samples 1 and 2, the corresponding proportions of UPCs and observations dropped are small by comparison. We retain a larger percentage of UPCs and observations because quite a few AGs account for only one UPC, especially at the observation level, which captures an AG at a particular store, in a particular two-week period. The histogram in Figure (1.2) displays the number of observations in samples 1 and 2 by the number of UPCs captured in the observation. As expected, sample 1, which removed far fewer low-selling AGs, has a much higher proportion of observations consisting of a single UPC. While Figure (1.2) displays a large range in the number of UPCs per observation, we utilize our aggregation method to capture the defining product characteristics.

Relative to their presence on store shelves, the products in sample 2 are purchased more frequently than the products in sample 1. We would expect results from sample 1 to be more representative of the average for all ice cream products than sample 2, since sample 1 includes products with both high and low sales volumes. Sample 2 will tend to capture the effects on the more popular product categories, and hence we refer to the products in sample 2 as staples.

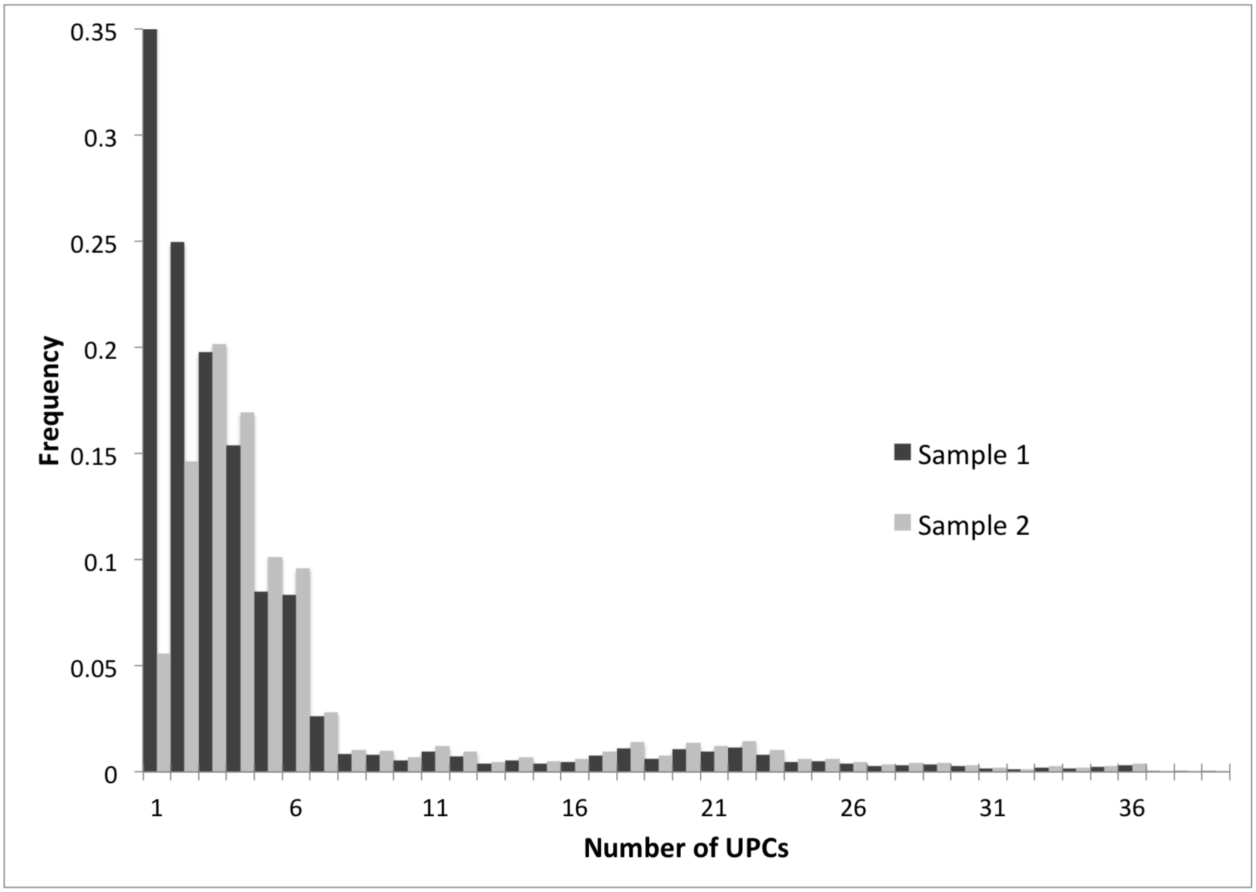
The raw data on revenue consist of net and gross values, which are equivalent for products not on promotion. When a product is under promotion, the net revenue captures the discounted total paid by customers. Unit prices are derived by dividing each revenue

Table 1.2: Comparison of Samples 1 and 2

	Sample 1		Sample 2	
	Count	Percent	Count	Percent
Total Observations	1937602	100%	1078015	100%
Zero Purchase Obs. (Dropped)	28883	1.5%	3384	0.3%
Obs. used in Estimations	1908719	98.5%	1074631	99.7%

Samples 1 and 2 formed by removing low-selling AGs from the original data. Sample 1 includes more of the original dataset, but drops more zero purchase observations. Sample 2 removes more low-selling AGs, thus focusing on high-selling products.

Figure 1.2: Distribution of number of UPCs in AG by observation



Aggregation groups (AGs) are product groups formed by brand, size, and subsubclass (see text). Observations are at the store, AG, two-week period level. Sample 1 is more representative of all ice cream products, while sample 2 removes more low-selling AGs.

variable by the quantity sold. Because the price data derive from transactions, the dataset has missing prices from the shelf data when a product is not purchased. In Section (1.3), we discuss how the retail setting allows us to utilize fixed effects in place of prices, and demonstrate that price changes do not bias effects identified in our results.

We now consider trends in the data in the before reset versus after reset periods. Table (1.3) displays the means and standard errors of the purchase quantities, number of ice cream UPCs in store, shelf schematics, average temperature, and change in the number of AGs per store for both samples, before reset versus in the eight week period after a reset. The observation level for these statistics is the same as in our tests, aggregated at the two-week, store, AG level. In general, standard errors are small due to the large number of observations in our dataset. In both samples, the average number of products purchased per AG increased after a reset, from 42.9 products purchased per week to 45.0 products purchased per week in sample 1, and from 67.1 products per week to 71.2 products per week in sample 2. However, our regression analysis will demonstrate that this change is not due to the reset itself. It is likely that these summary statistics simply capture average increases in ice cream purchases due to seasonal effects. Sample 2 has a higher average number of purchases per AG than sample 1, consistent with the fact that sample 2 excludes more the low-selling AGs.

In the eight-week period after a reset, there was a decrease in the average number of UPCs on shelves in store each week, from over 410 products to just below 400 products in both samples. The average number of ice cream facings wide, high, and deep allotted to a product group also tended to decrease after a reset, as a percentage of the total facings allocated to ice cream. The decrease in facings and decrease in total store UPCs could derive from a decrease in the size of the ice cream section of stores or from an increase in the allotment to larger sized ice cream products after resets.

In our analyses, we use the percent change in choices (UPCs) in an AG as a result of a reset, denoted the variable PCC_{is}. We also separate strictly positive changes and strictly negative changes to determine asymmetric effects from adding versus removing products. Tables (1.4) and (1.5) provide summary statistics of changes in variety through the control period (Before) and from a reset (After). In these tables, for the after reset treatment period we only include the two-week period including the reset for each store. Statistics for before and after a reset are shown separately for samples 1 and 2, with the significance between the means before versus after reset displayed on the estimates in the "Before" columns.

Table (1.4) displays the frequency of an increase or decrease in variety before and from a reset. In the control period, about 4.8% and 6.8% of observations had an increase in variety relative to the last two-week period, in samples 1 and 2 respectively. Both samples experienced a significant increase in the frequency of positive changes, to 9.9% and 13.0%, at resets. In both samples, the frequency of decreases in variety was higher in the control period relative to the frequency of increases in variety, at 5.9% and 8.3%. Again, the frequency of decreases in variety increased during resets, to 9.0% and 13.3% in samples 1 and 2. Thus, stores introduced and removed products outside of resets, but the changes were frequent pervasive than in the reset periods.

The statistics in Table (1.5) describe the magnitude of these variety changes. In the control period, AGs experienced a 0.8% average increase in products in sample 1 and a 1% increase in products in sample 2. During reset, however, these number jump to 2.4% and 1.2%, although this jump is only significant for sample 1. The "% Positive Change" and "% Negative Change" rows in Table (1.5) isolate strictly positive and negative changes. Except for positive changes in sample 2, the magnitude of all changes was significantly larger in a reset period than in the control period. In sample 1, the average percentage increase in variety increased from 45.6% to 47.3% at a reset, and the average percentage decrease moved from -23.8% to -25.6%. In sample 2, the average increase in variety decreased from 39.7% to 31.3% and the average decrease in variety also decreased from -20.4% to -21.9%.

In summary, the reset period experienced a significant increase in total quantity purchased as shown in Table (1.3). The overall number of UPCs in store also increased significantly after the resets, which might suggest that introducing more variety leads to increased sales. However, AGs did not experience a uniform increase in products, with the magnitudes

Table 1.3: Summary Statistics, Before and After Reset

	Sample 1				Sample 2			
	Before		After		Before		After	
	mean	se	mean	se	mean	se	mean	se
Total Quantity Purchased	42.863***	0.063	45.039	0.124	67.061***	0.102	71.178	0.207
No. UPCs in Store	417.07***	0.053	399.40	0.082	413.38***	0.073	396.28	0.115
AG Facings Wide (% of ice cream in store)	1.126***	0.001	1.107	0.003	1.757***	0.002	1.729	0.0024
AG Facings High (% of ice cream in store)	1.135***	0.001	1.114	0.002	1.748***	0.002	1.724	0.004
AG Facings Deep (% of ice cream in store)	1.124**	0.000	1.117	0.002	1.597	0.002	1.599	0.003
Avg. Temp (deg F, 000s)	53.01***	0.013	53.39	0.016	52.92***	0.017	53.39	0.022
% Change in AGs in Store	0.186***	0.031	-0.217	0.007	0.124***	0.040	-0.172	0.009
N	1489854		418865		842061		232570	

Two-sided t-tests on means before versus after the reset (within sample) were performed. Results indicated on the mean of the before period, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. Observations at the store, two-week period, AG level. Here, "After Reset" is the eight week period following the shelving reset.

Table 1.4: Frequency of Changes in Variety, Before and After Reset

	Sample 1				Sample 2			
	Before		After		Before		After	
	mean	se	mean	se	mean	se	mean	se
Positive Change	0.048***	0.000	0.099	0.001	0.068***	0.000	0.130	0.001
Negative Change	0.059***	0.000	0.090	0.001	0.083***	0.000	0.133	0.001
N	1474701		106063		837593		58365	

Observations at the store, two-week period, AG level. "After Reset" is the two-week period during which the shelving reset occurred. Changes in variety calculated for each two-week period relative to last two-week period. Two-sided t-tests on means before versus after the reset (within sample) were performed. Results indicated on the mean of the before period.* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 1.5: Percent Changes in Variety Relative to Last Two-Week Period, Before and After Reset

	Sample 1				Sample 2			
	Before		After		Before		After	
	mean	se	mean	se	mean	se	mean	se
% Change in UPCs (00s)	0.008***	0.000	0.024	0.001	0.010	0.000	0.012	0.001
% Positive Change (00s)	0.456**	0.002	0.473	0.004	0.397***	0.002	0.313	0.004
% Negative Change (00s)	-0.238***	0.001	-0.256	0.002	-0.204***	0.001	-0.219	0.002

Observations at the store, two-week period, AG level. "After Reset" is the two-week period during which the shelving reset occurred. Two-sided t-tests on means before versus after the reset (within sample) were performed. Results indicated on the mean of the before period, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

of both product introductions and removals increasing during resets, and these summary statistics clearly do not control for bias from seasonal variation, changes in shelf schematics, or store characteristics, to name a few. In addition, shelving schematics, such as the number of facings allocated to a product, and average temperatures were significantly different between the control and treatment periods. These facts motivate a formal regression analysis to determine the effects of changes in variety. Our retail setting, with resets rolled out over time, supports a difference in differences approach.

1.3 Empirical Strategy

In order to estimate the effect of the assortment size changes on quantity sold, we implement a difference in differences approach utilizing the shelving reset as treatment. Changes in sales trends in stores due to a reset are compared to trends in control stores that have not yet gone through a reset. The panel allows us to control for differences between treated and control stores and for differences between time periods. These controls are implemented with (1) store fixed effects that will control for observed and unobserved constant differences in determinants of demand at the store level and (2) seasonal two-week fixed effects that are common to all products at all stores. In all tests described below, we utilize the log of quantity as the dependent variable, allowing us to interpret coefficients as the percentage change in sales. We also use robust standard errors clustered at the store level, to correct for heteroskedasticity and within-store correlation in the error term.

Because price data from our retailer are only available from transactions, we do not have price data for each observation of products on the shelf. However, as discussed above, in our retail setting, stores change prices simultaneously each week, and prices are uniform among stores within a geographic division. In our difference in differences approach, we include store fixed effects and two-week period fixed effects. Consequently, the measured treatment effects derive from comparing trends in treated stores versus stores before treatment, and we omit prices from our tests. Our identification of the effects from changes in variety would be confounded with pricing changes if stores change price in the same period as a reset. We investigate this possibility with the following test:

$$(net\ price)_{usw} = \alpha_0 + \beta_1 reset_{sw} + d_s + c_w + c_u + \epsilon_{usw}, \tag{1.1}$$

where $(net\ price)_{usw}$ is the mode net price paid for UPC u, at store s in two week period w. We include the treatment dummy $reset_{sw}$ equal to one only if store s has a reset in the two week period w, and fixed effects for division (d_s), period (c_w), and UPC (c_u). Errors are clustered at the store level. Table (1.6) contains the results of specification (1.1).

The lack of a significant relationship between the reset dummy and net prices suggest that prices do not significantly change at the time of reset. In addition to this test, we later discuss a robustness test where we utilize division and period interactions in place of store fixed effects, that should capture any price changes, given our retail setting.

Table 1.6: Prices on Reset Week

	Net Unit Price
	b/se
resetwk2009	0.001
	(0.001)
Constant	3.546***
	(0.003)
Adjusted R^2	0.751
Observations	42093864

Standard errors in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Tables (1.4) and (1.5) demonstrated that the reset treatments included both variety extensions and reductions. Along with analyzing the average effect of a shelving reset, these differential types of shelf changes will allow us to estimate heterogeneity with respect to the type of shelf space change. We next discuss the regression specifications we utilize to investigate changes in variety.

1.3.1 Difference in Differences Specifications for Shelving Reset Average Effects

We utilize a difference in differences approach to find the effect of the changes in assortment on the quantity sold. The treatment effect is determined by comparing stores that have completed a shelving reset to stores that have not yet undergone a reset. In each regression, an observation describes a particular product aggregation group (AG, see description above) available on the shelf, and the outcome of interest is the log of the total quantity sold for the given AG. The data are by two-week period and by store for each AG.

We start by running the following difference in differences regression:

$$\begin{aligned} \log Q_{isw} = \alpha_0 + \beta_0 reset_{sw} + \gamma_1 width_{isw} + \gamma_3 temp_{sw} \\ + c_w + c_s + c_i + \epsilon_{isw}, \end{aligned} \tag{1.2}$$

where Q_{isw} is the quantity of UPCs sold in AG i in store s in the two-week period w. We include fixed effects for each two-week period, store, and AG: c_w, c_s, and c_i, respectively. Variable c_w captures any unobserved heterogeneity between each two-week period that is common among products and stores, variable c_s captures heterogeneity between stores but common across periods, and variable c_i captures heterogeneity between AGs, but common across periods and stores. The shelving variable $width_{isw}$ provides the amount of product facings wide allocated to an AG on the shelves, as a percentage of the total product facings wide allocated to all ice cream in the store that week.[3] Variable $temp_{sw}$ is the average

[3]Our data also include the shelving allocations for shelf depth, height, and capacity in square inches, but

temperature (degrees Fahrenheit) at the nearest available weather station for each store, derived from the National Climactic Data Centers daily surface data.[4]

The coefficient of interest is β_0, where dummy $reset_{sw}$ equals one only for products in stores that have undergone the shelving reset and only in the treatment period. Thus, the coefficient on $reset_{sw}$ will provide the average effect of the shelving reset on sales. Because consumers may slowly adjust to shelving changes, especially since resets occurred in the winter when demand for ice cream is relatively low, there may be a delayed effect on sales from a change in the ice cream assortment. In order to capture longer-term average effects, we test several different definitions of the treatment period: a reset in the past two weeks, four weeks, six weeks or eight weeks.

We chose eight weeks as the maximum treatment length for several reasons. The treatment rollout occupied a total of nearly nine weeks, after which time all stores in our dataset had been reset. Thus, observations for stores more than eight weeks past their reset dates primarily coincide with time periods when few, if any, comparison stores are in the pre-treatment state. More importantly, a treatment dummy defined by a long treatment period relative to the treatment rollout will approximate a seasonal dummy, absorbing the seasonal sales patterns of the months following the rollout period, when ice cream sales increase due to warmer weather. We therefore define a treatment period that is comparable to the rollout length. Observations for stores that have been treated and are beyond the given treatment period are dropped from the test, since these stores are neither viable as controls (they have undergone the yearly reset) nor as treatments (the treatment occurred a relatively long time ago, and seasonal changes are now dominating sales effects). Because the timing of the treatment rollout is unrelated to store sales characteristics, dropping stores after the treatment period should not introduce bias.

We next test the following lagged specification:

$$\begin{aligned}\log Q_{isw} = \alpha_0 + \sum_{l=0,1,2,3} \beta_{1,l}(resetlag)_{s(w-2l)} + \gamma_1 width_{isw} + \gamma_3 temp_{sw} \\ + c_w + c_s + c_i + \epsilon_{isw}.\end{aligned} \tag{1.3}$$

In specification (1.3), we replace variable $reset_{sw}$ from specification (1.2) with lagged reset dummies. Because w is a two-week period, specification (1.3) includes a dummy for the two-week period of a reset ($l = 0$), and dummies for the next three (two-week) periods. Therefore, the lagged specification covers the maximum treatment period length used in specification (1.2), but provides the effect on sales over time.

1.3.2 Specifications for Changes in Variety on Sales

Further, we test for different effects when resets increased versus decreased variety. We modify specification (1.2) to allow the average effect to differ by the sign of the change in

these variables are extremely correlated with the variables used, and are therefore omitted.

[4] http://hurricane.ncdc.noaa.gov/cdo/info.html.

variety:

$$\begin{aligned}\log Q_{isw} = \alpha_0 &+ \beta_1 reset_{sw} + \beta_2 (PCC_{is} > 0)^* reset_{sw} + \beta_3 (PCC_{is} < 0)^* reset_{sw} \\ &+ \beta_4 (PCC_{is} > 0) + \beta_5 (PCC_{is} < 0) \\ &+ \gamma_1 width_{isw} + \gamma_3 temp_{sw} \\ &+ \gamma_4 (AGchange)_{sw} + \gamma_5 (AGchange)_{sw}\, {}^* reset_{sw} \\ &+ c_w + c_s + c_i + \epsilon_{isw}\end{aligned} \tag{1.4}$$

The dummy variable $(PCC_{is} > 0)$ is equal to one only when the reset in store s led to a strict increase in the number of choices (we use PCC to denote the "percent change in choices," which will be used in a later test) in product group i, while $(PCC_{is} < 0)$ equals one only when the reset led to a strict decrease in the number of choices. We refer to $(PCC_{is} > 0)$ as the "Positive Group" dummy and $(PCC_{is} < 0)$ as the "Negative Group" dummy. Thus, β_1 provides the average effect of a reset on AGs that have no change in variety, while β_2 and β_3 decompose the effect of a nonzero change in variety by the sign of the change. The coefficients on the lower order variables $(PCC_{is} > 0)$ and $(PCC_{is} < 0)$ capture any differences during the control period between the product groups that had increases and decreases in choices at reset. As in specification (1.2), we separately test reset treatment periods of two, four, six, and eight weeks.

We also add the variable $(AGchange)_{sw}$, both independently and interacted with the reset dummy, and refer to it as the "% Change in AGs in store." This variable provides the period-to-period percent change in the total number of AGs in store, to control for losses or gains of entirely new product groups.

Again, we test a lagged version of specification (1.4):

$$\begin{aligned}\log Q_{isw} = \alpha_0 &+ \sum_{l=0,1,2,3} \Big(\beta_{1,l} reset_{s(w-2l)} + \beta_{2,l} (PCC_{is} > 0)^* reset_{s(w-2l)} \\ &\qquad + \beta_{3,l} (PCC_{is} < 0)^* reset_{s(w-2l)} \Big) \\ &+ \beta_4 (PCC_{is} > 0) + \beta_5 (PCC_{is} < 0) \\ &+ \gamma_1 width_{isw} + \gamma_3 temp_{sw} \\ &+ \gamma_4 (AGchange)_{sw} + \gamma_5 (AGchange)_{sw}\, {}^* reset_{sw} \\ &+ c_w + c_s + c_i + \epsilon_{isw},\end{aligned} \tag{1.5}$$

where $\beta_{1,l}$, $\beta_{2,l}$, and $\beta_{3,l}$ for $l = 0, 1, 2, 3$ provide the effect of resets on AGs that had no variety change, the effect on groups that increased variety during a reset, and the effect on groups that reduced variety during a reset, respectively, for three subsequent two-week periods.

Our preferred specification includes the percent change in variety from the store's reset to determine how the relative magnitude of the change in variety affects sales. We add variable PCC_{is}, the percent change in the number of choices in AG i from the reset in store s. Specifically, let r_s be the period that store s is reset, and let $(\overline{no.\ UPCs})_{ist}$ be the average number of UPCs in AG i, store s during the treatment period t. As discussed in Section (1.3), we test several different interval lengths as the after reset treatment periods in our difference in differences specifications. Then,

$$PCC_{is} = \frac{(\overline{no.\ UPCs})_{ist} - (\text{no. of UPCs in } i,\ s,\ r_s - 1)}{\text{no. of UPCs in } i,\ s,\ r_s - 1}.$$

Using this definition, we test the following:

$$\begin{aligned}\log Q_{isw} = \alpha_0 &+ \beta_1 reset_{sw} + \beta_2 (PCC_{is} > 0)^* reset_{sw} + \beta_3 (PCC_{is} < 0)^* reset_{sw} \\ &+ \beta_4 (PCC_{is} > 0) + \beta_5 (PCC_{is} < 0) \\ &+ \beta_6 (PCC_{is} > 0)^* PCC_{is}\, {}^* reset_{sw} + \beta_7 (PCC_{is} < 0)^* PCC_{is}\, {}^* reset_{sw} \\ &+ \beta_8 (PCC_{is} > 0)^* PCC_{is} + \beta_9 (PCC_{is} < 0)^* PCC_{is} \\ &+ \gamma_1 width_{isw} + \gamma_3 temp_{sw} \\ &+ \gamma_4 (AGchange)_{sw} + \gamma_5 (AGchange)_{sw}\, {}^* reset_{sw} \\ &+ c_w + c_s + c_i + \epsilon_{isw}.\end{aligned} \tag{1.6}$$

In our results, we refer to $(PCC_{is} > 0)^* PCC_{is}$ as the "% Positive Change" and to $(PCC_{is} < 0)^* PCC_{is}$ as the "% Negative Change," although the units of this variable are in fractional terms (i.e. $PCC_{is} = 1.00$ is equivalent to a 100% increase). Coefficient β_6 provides effect on the log of sales from a 100% increase in the number of UPCs in a product group during the treatment period, while β_8 provides the same measure during the control period. The equivalent measures for decreases in variety are given by β_7 and β_9.

The lagged version of (1.6) as follows:

$$\begin{aligned}\log Q_{isw} = \alpha_0 & \\ &+ \sum_{l=0,1,2,3} \Big(\beta_{1,t} reset_{sw} \\ &\qquad + \beta_{2,t} (PCC_{is} > 0)^* reset_{sw} + \beta_{3,t} (PCC_{is} < 0)^* reset_{sw} \Big) \\ &+ \beta_4 (PCC_{is} > 0) + \beta_5 (PCC_{is} < 0) \\ &+ \sum_{l=0,1,2,3} \Big(\beta_{6,t} (PCC_{is} > 0)^* PCC_{is}\, {}^* reset_{sw} \\ &\qquad + \beta_{7,t} (PCC_{is} < 0)^* PCC_{is}\, {}^* reset_{sw} \Big) \\ &+ \beta_8 (PCC_{is} > 0)^* PCC_{is} + \beta_9 (PCC_{is} < 0)^* PCC_{is} \\ &+ \gamma_1 width_{isw} + \gamma_3 temp_{sw} \\ &+ \gamma_4 (AGchange)_{sw} + \gamma_5 (AGchange)_{sw}\, {}^* reset_{sw} \\ &+ c_w + c_s + c_i + \epsilon_{isw},\end{aligned} \tag{1.7}$$

where we again use three lags of the two-week reset period.

1.3.3 Robustness Tests for Price Changes

We showed above that the resets do not significantly predict price changes in our dataset. We perform an additional test to ensure our results are qualitatively robust to possible price changes. Specifically, we test the following specification:

$$\begin{aligned}\log Q_{isw} = \alpha_0 &+ \beta_1 reset_{sw} + \beta_2(PCC_{is} > 0)^* reset_{sw} + \beta_3(PCC_{is} < 0)^* reset_{sw} \\ &+ \beta_4(PCC_{is} > 0) + \beta_5(PCC_{is} < 0) \\ &+ \beta_6(PCC_{is} > 0)^* PCC_{is}\,^* reset_{sw} + \beta_7(PCC_{is} < 0)^* PCC_{is}\,^* reset_{sw} \\ &+ \beta_8(PCC_{is} > 0)^* PCC_{is} + \beta_9(PCC_{is} < 0)^* PCC_{is} \\ &+ \gamma_1 width_{isw} + \gamma_3 temp_{sw} \\ &+ \gamma_4(AGchange)_{sw} + \gamma_5(AGchange)_{sw}\,^* reset_{sw} \\ &+ c_w + d_s + d_s\,^* c_w + c_i + \epsilon_{isw},\end{aligned} \tag{1.8}$$

where d_s is the geographic division that store s is located in. The analogous lagged specification is

$$\begin{aligned}\log Q_{isw} = \alpha_0 & \\ &+ \sum_{l=0,1,2,3} \Big(\beta_{1,t} reset_{sw} \\ &\qquad + \beta_{2,t}(PCC_{is} > 0)^* reset_{sw} + \beta_{3,t}(PCC_{is} < 0)^* reset_{sw}\Big) \\ &+ \beta_4(PCC_{is} > 0) + \beta_5(PCC_{is} < 0) \\ &+ \sum_{l=0,1,2,3} \Big(\beta_{6,t}(PCC_{is} > 0)^* PCC_{is}\,^* reset_{sw} \\ &\qquad + \beta_{7,t}(PCC_{is} < 0)^* PCC_{is}\,^* reset_{sw}\Big) \\ &+ \beta_8(PCC_{is} > 0)^* PCC_{is} + \beta_9(PCC_{is} < 0)^* PCC_{is} \\ &+ \gamma_1 width_{isw} + \gamma_3 temp_{sw} \\ &+ \gamma_4(AGchange)_{sw} + \gamma_5(AGchange)_{sw}\,^* reset_{sw} \\ &+ c_w + d_s + d_s\,^* c_w + c_i + \epsilon_{isw}.\end{aligned} \tag{1.9}$$

Equations (1.8) and (1.9) replace store fixed effects with division fixed effects and add an interaction between division and (two-week) period fixed effects. Because prices are uniform within geographic division and any price changes occur simultaneously for all stores on the same day each week, the interaction between division and period fixed effects should capture any price changes.

1.4 Results

We first present the results from the difference in differences identification of the average ice cream quantity responses due to the shelving reset. Next we present the effects from an increase in assortment versus a decrease in assortment. Then we determine how the relative magnitude of an increase versus a decrease affects sales, followed by results of robustness tests to rule out bias from omitted price variables. When discussing results in the text, percent effects will always be the exponentiated predicted values, corrected for bias (**?**).

1.4.1 Average Effects of the Shelving Reset

The results from the difference in differences test on the average effect of a shelving reset are presented in Tables (1.7) and (1.8), for samples 1 and 2, respectively. In both tables, each column shows the results for a different definition of the treatment period: two weeks (column (1)), four weeks (column (2)), six weeks (column (3)), and eight weeks (column (4)). The estimate on the reset variable is negative and significant in both samples, implying the average effect of a reset was a reduction in sales. Within each sample, the estimated effect of a reset is relatively stable between different definitions of treatment period length. In sample 1, resets led to an average decrease in sales of about 1%, while the decrease was slightly larger in sample 2, at about 1.6%.

The number of UPCs in a store was also significantly related to sales, with an increase of 100 UPCs of ice cream correlated with an approximate increase in sales of 10-11% in sample 1 and nearly 15-16% in sample 2.[5] While we cannot use the specifications in this paper for a causal interpretation of the three facings schematics variables, generally a larger allocation to a product group is correlated with higher sales in that group (an exception is the facings wide statistic for sample 1, which was not significantly different from zero for any of the treatment period definitions). However, care should be taken when considering the actual point estimates, since the three facings variables were strongly collinear, with variance inflation factors of over 5 in both samples 1 and 2. Finally, the average temperature variable indicates that a temperature increase of one degree is correlated with increased sales of about 0.16% (correcting for scale).

The results of the lagged specification in Tables (1.9) and (1.10) show that the reduced sales effect persisted longer for sample 2 than for sample 1. Within a sample, each column in Tables (1.9) and (1.10) shows results from iteratively adding lags to column (1) of Tables (1.7) and (1.8). In both tables, column (1) adds a one period (two-week) lag, and column (3) includes lags for up to three two-week periods, therefore covering the effects of any reset within the past eight weeks. In sample 1, resets significantly affect sales only in the two-week period that a reset takes place, implying the significance of the reset variables in all the non-lagged specifications are primarily from effects in the first two weeks. In sample 2, the effect remains significant through the four-week lag, with the magnitude of the effect

[5] Percentage effects in text are exponentiated and corrected for bias (Kennedy, 1983)

Table 1.7: Log Total Quantity on Reset, Sample 1

	(1)	(2)	(3)	(4)
Treatment Length (weeks)	2	4	6	8
Reset	-0.010*	-0.010*	-0.010*	-0.010*
	(0.005)	(0.005)	(0.005)	(0.005)
Facings wide, % of total ice cream	-0.011	-0.007	-0.004	0.000
	(0.020)	(0.019)	(0.018)	(0.017)
Facings deep, % of total ice cream	0.159**	0.163***	0.165***	0.166***
	(0.050)	(0.048)	(0.047)	(0.045)
Facings high, % of total ice cream	0.156***	0.154***	0.154***	0.153***
	(0.017)	(0.016)	(0.016)	(0.015)
Avg. temperature (deg F, 00s)	0.165***	0.155***	0.155***	0.164***
	(0.027)	(0.027)	(0.026)	(0.025)
No. of UPCs in store (00s)	0.097***	0.100***	0.102***	0.103***
	(0.017)	(0.016)	(0.016)	(0.015)
Constant	2.339***	2.317***	2.297***	2.277***
	(0.077)	(0.074)	(0.071)	(0.068)
Adjusted R^2	0.762	0.762	0.761	0.760
Observations	1595917	1700658	1804712	1908719
Store FEs	Y	Y	Y	Y
Two-week FEs	Y	Y	Y	Y
AG FEs	Y	Y	Y	Y

Standard errors in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 1.8: Log Total Quantity on Reset, Sample 2

	(1)	(2)	(3)	(4)
Treatment Length (weeks)	2	4	6	8
Reset	-0.016***	-0.017***	-0.016***	-0.016***
	(0.005)	(0.005)	(0.005)	(0.005)
Facings wide, % of total ice cream	0.039*	0.041*	0.043**	0.047**
	(0.017)	(0.017)	(0.016)	(0.015)
Facings deep, % of total ice cream	0.084	0.089*	0.093*	0.094*
	(0.043)	(0.042)	(0.040)	(0.039)
Facings high, % of total ice cream	0.147***	0.144***	0.143***	0.142***
	(0.014)	(0.013)	(0.013)	(0.012)
Avg. temperature (deg F, 00s)	0.169***	0.156***	0.151***	0.150***
	(0.028)	(0.027)	(0.026)	(0.026)
No. of UPCs in store (00s)	0.144***	0.146***	0.146***	0.147***
	(0.023)	(0.022)	(0.021)	(0.020)
Constant	2.078***	2.067***	2.059***	2.049***
	(0.094)	(0.090)	(0.086)	(0.082)
Adjusted R^2	0.774	0.773	0.773	0.772
Observations	900,426	958,568	1,016,593	1,074,631
Store FEs	Y	Y	Y	Y
Two-week FEs	Y	Y	Y	Y
AG FEs	Y	Y	Y	Y

Standard errors in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

actually increasing during this period, although the difference between the estimates on the lags is not statistically significant. The significant effect on higher order lags in sample 2 suggests that the reset effect persisted longer for ice cream staples than for ice cream in general. The estimates of the control variables are qualitatively similar as in the results for the non-lagged specification.

Table 1.9: Log Total Quantity on Reset with Lags, Sample 1

	(1) b/se	(2) b/se	(3) b/se
Reset	-0.010*	-0.010*	-0.010*
	(0.005)	(0.005)	(0.005)
Reset, lag 2 weeks	-0.012	-0.011	-0.011
	(0.007)	(0.007)	(0.007)
Reset, lag 4 weeks		-0.016	-0.016
		(0.010)	(0.010)
Reset, lag 6 weeks			-0.015
			(0.012)
Facings wide, % of total ice cream	-0.007	-0.004	0.000
	(0.019)	(0.018)	(0.017)
Facings deep, % of total ice cream	0.163***	0.165***	0.166***
	(0.048)	(0.047)	(0.045)
Facings high, % of total ice cream	0.154***	0.154***	0.153***
	(0.016)	(0.016)	(0.015)
Avg. temperature (deg F, 00s)	0.155***	0.155***	0.164***
	(0.027)	(0.026)	(0.025)
No. of UPCs in store (00s)	0.100***	0.102***	0.103***
	(0.016)	(0.016)	(0.015)
Constant	2.318***	2.297***	2.277***
	(0.074)	(0.071)	(0.068)
Adjusted R^2	0.762	0.761	0.760
Observations	1,700,658	1,804,712	1,908,719
Store FEs	Y	Y	Y
Two-week FEs	Y	Y	Y
AG FEs	Y	Y	Y

Standard errors in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Because many changes are made during a reset, we cannot attribute the average effect of a reset simply to a change in variety. The negative effect on sales could be due to confusion in response to the general rearrangement of products, although we have controlled for changes in the amount of space allocated to each product group. In the next section, we therefore break down the reset variable into increases versus decreases in variety.

Table 1.10: Log Total Quantity on Reset with Lags, Sample 2

	(1) b/se	(2) b/se	(3) b/se
Reset	-0.016***	-0.016***	-0.016***
	(0.005)	(0.005)	(0.005)
Reset, lag 2 weeks	-0.019**	-0.018*	-0.018*
	(0.007)	(0.007)	(0.007)
Reset, lag 4 weeks		-0.023*	-0.022*
		(0.010)	(0.010)
Reset, lag 6 weeks			-0.021
			(0.013)
Facings wide, % of total ice cream	0.041*	0.043**	0.047**
	(0.017)	(0.016)	(0.015)
Facings deep, % of total ice cream	0.089*	0.093*	0.094*
	(0.042)	(0.040)	(0.039)
Facings high, % of total ice cream	0.144***	0.143***	0.142***
	(0.013)	(0.013)	(0.012)
Avg. temperature (deg F, 00s)	0.156***	0.151***	0.150***
	(0.027)	(0.026)	(0.026)
No. of UPCs in store (00s)	0.146***	0.146***	0.147***
	(0.022)	(0.021)	(0.020)
Constant	2.067***	2.059***	2.049***
	(0.090)	(0.086)	(0.082)
Adjusted R^2	0.773	0.773	0.772
Observations	958,568	1,016,593	1,074,631
Store FEs	Y	Y	Y
Two-week FEs	Y	Y	Y
AG FEs	Y	Y	Y

Standard errors in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

1.4.2 Sales Effects from Changes in Variety

We now investigate the differential effects of an increase versus a decrease in product assortment. Tables (1.11) and (1.12) provide the results of specification (1.4) for samples 1 and 2 separately, and are organized analogously to Tables (1.7) and (1.8). Successive columns within a sample utilize treatment periods of two, four, six, and finally eight weeks. The corresponding results for lagged specification (1.5) are presented in Tables (1.13) and (1.14).

In these regressions, the estimates on the reset variables in Tables (1.11) through (1.14) only indicate the effect of a shelving reset on AGs that did not change variety during the given store's reset. Thus, the reset dummy captures all changes to sales of a product group from a reset, excluding effects from a strict change in variety to the given product group (equivalently, excluding a nonzero change to the number of products in the given product group). Using this definition, the estimate of the reset variable may include changes in shelf location or cross effects from introducing or removing products in other AGs.

Tables (1.11) and (1.12) show for both samples 1 and 2, the estimate of the reset variable only becomes significant for longer definitions of the treatment period, where the resets had a positive effect on sales of product groups that had no change in variety. For the six-week definition of the reset treatment, resets increased sales by approximately 1.9% and 2.6% in samples 1 and 2, respectively. The results of the corresponding lagged reset variables in Tables (1.13) and (1.14) are not significant for any of the individual lags, suggesting that the significance of resets for long definitions of the reset variable derives from cumulative effects during the post-treatment weeks.

In the non-lagged results shown in Tables (1.11) and (1.12), the signs of the estimates from the change in variety variable "(Negative Group)*Reset" demonstrate that within a product group, decreasing variety decreased sales on average for all definitions of the treatment period. For "(Positive Group)*Reset", the estimates in sample 1 demonstrate that for all ice cream, variety increases led to increased sales between 2.4-5.8% in the first six weeks, and had negligible effects over the eight week treatment. For sample 2, adding choices significantly increased sales only in the four week treatment (by 2.7%), and significantly decreased sales in the eight week treatment (by 3.5%). The notable differences between the treatment definitions imply there may be important period-to-period changes in effects.

Results of the lagged specifications explain the opposing results between the eight week treatment period and the shorter treatment definitions in the non-lagged tests. We focus on the results in column (3) of Tables (1.13) and (1.14), which contains the fully lagged model. In sample 1, increasing variety led to a sales increase of about 4% in the first two-week period, but led to 5.9% decreased sales after an additional four weeks and 7.3% decreased sales in the six week lag period (the effects of the four-week and six-week lags are significantly different at the 95% level). In sample 2, both the immediate and the two-week lag effect are not significant, but the four and six week lags experience negative sales effect of 5.4% and 5.0% (the estimates on these lags are not significantly different from each other). Hence, decreases in sales followed mild or nonexistent increases in sales, translating to a negative

Table 1.11: Log Total Quantity on PCC Dummies, Sample 1

	(1)	(2)	(3)	(4)
Treatment Length (weeks)	2	4	6	8
	b/se	b/se	b/se	b/se
Reset	-0.003	0.010	0.019***	0.031***
	(0.006)	(0.006)	(0.005)	(0.005)
(Positive Group) * Reset	0.046***	0.056***	0.024*	-0.015
	(0.012)	(0.010)	(0.010)	(0.010)
Positive Group	-0.020**	-0.111***	-0.126***	-0.118***
	(0.006)	(0.007)	(0.007)	(0.007)
(Negative Group) * Reset	-0.130***	-0.167***	-0.161***	-0.173***
	(0.009)	(0.007)	(0.007)	(0.007)
Negative Group	-0.027***	-0.053***	-0.054***	-0.043***
	(0.006)	(0.005)	(0.005)	(0.005)
% Change in Store AGs	-0.000	-0.000	-0.000	-0.000
	(0.000)	(0.000)	(0.000)	(0.000)
(% Change in Store AGs)*Reset	-0.000	-0.001	-0.001	-0.001
	(0.001)	(0.001)	(0.001)	(0.001)
Facings wide, % of total ice cream	-0.010	-0.004	0.001	0.005
	(0.020)	(0.019)	(0.018)	(0.017)
Facings deep, % of total ice cream	0.158**	0.158**	0.158***	0.159***
	(0.050)	(0.048)	(0.046)	(0.045)
Facings high, % of total ice cream	0.156***	0.152***	0.151***	0.150***
	(0.017)	(0.016)	(0.015)	(0.015)
Avg. temperature (deg F, 00s)	0.166***	0.156***	0.157***	0.168***
	(0.027)	(0.027)	(0.026)	(0.025)
No. of UPCs in store (00s)	0.096***	0.098***	0.099***	0.100***
	(0.018)	(0.017)	(0.016)	(0.015)
Constant	2.342***	2.330***	2.312***	2.288***
	(0.078)	(0.075)	(0.071)	(0.068)
Adjusted R^2	0.762	0.762	0.762	0.761
Observations	1,595,917	1,700,658	1,804,712	1,908,719
Store FEs	Y	Y	Y	Y
Two-week FEs	Y	Y	Y	Y
AG FEs	Y	Y	Y	Y

Standard errors in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 1.12: Log Total Quantity on PCC Dummies, Sample 2

	(1)	(2)	(3)	(4)
Treatment Length (weeks)	2	4	6	8
	b/se	b/se	b/se	b/se
Reset	-0.006	0.014*	0.026***	0.040***
	(0.006)	(0.006)	(0.006)	(0.006)
(Positive Group) * Reset	0.020	0.027**	-0.001	-0.036***
	(0.012)	(0.010)	(0.010)	(0.010)
Positive Group	-0.019**	-0.094***	-0.105***	-0.095***
	(0.007)	(0.007)	(0.007)	(0.007)
(Negative Group) * Reset	-0.094***	-0.141***	-0.137***	-0.148***
	(0.009)	(0.007)	(0.007)	(0.006)
Negative Group	-0.032***	-0.058***	-0.057***	-0.046***
	(0.006)	(0.005)	(0.006)	(0.006)
% Change in Store AGs	0.000	0.000	0.000	-0.000
	(0.000)	(0.000)	(0.000)	(0.000)
(% Change in Store AGs)*Reset	-0.000	-0.001	-0.001	-0.001
	(0.001)	(0.001)	(0.001)	(0.001)
Facings wide, % of total ice cream	0.040*	0.044**	0.046**	0.050**
	(0.017)	(0.017)	(0.016)	(0.016)
Facings deep, % of total ice cream	0.084	0.086*	0.089*	0.090*
	(0.043)	(0.042)	(0.040)	(0.039)
Facings high, % of total ice cream	0.146***	0.142***	0.141***	0.139***
	(0.014)	(0.013)	(0.013)	(0.012)
Avg. temperature (deg F, 00s)	0.169***	0.157***	0.153***	0.154***
	(0.029)	(0.028)	(0.026)	(0.026)
No. of UPCs in store (00s)	0.144***	0.143***	0.143***	0.143***
	(0.024)	(0.022)	(0.021)	(0.019)
Constant	2.080***	2.079***	2.072***	2.057***
	(0.096)	(0.091)	(0.086)	(0.082)
Adjusted R^2	0.774	0.774	0.774	0.773
Observations	900,426	958,568	1,016,593	1,074,631
Store FEs	Y	Y	Y	Y
Two-week FEs	Y	Y	Y	Y
AG FEs	Y	Y	Y	Y

Standard errors in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 1.13: Log Total Quantity on PCC Dummies with Lags, Sample 1

	(1)		(2)		(3)	
	b	se	b	se	b	se
Reset	-0.004	(0.005)	-0.004	(0.005)	-0.004	(0.005)
Reset, lag 2 weeks	0.004	(0.007)	0.004	(0.007)	0.004	(0.007)
Reset, lag 4 weeks			0.000	(0.010)	0.000	(0.010)
Reset, lag 6 weeks					0.009	(0.012)
(Positive Group) * Reset	0.043***	(0.012)	0.041***	(0.011)	0.040***	(0.011)
(Positive Group) * Reset, lag 2 weeks	-0.020	(0.014)	-0.022	(0.014)	-0.024	(0.013)
(Positive Group) * Reset, lag 4 weeks			-0.060***	(0.014)	-0.061***	(0.013)
(Positive Group) * Reset, lag 6 weeks					-0.076***	(0.013)
Positive Group	-0.020**	(0.007)	-0.017**	(0.007)	-0.013	(0.007)
(Negative Group) * Reset	-0.129***	(0.009)	-0.128***	(0.009)	-0.127***	(0.009)
(Negative Group) * Reset, lag 2 weeks	-0.149***	(0.009)	-0.148***	(0.009)	-0.147***	(0.009)
(Negative Group) * Reset, lag 4 weeks			-0.117***	(0.009)	-0.116***	(0.009)
(Negative Group) * Reset, lag 6 weeks					-0.177***	(0.009)
Negative Group	-0.020***	(0.006)	-0.013*	(0.006)	-0.009	(0.006)
% Change in Store AGs	-0.000	(0.000)	-0.000	(0.000)	-0.000	(0.000)
(% Change in Store AGs)*Reset	-0.001	(0.001)	-0.001	(0.001)	-0.001	(0.001)
Facings wide, % of total ice cream	-0.006	(0.019)	-0.003	(0.018)	0.001	(0.017)
Facings deep, % of total ice cream	0.162***	(0.049)	0.164***	(0.047)	0.166***	(0.045)
Facings high, % of total ice cream	0.153***	(0.016)	0.153***	(0.016)	0.152***	(0.015)
Avg. temperature (deg F, 00s)	0.156***	(0.027)	0.157***	(0.026)	0.167***	(0.025)
No. of UPCs in store (00s)	0.099***	(0.017)	0.100***	(0.016)	0.101***	(0.015)
Constant	2.323***	(0.074)	2.303***	(0.071)	2.282***	(0.068)
Adjusted R^2	0.762		0.761		0.760	
Observations	1,700,658		1,804,712		1,908,719	
Store FEs	Y		Y		Y	
Two-week FEs	Y		Y		Y	
AG FEs	Y		Y		Y	

Standard errors in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 1.14: Log Total Quantity on PCC Dummies with Lags, Sample 2

	(1)		(2)		(3)	
	b	se	b	se	b	se
Reset	-0.008	(0.006)	-0.007	(0.005)	-0.007	(0.005)
Reset, lag 2 weeks	0.000	(0.008)	-0.000	(0.008)	0.000	(0.007)
Reset, lag 4 weeks			-0.006	(0.010)	-0.006	(0.010)
Reset, lag 6 weeks					0.005	(0.013)
(Positive Group) * Reset	0.018	(0.011)	0.016	(0.011)	0.015	(0.011)
(Positive Group) * Reset, lag 2 weeks	-0.020	(0.014)	-0.022	(0.013)	-0.023	(0.013)
(Positive Group) * Reset, lag 4 weeks			-0.054***	(0.014)	-0.055***	(0.013)
(Positive Group) * Reset, lag 6 weeks					-0.050***	(0.013)
Positive Group	-0.018**	(0.007)	-0.017*	(0.007)	-0.013	(0.007)
(Negative Group) * Reset	-0.093***	(0.009)	-0.093***	(0.009)	-0.092***	(0.009)
(Negative Group) * Reset, lag 2 weeks	-0.121***	(0.009)	-0.119***	(0.009)	-0.119***	(0.008)
(Negative Group) * Reset, lag 4 weeks			-0.068***	(0.009)	-0.067***	(0.009)
(Negative Group) * Reset, lag 6 weeks					-0.143***	(0.009)
Negative Group	-0.026***	(0.006)	-0.021***	(0.006)	-0.017**	(0.006)
% Change in Store AGs	0.000	(0.000)	0.000	(0.000)	0.000	(0.000)
(% Change in Store AGs)*Reset	-0.001	(0.001)	-0.001	(0.001)	-0.001	(0.001)
Facings wide, % of total ice cream	0.042*	(0.017)	0.044**	(0.016)	0.047**	(0.016)
Facings deep, % of total ice cream	0.089*	(0.042)	0.093*	(0.041)	0.094*	(0.040)
Facings high, % of total ice cream	0.144***	(0.013)	0.143***	(0.013)	0.141***	(0.012)
Avg. temperature (deg F, 00s)	0.157***	(0.027)	0.152***	(0.026)	0.152***	(0.026)
No. of UPCs in store (00s)	0.145***	(0.022)	0.145***	(0.021)	0.144***	(0.019)
Constant	2.071***	(0.090)	2.063***	(0.086)	2.054***	(0.082)
Adjusted R^2	0.774		0.773		0.772	
Observations	958,568		1,016,593		1,074,631	
Store FEs	Y		Y		Y	
Two-week FEs	Y		Y		Y	
AG FEs	Y		Y		Y	

Standard errors in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

average effect over the eight week treatment period.

Several mechanisms may explain the trend from increases in variety. First, immediate increases in ice cream purchases may have been compensated by a decrease in later sales. That is, consumers may simply substitute purchases within a product group over time. A second possible mechanism involves switching between product groups. Specifically, note that the average effect of a reset is more significant and positive for longer definitions of the treatment period. Because this variable captures sales in AGs that do not change variety, it's possible that consumers switched from product groups that have an increase in variety to product groups with no change in variety. However, sales increases for product groups with no changes in variety may also come from consumers switching away from products that had decreases in variety, or from consumer entry into the ice cream category. Lastly, some consumers may exit the ice cream category altogether in response to increases in variety, which is consistent with the overall negative effect of a reset on all ice cream products, as shown in Tables (1.7) and (1.8).

For decreases in variety, the results of the non-lagged specification on both samples 1 and 2 indicate the average effect led to a significant decrease in sales, for all definitions of the treatment period. Depending on the treatment period length, the magnitude of the effect from decreased variety ranges from -12.2% to -15.9% in sample 1 (a change of -0.130 and -0.173 in log sales), and from -9.0% to -13.8% (changes of -0.094 to -0.148 in log sales) in sample 2, with the size of the effect smallest for the two-week treatment and largest in the eight-week treatment in both samples. Results of the lagged specifications in Tables (1.13) and (1.14) also reveal that reductions in variety led to significant sales decreases in each of the four two-week periods in the treatment period, for both samples 1 and 2. In sample 1, the two and six week lags on the "(Negative Group)*Reset" variable are both significantly more negative than the initial effect, although the four week lag is not significantly different at the 95% confidence level. In sample 2, all differences between the estimate on the "(Negative Group)*Reset" variable and its lags are significant, and again the magnitude of the effect on sales is non-monotonic over time.

The lower order terms "Positive Group" and "Negative Group" provide the average sales on the AGs that experienced non-zero variety changes during the control period. Prior to their resets, these AGs had lower sales relative to the AGs that had no variety change, for both samples 1 and 2. In the lagged specifications with all four lags, however, this reduction is only significant for the negative group in sample 2. For these tests, we also added the control "% Change in Store AGs" and its interaction with the reset dummy, although these variables were not significant. The control variable on the facings allocated to each AG, temperature, and the total number of UPCs in store remain significant, with estimates comparable to those from Tables (1.7) and (1.8).

We now discuss the results from specifications (1.6) and (1.7), and explore how the magnitude of variety changes affects sales. Tables (1.15)-(1.19) contain the results for these specifications. The inclusion of the percent change variables "(% Positive Change)*Reset" and "(% Negative Change)*Reset" allow the treatment effect to depend linearly on the percent change in variety relative to the number of UPCs in an AG prior to reset, with separate effects for increases versus decreases in variety. For reference, we present the coefficients for the reset and change in variety variables recalculated as percent changes in sales in Tables (1.17) and (1.20) (Kennedy, 1983).

We first discuss the results of the non-lagged tests. Again, the reset dummies now provide the effect of a reset on AGs that have no change in variety, and the results are qualitatively similar to those in the last specification, displayed in Tables (1.11) and (1.12). For longer definitions of the treatment period, resets exhibit a significantly positive effect on sales, although the results of the lagged specifications in Tables (1.18) and (1.19) demonstrate that no individual two-week period has a significant increase in sales relative to the control period, for both samples 1 and 2.

To better illustrate the effects of variety changes, Figure (1.3) contains plots of the relevant estimates and confidence intervals from the non-lagged specifications corresponding to Tables (1.15) and (1.16), with separate sub charts for each of the treatment length definitions. The estimates for the "(% Positive Change)*Reset" and "(% Negative Change)*Reset" are positive and statistically significant, except for the "(% Negative Change)*Reset" variable in the two week treatment period for sample 2, which is not significantly different from zero. For these variables, positive signs indicate that a larger magnitude increase will increase sales more, and a larger magnitude decrease will decrease sales more. Quantitatively, for a given increase in variety, sales will increase by about 0.179 to 0.232 in log sales, or about 19.6-26.1% more in sample 1 for an additional 1% increase in variety (using the exponentiated and estimates corrected for bias). For a given decrease in variety, a 1% reduction in the decrease in variety will reduce the potential sales loss by 45.8-60.1%. Note that in sample 1, the estimate on "(% Negative Change)*Reset" is significantly larger than the estimate on "(% Positive Change)*Reset", for all treatment lengths. Thus, an equivalent decrease in variety reduces sales more than an equivalent increase in variety increases sales. The asymmetry does not hold in sample 2, where the estimates for "(% Positive Change)*Reset" and "(% Negative Change)*Reset" are not significantly different, with the sales effect of a 1% increase (decrease) in increased (decreased) variety ranging from 16.8 to 29.9% for '(% Positive Change)*Reset" and from 12.9 to 28.4% for "(% Negative Change)*Reset".

The "(Positive Group)*Reset" and "(Negative Group)*Reset" variables are all less than zero (although not all estimates are significantly different than zero). In combination with the positive estimates for "(% Positive Change)*Reset" and "(% Negative Change)*Reset", these negative estimates predict a decrease in sales for both small percentage increases in variety and decreases in variety, but increases in sales for relatively large increases in variety. For very small increases in variety, the "(Positive Group)*Reset" estimates predict that sales may decrease by up to approximately 9.8% on average over the eight weeks after a reset, although the immediate effect predicted in the two and four week treatment specifications is much smaller.

Note that results from specification (1.4) predict an average increase in sales from increased variety due to the prevalence of sufficiently large variety increases in the data. At the mean percent positive increase (0.473), the estimates predict an increase in sales for both samples, however, the 10th percentile of the non-zero observations for "(% Positive

Table 1.15: Log Total Quantity on Percent Change in Choices, Sample 1

	(1)	(2)	(3)	(4)
Treatment Length (weeks)	2	4	6	8
	b/se	b/se	b/se	b/se
Reset	-0.004	0.007	0.018***	0.030***
	(0.006)	(0.006)	(0.005)	(0.005)
(% Positive Change) * Reset	0.179***	0.191***	0.213***	0.232***
	(0.016)	(0.012)	(0.011)	(0.011)
(Positive Group) * Reset	-0.034*	-0.008	-0.053***	-0.103***
	(0.014)	(0.011)	(0.010)	(0.010)
% Positive Change	-0.145***	-0.170***	-0.176***	-0.191***
	(0.015)	(0.013)	(0.013)	(0.012)
Positive Group	0.046***	-0.047***	-0.058***	-0.042***
	(0.008)	(0.007)	(0.007)	(0.007)
(% Negative Change) * Reset	0.429***	0.440***	0.473***	0.379***
	(0.067)	(0.062)	(0.059)	(0.055)
(Negative Group) * Reset	-0.021	-0.073***	-0.064***	-0.095***
	(0.012)	(0.009)	(0.008)	(0.008)
% Negative Change	0.224***	0.071**	-0.011	-0.021
	(0.032)	(0.025)	(0.024)	(0.024)
Negative Group	0.035***	-0.032***	-0.051***	-0.041***
	(0.010)	(0.008)	(0.008)	(0.008)
% Change in Store AGs	-0.000	-0.000	-0.000	-0.000
	(0.000)	(0.000)	(0.000)	(0.000)
(% Change in Store AGs)*Reset	-0.001	-0.001	-0.001	-0.001
	(0.001)	(0.001)	(0.001)	(0.001)
Facings wide, % of total ice cream	-0.011	-0.003	0.001	0.006
	(0.020)	(0.019)	(0.018)	(0.017)
Facings deep, % of total ice cream	0.157**	0.154**	0.153***	0.153***
	(0.050)	(0.048)	(0.046)	(0.044)
Facings high, % of total ice cream	0.156***	0.152***	0.151***	0.150***
	(0.017)	(0.016)	(0.015)	(0.014)
Avg. temperature (deg F, 00s)	0.166***	0.157***	0.158***	0.168***
	(0.027)	(0.027)	(0.025)	(0.025)
No. of UPCs in store (00s)	0.095***	0.096***	0.096***	0.097***
	(0.018)	(0.017)	(0.015)	(0.015)
Constant	2.347***	2.340***	2.325***	2.303***
	(0.078)	(0.074)	(0.071)	(0.068)
Adjusted R^2	0.763	0.763	0.762	0.762
Observations	1,595,917	1,700,658	1,804,712	1,908,719
Store FEs	Y	Y	Y	Y
Two-week FEs	Y	Y	Y	Y
AG FEs	Y	Y	Y	Y

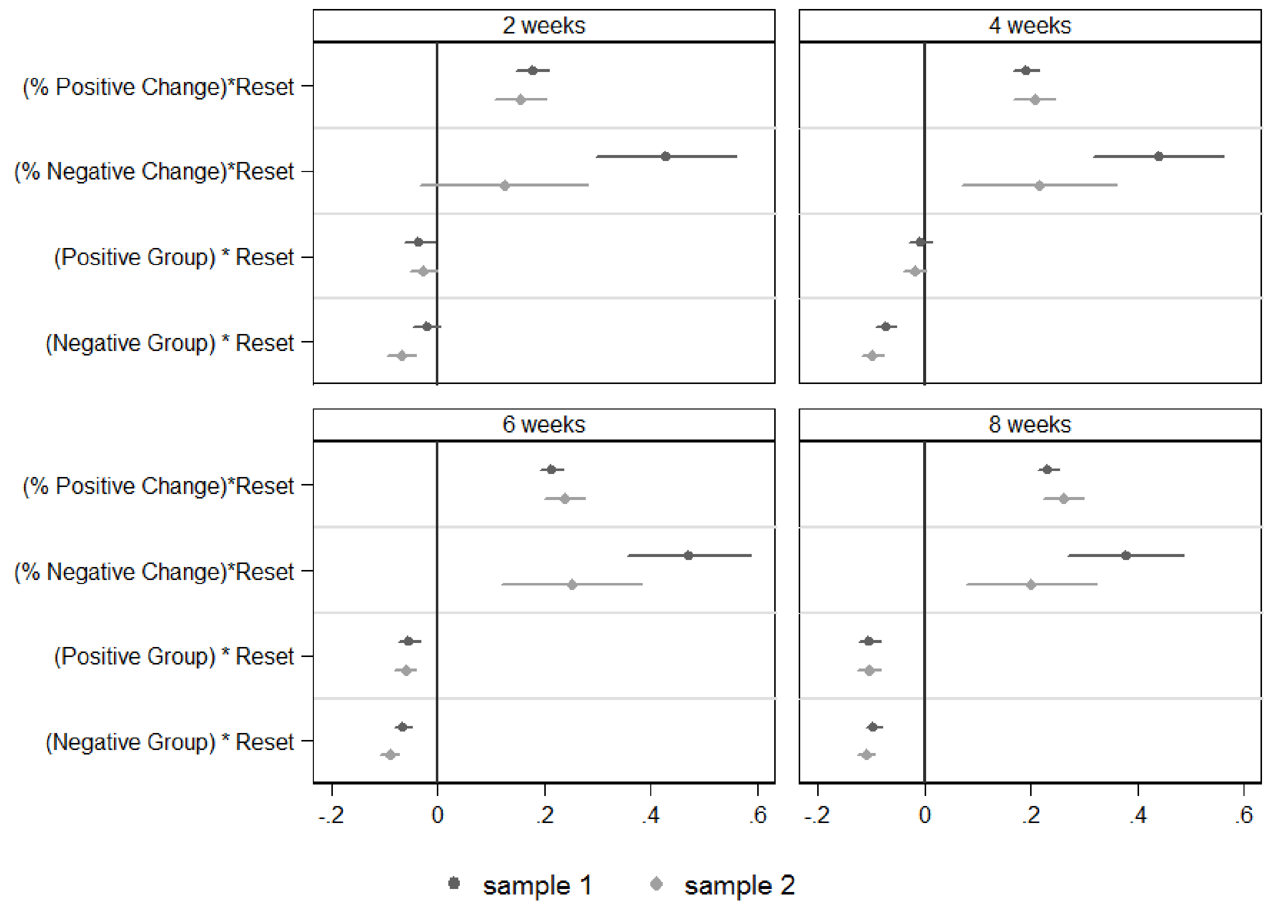
Standard errors in parentheses. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 1.16: Log Total Quantity on Percent Change in Choices, Sample 2

	(1)	(2)	(3)	(4)
Treatment Length (weeks)	2	4	6	8
	b/se	b/se	b/se	b/se
Reset	-0.006	0.011	0.025***	0.040***
	(0.006)	(0.006)	(0.006)	(0.006)
(% Positive Change) * Reset	0.155***	0.207***	0.238***	0.262***
	(0.025)	(0.020)	(0.020)	(0.019)
(Positive Group) * Reset	-0.027*	-0.018	-0.061***	-0.105***
	(0.013)	(0.011)	(0.011)	(0.011)
% Positive Change	-0.171***	-0.215***	-0.223***	-0.242***
	(0.026)	(0.022)	(0.021)	(0.020)
Positive Group	0.039***	-0.032***	-0.039***	-0.022*
	(0.009)	(0.008)	(0.008)	(0.008)
(% Negative Change) * Reset	0.125	0.215**	0.252***	0.200**
	(0.081)	(0.073)	(0.068)	(0.063)
(Negative Group) * Reset	-0.068***	-0.098***	-0.090***	-0.111***
	(0.013)	(0.010)	(0.009)	(0.008)
% Negative Change	0.220***	0.089**	0.026	0.015
	(0.038)	(0.030)	(0.028)	(0.027)
Negative Group	0.020*	-0.036***	-0.046***	-0.036***
	(0.010)	(0.008)	(0.008)	(0.008)
% Change in Store AGs	0.000	0.000	0.000	-0.000
	(0.000)	(0.000)	(0.000)	(0.000)
(% Change in Store AGs)*Reset	-0.001	-0.001	-0.001	-0.001
	(0.001)	(0.001)	(0.001)	(0.001)
Facings wide, % of total ice cream	0.039*	0.044**	0.047**	0.051***
	(0.017)	(0.016)	(0.016)	(0.015)
Facings deep, % of total ice cream	0.083	0.083*	0.084*	0.086*
	(0.043)	(0.041)	(0.040)	(0.039)
Facings high, % of total ice cream	0.146***	0.142***	0.140***	0.138***
	(0.014)	(0.013)	(0.012)	(0.012)
Avg. temperature (deg F, 00s)	0.169***	0.158***	0.154***	0.154***
	(0.029)	(0.027)	(0.026)	(0.026)
No. of UPCs in store (00s)	0.143***	0.141***	0.139***	0.138***
	(0.024)	(0.022)	(0.020)	(0.019)
Constant	2.086***	2.093***	2.092***	2.083***
	(0.095)	(0.090)	(0.085)	(0.081)
Adjusted R^2	0.774	0.775	0.775	0.774
Observations	900,426	958,568	1,016,593	1,074,631
Store FEs	Y	Y	Y	Y
Two-week FEs	Y	Y	Y	Y
AG FEs	Y	Y	Y	Y

Standard errors in parentheses. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Figure 1.3: Estimates of change in variety variables



Point estimates and 95% confidence intervals (not exponentiated) of change in variety variables from specification (1.6). Subcharts correspond to a specific treatment period length and show estimates for both samples. All estimates for "(Positive Group)*Reset" and "(Negative Group)*Reset" are negative, all estimates of "(% Positive Change)*Reset" and "(% Negative Change)* Reset" are positive, and nearly all estimates are significant. For all treatment lengths, "(% Negative Change)* Reset" estimate is significantly greater than that of "(% Positive Change)*Reset", implying an asymmetric effect in increases versus decreases in variety. In Sample 2, slope effects are not significantly different.

Table 1.17: Estimated Percent Change in Variety (00s) from Tables (1.15) and (1.16)

	Sample 1				Sample 2			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Treatment Length (weeks)	2	4	6	8	2	4	6	8
Reset	-0.004	0.007	0.018***	0.031***	-0.006	0.011	0.025***	0.041***
(% Positive Change) * Reset	0.196***	0.210***	0.238***	0.261***	0.168***	0.229***	0.269***	0.299***
(Positive Group) * Reset	-0.034*	-0.009	-0.052***	-0.098***	-0.027*	-0.018	-0.059***	-0.100***
% Positive Change	-0.135***	-0.156***	-0.162***	-0.174***	-0.157***	-0.194***	-0.200***	-0.215***
Positive Group	0.047***	-0.046***	-0.056***	-0.041***	0.040***	-0.032***	-0.038***	-0.021*
(% Negative Change) * Reset	0.532***	0.550***	0.601***	0.458***	0.129	0.237**	0.284***	0.219**
(Negative Group) * Reset	-0.021	-0.070***	-0.062***	-0.091***	-0.066***	-0.093***	-0.086***	-0.105***
% Negative Change	0.251***	0.074**	-0.012	-0.021	0.245***	0.093**	0.025	0.015
Negative Group	0.036***	-0.032***	-0.050***	-0.040***	0.020*	-0.035***	-0.045***	-0.036***

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Estimated coefficients from Tables (1.15) and (1.16) exponentiated and corrected for bias (Kennedy, 1983)

Change)*Reset" lies at 0.059, where our estimates would predict a negative or near zero change in sales for all definitions of the treatment period. Thus, the value range that predicts a decrease in sales from small percentage increases in variety is indeed in the feasible range. However, the prevalence of larger increases in variety from the resets accounts for the average increase in sales from increased variety.

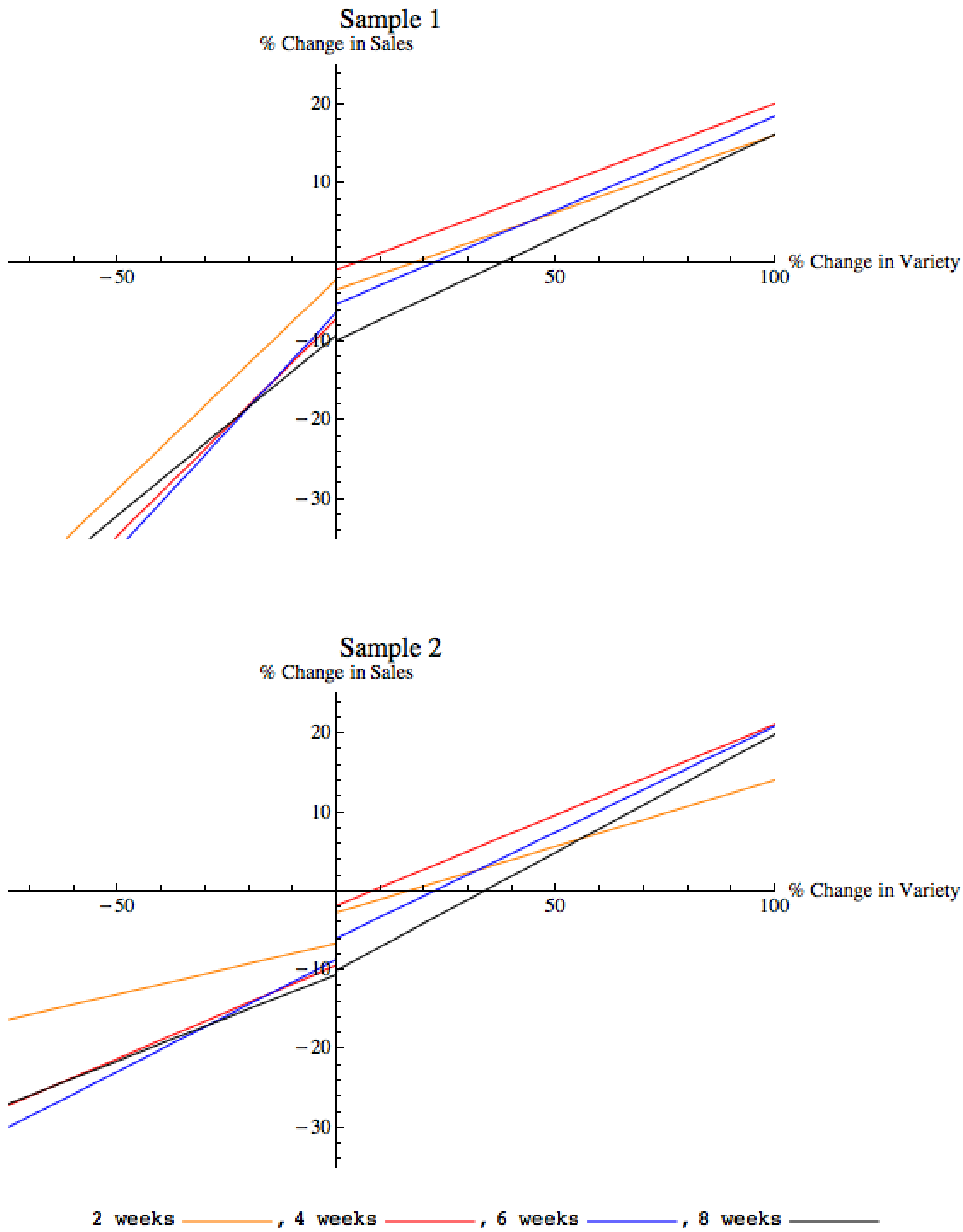
We illustrate the above features in Figure (1.4), including plots of the marginal predicted percent changes in sales from (non-zero) changes in variety for samples 1 and 2, using the predicted percent changes in Table (1.17). Each chart includes a separate line for the four different treatment lengths used. Note the graphs do not plot the effect of resets, but plot effects from changes in variety as identified by the difference in differences methodology. Thus, the intercepts derive from the coefficients on the "(Positive Group)*Reset" and "(Negative Group)*Reset" variables, while the slopes derive from the "(% Positive Change)*Reset" and "(% Negative Change)*Reset" estimates. In sample 1, the estimated effects visibly kink at the vertical axis, reflecting the asymmetric effect of positive versus negative percent changes. In addition, for both samples, the estimated percent change in sales is negative as the percent change in sales reduces to zero.

Results from the lagged specification are provided in Tables (1.18) and (1.19), with corresponding percent changes in Table (1.20). We focus on column (3) in each table, containing the estimates from the fully lagged model, and plot change of variety estimates and 95% confidence intervals in Figure (1.5). Consistent with the non-lagged results, all point estimates for the lags on "(Positive Group)*Reset" and "(Negative Group)*Reset" are negative, and almost entirely significant, in both samples. The magnitude of the estimates for these variables is always smallest in lag 0, and largest in the four-week lag for "(Positive Group)*Reset" and in the six-week lag for "(Negative Group)*Reset". Thus, sales tend to respond more negatively to small magnitude changes in variety six to eight weeks after a reset.

The pattern in the effects from the magnitude of a positive percent change are similar between the lagged and non-lagged specifications. The estimates on all lags for "(% Positive Change)*Reset" are significantly positive for both samples. Thus, larger percent increases will boost sales more than smaller percent increases, for each two-week period in the eight weeks after a reset, although there is no monotonic trend in the size of the effect over time. Turning to the "(% Negative Change)*Reset" variables, while the results in Table (1.15) and (1.16) demonstrate that the average effect from an incremental decrease in the magnitude of a variety reduction was to decrease sales, the lagged specification demonstrates that this pattern holds for the initial weeks after reset, but in later weeks (by lag 6 weeks for sample 1 and lag 2 weeks for sample 2), larger decreases in variety actually decrease sales less.

Figures (1.6) and (1.7) demonstrate these phenomenon for samples 1 and 2 respectively. These figures show the predicted marginal effects of non-zero changes in variety at the various lags, for the 10th, 50th, and 90th percentiles of the "(% Positive Change)*Reset" and "(% Negative Change)*Reset" variables. Lines are drawn between predicted values solely to identify a given percentile at various lags, with the values corresponding to the percentiles provided on either side of each line. Increases in variety are shown in black while decreases

Figure 1.4: Marginal Predicted Effects of Change in Variety on Sales



Plots derived from results of specification (1.6). The top plot shows results from tests on sample 1, the representative sample, while lower plot shows results from sample 2, the ice cream staples (see text). Lines are point estimates of the marginal predicted effect of percent changes in variety on sales, for the four definitions of treatment length: reset in the past 2 weeks, 4 weeks, 6 weeks, or 8 weeks. Axes rescaled as percentage changes.

Table 1.18: Log Total Quantity on Percent Change in Choices with Lags, Sample 1

	(1)		(2)		(3)	
	b	se	b	se	b	se
Reset	-0.005	(0.005)	-0.005	(0.005)	-0.005	(0.005)
Reset, lag 2 weeks	0.003	(0.007)	0.003	(0.007)	0.003	(0.007)
Reset, lag 4 weeks			-0.001	(0.010)	-0.001	(0.010)
Reset, lag 6 weeks					0.008	(0.012)
(% Positive Change) * Reset	0.179***	(0.016)	0.179***	(0.016)	0.180***	(0.016)
(% Positive Change) * Reset, lag 2 weeks	0.141***	(0.015)	0.141***	(0.015)	0.141***	(0.015)
(% Positive Change) * Reset, lag 4 weeks			0.170***	(0.016)	0.171***	(0.016)
(% Positive Change) * Reset, lag 6 weeks					0.131***	(0.016)
(Positive Group) * Reset	-0.036**	(0.013)	-0.038**	(0.013)	-0.040**	(0.013)
(Positive Group) * Reset, lag 2 weeks	-0.081***	(0.015)	-0.083***	(0.015)	-0.085***	(0.015)
(Positive Group) * Reset, lag 4 weeks			-0.135***	(0.015)	-0.136***	(0.015)
(Positive Group) * Reset, lag 6 weeks					-0.132***	(0.015)
% Positive Change	-0.153***	(0.015)	-0.157***	(0.015)	-0.162***	(0.015)
Positive Group	0.050***	(0.008)	0.054***	(0.008)	0.061***	(0.008)
(% Negative Change) * Reset	0.421***	(0.065)	0.414***	(0.064)	0.409***	(0.062)
(% Negative Change) * Reset, lag 2 weeks	0.186**	(0.061)	0.180**	(0.060)	0.176**	(0.059)
(% Negative Change) * Reset, lag 4 weeks			0.202***	(0.054)	0.198***	(0.053)
(% Negative Change) * Reset, lag 6 weeks					-0.062	(0.054)
(Negative Group) * Reset	-0.022	(0.012)	-0.022	(0.012)	-0.023	(0.012)
(Negative Group) * Reset, lag 2 weeks	-0.103***	(0.011)	-0.103***	(0.011)	-0.103***	(0.011)
(Negative Group) * Reset, lag 4 weeks			-0.066***	(0.011)	-0.067***	(0.011)
(Negative Group) * Reset, lag 6 weeks					-0.194***	(0.011)
% Negative Change	0.215***	(0.031)	0.217***	(0.031)	0.222***	(0.031)
Negative Group	0.040***	(0.010)	0.047***	(0.010)	0.053***	(0.009)
% Change in Store AGs	-0.000	(0.000)	-0.000	(0.000)	-0.000	(0.000)
(% Change in Store AGs)*Reset	-0.001	(0.001)	-0.001	(0.001)	-0.001	(0.001)
Facings wide, % of total ice cream	-0.006	(0.019)	-0.003	(0.018)	0.001	(0.017)
Facings deep, % of total ice cream	0.159***	(0.048)	0.162***	(0.047)	0.163***	(0.045)
Facings high, % of total ice cream	0.154***	(0.016)	0.153***	(0.015)	0.152***	(0.015)
Avg. temperature (deg F, 00s)	0.156***	(0.027)	0.157***	(0.026)	0.167***	(0.025)
No. of UPCs in store (00s)	0.098***	(0.016)	0.100***	(0.015)	0.100***	(0.015)
Constant	2.327***	(0.074)	2.307***	(0.071)	2.287***	(0.068)
Adjusted R^2	0.762		0.761		0.761	
Observations	1,700,658		1,804,712		1,908,719	
Store FEs	Y		Y		Y	
Two-week FEs	Y		Y		Y	
AG FEs	Y		Y		Y	

Standard errors in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 1.19: Log Total Quantity on Percent Change in Choices with Lags, Sample 2

	(1)		(2)		(3)	
	b	se	b	se	b	se
Reset	-0.008	(0.006)	-0.008	(0.005)	-0.007	(0.005)
Reset, lag 2 weeks	-0.000	(0.008)	-0.000	(0.008)	0.000	(0.008)
Reset, lag 4 weeks			-0.006	(0.010)	-0.006	(0.010)
Reset, lag 6 weeks					0.005	(0.013)
(% Positive Change) * Reset	0.158***	(0.025)	0.160***	(0.025)	0.162***	(0.025)
(% Positive Change) * Reset, lag 2 weeks	0.128***	(0.026)	0.129***	(0.026)	0.131***	(0.026)
(% Positive Change) * Reset, lag 4 weeks			0.207***	(0.032)	0.210***	(0.032)
(% Positive Change) * Reset, lag 6 weeks					0.212***	(0.030)
(Positive Group) * Reset	-0.029*	(0.013)	-0.031*	(0.012)	-0.033**	(0.012)
(Positive Group) * Reset, lag 2 weeks	-0.057***	(0.015)	-0.060***	(0.015)	-0.062***	(0.014)
(Positive Group) * Reset, lag 4 weeks			-0.117***	(0.016)	-0.118***	(0.016)
(Positive Group) * Reset, lag 6 weeks					-0.114***	(0.016)
% Positive Change	-0.178***	(0.026)	-0.184***	(0.027)	-0.193***	(0.027)
Positive Group	0.042***	(0.009)	0.044***	(0.009)	0.050***	(0.009)
(% Negative Change) * Reset	0.111	(0.078)	0.099	(0.076)	0.090	(0.075)
(% Negative Change) * Reset, lag 2 weeks	-0.076	(0.071)	-0.089	(0.069)	-0.097	(0.068)
(% Negative Change) * Reset, lag 4 weeks			-0.163*	(0.064)	-0.170**	(0.063)
(% Negative Change) * Reset, lag 6 weeks					-0.345***	(0.062)
(Negative Group) * Reset	-0.070***	(0.013)	-0.072***	(0.013)	-0.073***	(0.013)
(Negative Group) * Reset, lag 2 weeks	-0.138***	(0.012)	-0.139***	(0.012)	-0.141***	(0.012)
(Negative Group) * Reset, lag 4 weeks			-0.104***	(0.011)	-0.105***	(0.011)
(Negative Group) * Reset, lag 6 weeks					-0.219***	(0.011)
% Negative Change	0.220***	(0.037)	0.229***	(0.037)	0.245***	(0.037)
Negative Group	0.025**	(0.010)	0.033***	(0.010)	0.040***	(0.010)
% Change in Store AGs	0.000	(0.000)	0.000	(0.000)	-0.000	(0.000)
(% Change in Store AGs)*Reset	-0.001	(0.001)	-0.001	(0.001)	-0.001	(0.001)
Facings wide, % of total ice cream	0.042*	(0.017)	0.044**	(0.016)	0.047**	(0.016)
Facings deep, % of total ice cream	0.087*	(0.042)	0.091*	(0.041)	0.093*	(0.040)
Facings high, % of total ice cream	0.144***	(0.013)	0.142***	(0.013)	0.140***	(0.012)
Avg. temperature (deg F, 00s)	0.157***	(0.027)	0.152***	(0.026)	0.152***	(0.026)
No. of UPCs in store (00s)	0.144***	(0.022)	0.144***	(0.020)	0.143***	(0.019)
Constant	2.078***	(0.090)	2.070***	(0.085)	2.062***	(0.082)
Adjusted R^2	0.774		0.773		0.773	
Observations	958,568		1,016,593		1,074,631	
Store FEs	Y		Y		Y	
Two-week FEs	Y		Y		Y	
AG FEs	Y		Y		Y	

Standard errors in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 1.20: Estimated Percent Change in Variety (00s) from Tables (1.18) and (1.19)

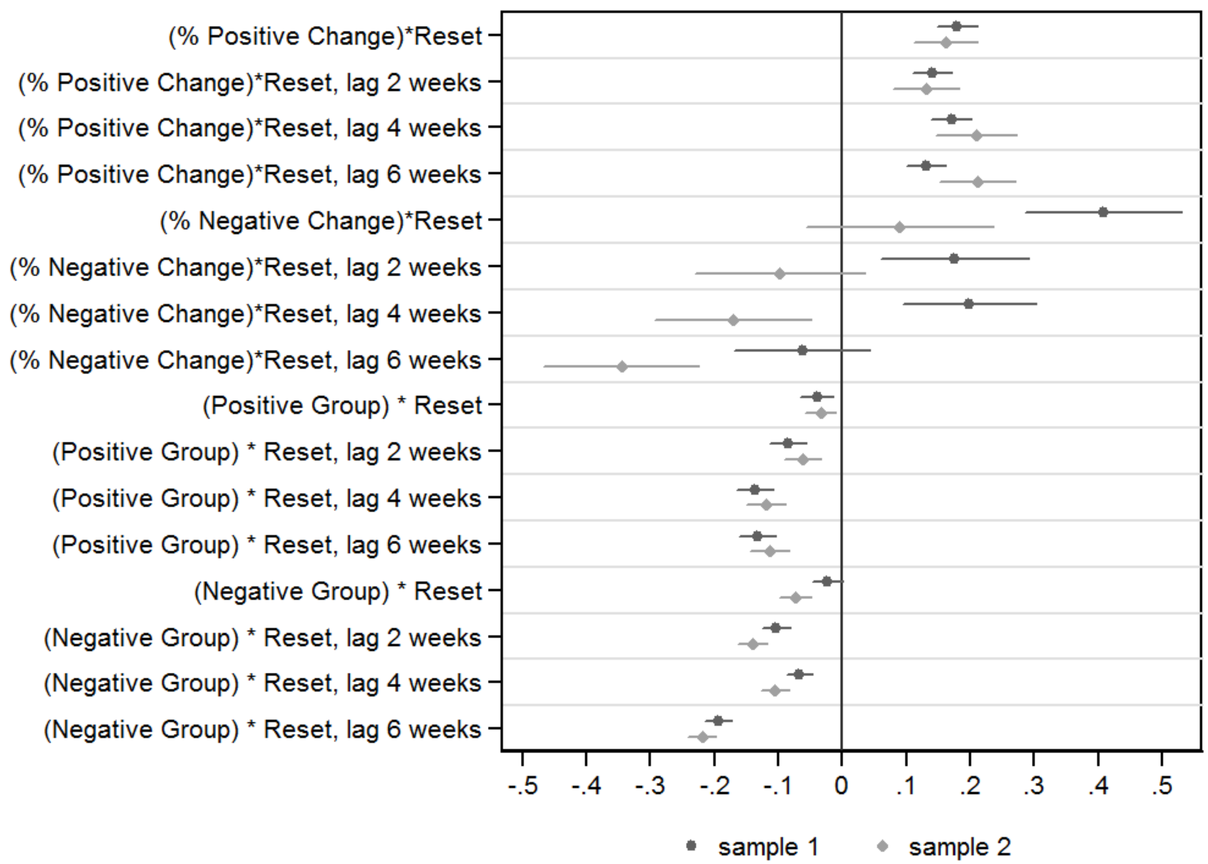
	Sample 1	Sample 2
Reset	-0.005	-0.007
Reset, lag 2 weeks	0.003	0.000
Reset, lag 4 weeks	-0.001	-0.006
Reset, lag 6 weeks	0.008	0.005
(% Positive Change) * Reset	0.197***	0.175***
(% Positive Change) * Reset, lag 2 weeks	0.152***	0.140***
(% Positive Change) * Reset, lag 4 weeks	0.186***	0.232***
(% Positive Change) * Reset, lag 6 weeks	0.139***	0.235***
(Positive Group) * Reset	-0.039**	-0.033**
(Positive Group) * Reset, lag 2 weeks	-0.082***	-0.060***
(Positive Group) * Reset, lag 4 weeks	-0.128***	-0.112***
(Positive Group) * Reset, lag 6 weeks	-0.124***	-0.108***
% Positive Change	-0.149***	-0.176***
Positive Group	0.062***	0.052***
(% Negative Change) * Reset	0.502***	0.091
(% Negative Change) * Reset, lag 2 weeks	0.190**	-0.095
(% Negative Change) * Reset, lag 4 weeks	0.218***	-0.158**
(% Negative Change) * Reset, lag 6 weeks	-0.061	-0.293***
(Negative Group) * Reset	-0.023	-0.071***
(Negative Group) * Reset, lag 2 weeks	-0.098***	-0.131***
(Negative Group) * Reset, lag 4 weeks	-0.064***	-0.100***
(Negative Group) * Reset, lag 6 weeks	-0.176***	-0.197***
% Negative Change	0.248***	0.277***
Negative Group	0.054***	0.041***

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Estimated from Tables (1.18) and (1.19).

Exponentiated and corrected for bias (Kennedy, 1983)

Figure 1.5: Estimates of Change in Variety Variables with Lags



Point estimates and 95% confidence intervals (not exponentiated) of the change in variety variables from specification (1.7) are plotted for each definition of the treatment period. The six-week lags of the "(Positive Group)*Reset" and "(Negative Group)*Reset" variables are significantly more negative than the lag 0 versions, implying that small changes in variety reduce sales more over time. Estimates for "(% Negative Change)*Reset" change from positive in early weeks to negative in later weeks, implying larger reductions in variety reduce sales less than small changes in variety in these later weeks.

in variety are shown in red, with the line for the 50th percentile of each variable in bold.

In Sample 1, Figure (1.6) demonstrates a downward trend in the effect of increases in variety on the change in sales over time. The 10th, 50th, and 90th percentile increases in variety respectively correspond to a 6%, 33%, and 100% increase in the number of UPCs in an AG. The large increase in variety continues to boost sales for the next six weeks, but less so over time. For the median increase at 33%, sales actually decrease two weeks after a reset, and continue to decrease for the next four weeks. For the small increase in variety at 6%, sales decrease in all periods after the change. While less drastic, the downward patterns for the 33% and 6% increases in variety are similar in sample 2. For the large increase at 100%, the initial drop in the change in sales after lag 0 tends to level out or reverse by lag 6. In both samples, the moderate and large increases in variety (33% and 100%) lead to increased sales immediately, but the 33% increase in variety decreases sales by lags 2 and onward, and the small increase in variety reduces sales for all periods.

For variety decreases in sample 1, Figure (1.6) demonstrates that sales reduce in response to changes at the 10th, 50th, and 90t percentile levels (-5%, -25%, and -50% changes in variety, respectively) for all periods. For lags of 0-4 weeks, larger sales reductions correspond to larger decreases in variety. However, the effect of magnitude is not significant for the lag of 6 weeks, and the point estimate is actually negative. In sample 2, the point estimate for lag 0 of "(% Negative Change)*Reset" is significantly greater than zero, but lags 4 and 6 are significant and negative.

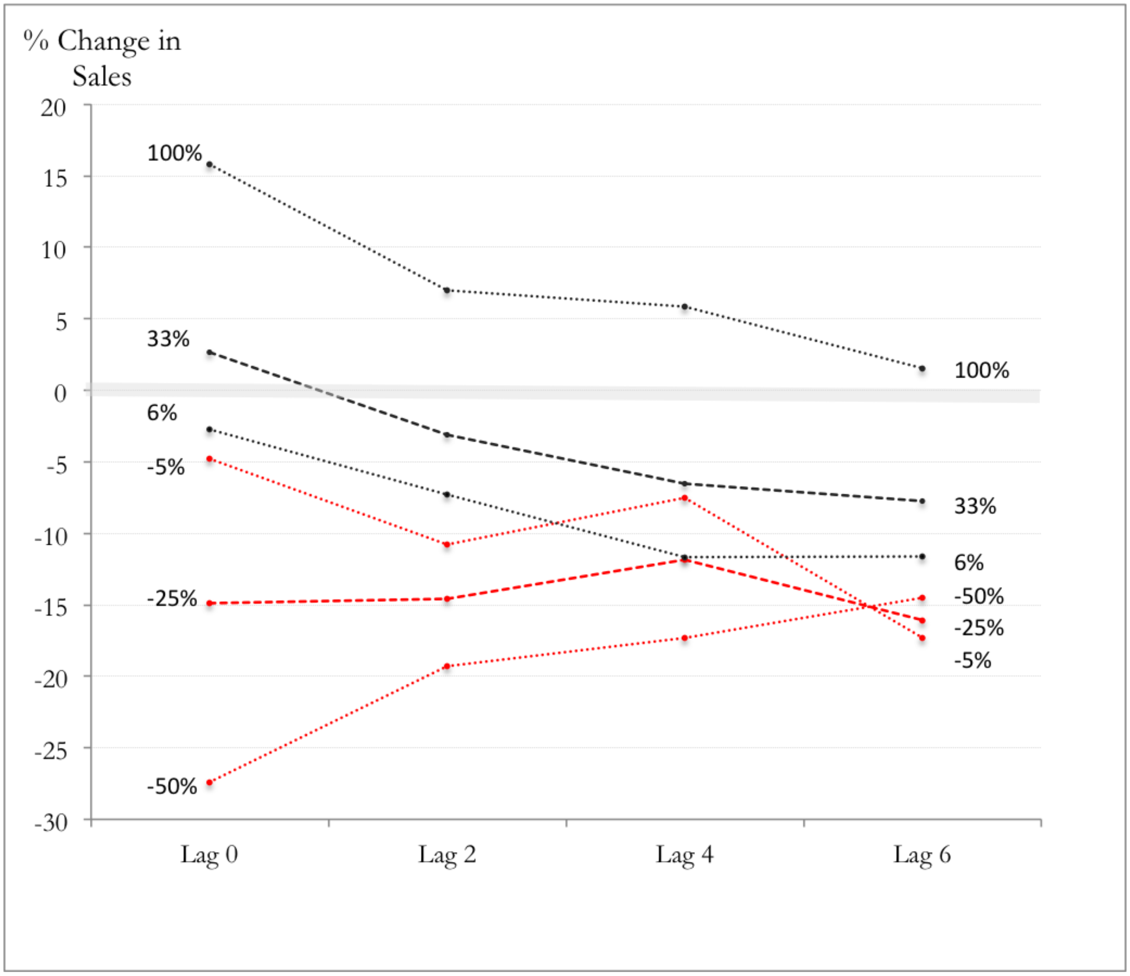
As illustrated by Figure (1.7), larger decreases in variety reduce sales more initially, but in later periods, larger decreases in variety still reduce sales, but less so compared to smaller decreases in variety. In sample 2, sales for product groups with large decreases in variety recover to the extent that the percentage sales effects fair better than for product groups with small increases in variety. The alleviation in sales for larger reductions may again reflect substitution over time if consumers tend to buy the same products before and after changes in variety. Alternatively, consumers may switch to new products over time, as they adjust to the removal of a preferred product, or, similar to the findings in Sloot, Fok, and Verhoef (2006), the entry of new consumers to the ice cream category may be compensating for initial losses in sales.

1.4.3 Results of Robustness Tests

Tables (1.21)-(1.24) present the results of specifications (1.8) and (1.9). As described in section (1.2), store prices are set simultaneously each week, with uniform pricing for stores in the same geographic division. Thus, including an interaction term consisting of division fixed effects and period fixed effects should capture price changes. In place of store fixed effects, the robustness tests use division fixed effects and division-period interactions to capture changes in price. Again, we plot the estimates of the change in variety variables in Figures (1.8) and (1.9).

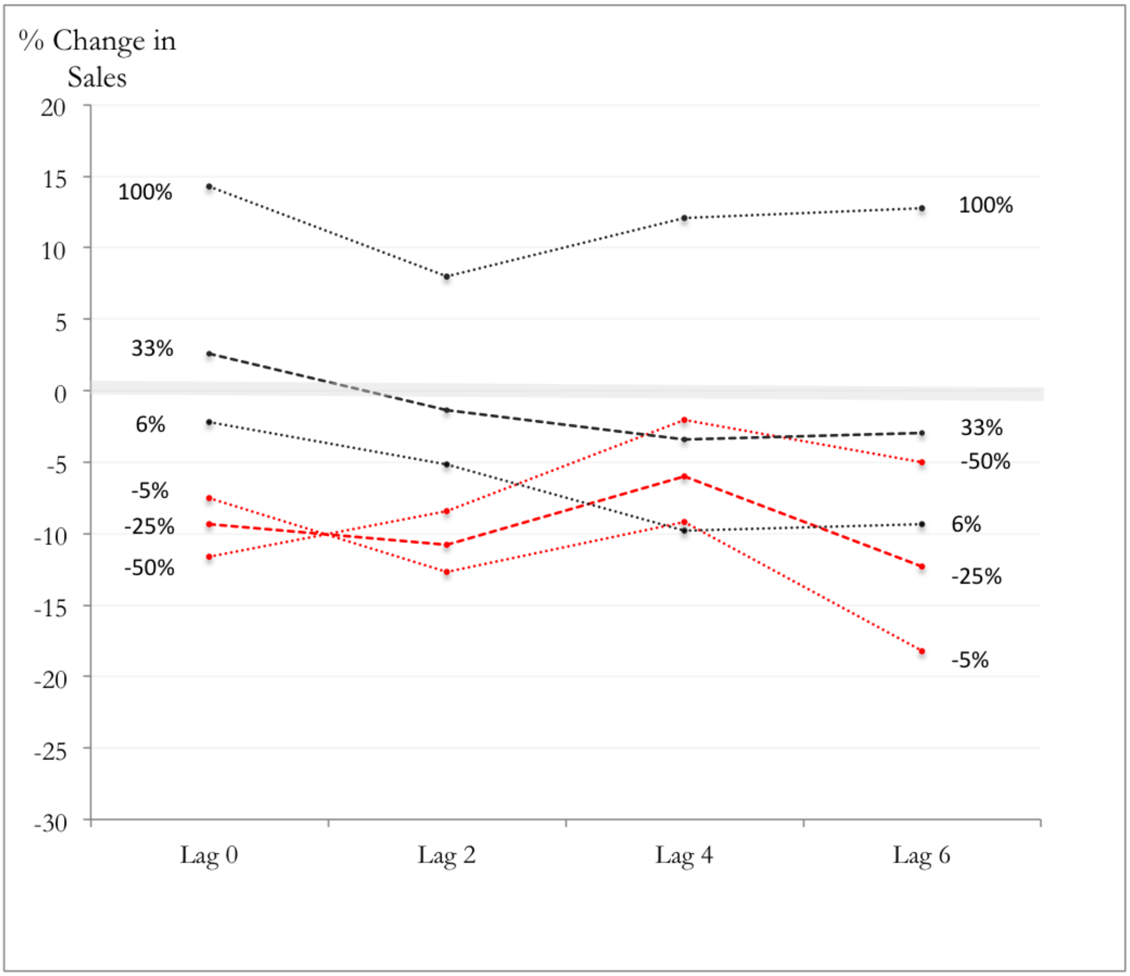
The estimates on the change in variety variables from the robustness specification are very similar to those from the analogous specification with store fixed effects, as shown by

Figure 1.6: Marginal Treatment Effects of Change in Variety Variables with Lags, sample 1



Predicted marginal effects for six levels of percent change in variety are shown for sample 1. Axes rescaled as percentage changes. Each lag value corresponds to the number of weeks since a reset. Levels are the 10th, 50th, and 90th percentiles of the nonzero increases and decreases in variety in our dataset, with the corresponding values of each percentage change shown on either side of the line. Large increases in variety (100%) increase sales, but less so over time. Moderate and small increases in variety reduce sales in the later periods. Decreases in variety reduce sales, but larger decreases in variety reduce sales less by lag 6.

Figure 1.7: Marginal Treatment Effects of Change in Variety Variables with Lags, sample 2



Predicted marginal effects for six levels of percent change in variety are shown for sample 2. Axes rescaled as percentage changes. Each lag value corresponds to the number of weeks since a reset. Levels are the 10th, 50th, and 90th percentiles of the nonzero increases and decreases in variety in our dataset, with the corresponding values of each percentage change shown on either side of the line. Large increases in variety (100%) increase sales. Moderate and small increases in variety reduce sales in the later periods. Decreases in variety reduce sales, but larger decreases in variety reduce sales less by lag 2, and reduce percent sales less than small increases in variety in lags 4 and 6.

Table 1.21: Robustness Test of PCC, Sample 1

	(1)	(2)	(3)	(4)
Treatment Length (weeks)	2	4	6	8
	b/se	b/se	b/se	b/se
Reset	0.031*	0.053**	0.056***	0.065***
	(0.015)	(0.016)	(0.016)	(0.015)
(% Positive Change) * Reset	0.178***	0.179***	0.205***	0.229***
	(0.016)	(0.013)	(0.012)	(0.011)
(Positive Group) * Reset	-0.037*	-0.005	-0.050***	-0.106***
	(0.014)	(0.012)	(0.010)	(0.010)
% Positive Change	-0.181***	-0.177***	-0.176***	-0.191***
	(0.019)	(0.016)	(0.015)	(0.015)
Positive Group	0.037***	-0.083***	-0.107***	-0.094***
	(0.010)	(0.009)	(0.009)	(0.009)
(% Negative Change) * Reset	0.387***	0.435***	0.481***	0.369***
	(0.072)	(0.066)	(0.061)	(0.058)
(Negative Group) * Reset	-0.018	-0.069***	-0.056***	-0.094***
	(0.013)	(0.010)	(0.009)	(0.008)
% Negative Change	0.261***	0.068	-0.064	-0.100**
	(0.040)	(0.036)	(0.035)	(0.033)
Negative Group	0.006	-0.078***	-0.113***	-0.110***
	(0.012)	(0.010)	(0.010)	(0.010)
% Change in Store AGs	-0.000	-0.000	-0.000	-0.000
	(0.000)	(0.000)	(0.000)	(0.000)
(% Change in Store AGs)*Reset	0.004	0.002	-0.001	-0.003*
	(0.002)	(0.001)	(0.001)	(0.001)
Facings wide, % of total ice cream	-0.008	-0.001	0.004	0.008
	(0.020)	(0.019)	(0.018)	(0.017)
Facings deep, % of total ice cream	0.158**	0.156**	0.156***	0.157***
	(0.050)	(0.048)	(0.046)	(0.045)
Facings high, % of total ice cream	0.157***	0.153***	0.152***	0.151***
	(0.018)	(0.016)	(0.016)	(0.015)
Avg. temperature (deg F, 00s)	0.436**	0.430**	0.421**	0.402**
	(0.146)	(0.146)	(0.147)	(0.148)
No. of UPCs in store (00s)	0.321***	0.321***	0.322***	0.323***
	(0.015)	(0.015)	(0.015)	(0.015)
Constant	1.096***	1.138***	1.143***	1.143***
	(0.129)	(0.128)	(0.128)	(0.128)
Adjusted R^2	0.714	0.714	0.714	0.713
Observations	1,595,917	1,700,658	1,804,712	1,908,719
Division FEs	Y	Y	Y	Y
Two-week FEs	Y	Y	Y	Y
(Division FEs)* (Two-week FEs)	Y	Y	Y	Y
AG FEs	Y	Y	Y	Y

Standard errors in parentheses. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 1.22: Robustness Test of PCC, Sample 2

	(1)	(2)	(3)	(4)
Treatment Length (weeks)	2	4	6	8
	b/se	b/se	b/se	b/se
Reset	0.023	0.053**	0.060***	0.072***
	(0.015)	(0.016)	(0.016)	(0.016)
(% Positive Change) * Reset	0.145***	0.189***	0.224***	0.254***
	(0.026)	(0.021)	(0.021)	(0.020)
(Positive Group) * Reset	-0.026	-0.013	-0.058***	-0.109***
	(0.014)	(0.012)	(0.011)	(0.011)
% Positive Change	-0.221***	-0.233***	-0.230***	-0.249***
	(0.033)	(0.026)	(0.025)	(0.024)
Positive Group	0.035**	-0.060***	-0.079***	-0.063***
	(0.012)	(0.010)	(0.010)	(0.011)
(% Negative Change) * Reset	0.080	0.203**	0.261***	0.195**
	(0.088)	(0.078)	(0.071)	(0.066)
(Negative Group) * Reset	-0.067***	-0.096***	-0.084***	-0.111***
	(0.014)	(0.011)	(0.009)	(0.009)
% Negative Change	0.261***	0.098*	-0.014	-0.051
	(0.048)	(0.043)	(0.041)	(0.040)
Negative Group	-0.010	-0.080***	-0.104***	-0.099***
	(0.012)	(0.010)	(0.010)	(0.010)
% Change in Store AGs	-0.000	-0.000	-0.000	-0.000
	(0.000)	(0.000)	(0.000)	(0.000)
(% Change in Store AGs)*Reset	0.003	0.001	-0.001	-0.003*
	(0.002)	(0.001)	(0.001)	(0.001)
Facings wide, % of total ice cream	0.042*	0.046**	0.049**	0.053***
	(0.017)	(0.016)	(0.016)	(0.015)
Facings deep, % of total ice cream	0.082	0.083*	0.085*	0.087*
	(0.042)	(0.041)	(0.040)	(0.039)
Facings high, % of total ice cream	0.146***	0.141***	0.139***	0.138***
	(0.014)	(0.013)	(0.013)	(0.012)
Avg. temperature (deg F, 00s)	0.419**	0.414**	0.408**	0.391*
	(0.152)	(0.152)	(0.154)	(0.155)
No. of UPCs in store (00s)	0.371***	0.368***	0.369***	0.370***
	(0.014)	(0.014)	(0.014)	(0.014)
Constant	1.418***	1.477***	1.481***	1.476***
	(0.134)	(0.133)	(0.132)	(0.132)
Adjusted R^2	0.714	0.715	0.715	0.713
Observations	900,426	958,568	1,016,593	1,074,631
Division FEs	Y	Y	Y	Y
Two-week FEs	Y	Y	Y	Y
(Division FEs)* (Two-week FEs)	Y	Y	Y	Y
AG FEs	Y	Y	Y	Y

Standard errors in parentheses. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 1.23: Robustness Test of PCC with Lags, Sample 1

	(1)		(2)		(3)	
	b	se	b	se	b	se
Reset	0.026	(0.014)	0.026	(0.014)	0.025	(0.014)
Reset, lag 2 weeks	0.072**	(0.024)	0.072**	(0.024)	0.072**	(0.024)
Reset, lag 4 weeks			0.115**	(0.035)	0.114**	(0.036)
Reset, lag 6 weeks					0.171***	(0.046)
(% Positive Change) * Reset	0.177***	(0.016)	0.177***	(0.016)	0.178***	(0.016)
(% Positive Change) * Reset, lag 2 weeks	0.138***	(0.016)	0.138***	(0.016)	0.138***	(0.016)
(% Positive Change) * Reset, lag 4 weeks			0.166***	(0.017)	0.166***	(0.017)
(% Positive Change) * Reset, lag 6 weeks					0.127***	(0.016)
(Positive Group) * Reset	-0.040**	(0.014)	-0.042**	(0.013)	-0.044***	(0.013)
(Positive Group) * Reset, lag 2 weeks	-0.077***	(0.016)	-0.080***	(0.015)	-0.083***	(0.015)
(Positive Group) * Reset, lag 4 weeks			-0.129***	(0.016)	-0.131***	(0.015)
(Positive Group) * Reset, lag 6 weeks					-0.128***	(0.015)
% Positive Change	-0.189***	(0.019)	-0.193***	(0.019)	-0.197***	(0.019)
Positive Group	0.041***	(0.010)	0.045***	(0.010)	0.051***	(0.010)
(% Negative Change) * Reset	0.387***	(0.070)	0.381***	(0.069)	0.375***	(0.068)
(% Negative Change) * Reset, lag 2 weeks	0.167**	(0.064)	0.168**	(0.062)	0.165**	(0.061)
(% Negative Change) * Reset, lag 4 weeks			0.207***	(0.055)	0.209***	(0.054)
(% Negative Change) * Reset, lag 6 weeks					-0.075	(0.056)
(Negative Group) * Reset	-0.018	(0.013)	-0.019	(0.013)	-0.020	(0.013)
(Negative Group) * Reset, lag 2 weeks	-0.093***	(0.012)	-0.094***	(0.012)	-0.094***	(0.012)
(Negative Group) * Reset, lag 4 weeks			-0.053***	(0.011)	-0.053***	(0.011)
(Negative Group) * Reset, lag 6 weeks					-0.182***	(0.011)
% Negative Change	0.253***	(0.039)	0.254***	(0.039)	0.260***	(0.039)
Negative Group	0.010	(0.012)	0.016	(0.012)	0.022	(0.012)
% Change in Store AGs	-0.000	(0.000)	-0.000	(0.000)	-0.000	(0.000)
(% Change in Store AGs)*Reset	0.002	(0.002)	0.002	(0.002)	0.002	(0.002)
Facings wide, % of total ice cream	-0.003	(0.019)	-0.001	(0.018)	0.003	(0.018)
Facings deep, % of total ice cream	0.162***	(0.049)	0.165***	(0.047)	0.167***	(0.046)
Facings high, % of total ice cream	0.155***	(0.017)	0.154***	(0.016)	0.154***	(0.015)
Avg. temperature (deg F, 00s)	0.432**	(0.147)	0.423**	(0.149)	0.404**	(0.150)
No. of UPCs in store (00s)	0.324***	(0.015)	0.326***	(0.015)	0.328***	(0.015)
Constant	1.152***	(0.135)	1.139***	(0.135)	1.138***	(0.135)
Adjusted R^2	0.713		0.712		0.711	
Observations	1,700,658		1,804,712		1,908,719	
Division FEs	Y		Y		Y	
Two-week FEs	Y		Y		Y	
(Division FE)*(Period FE)	Y		Y		Y	
AG FEs	Y		Y		Y	

Standard errors in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

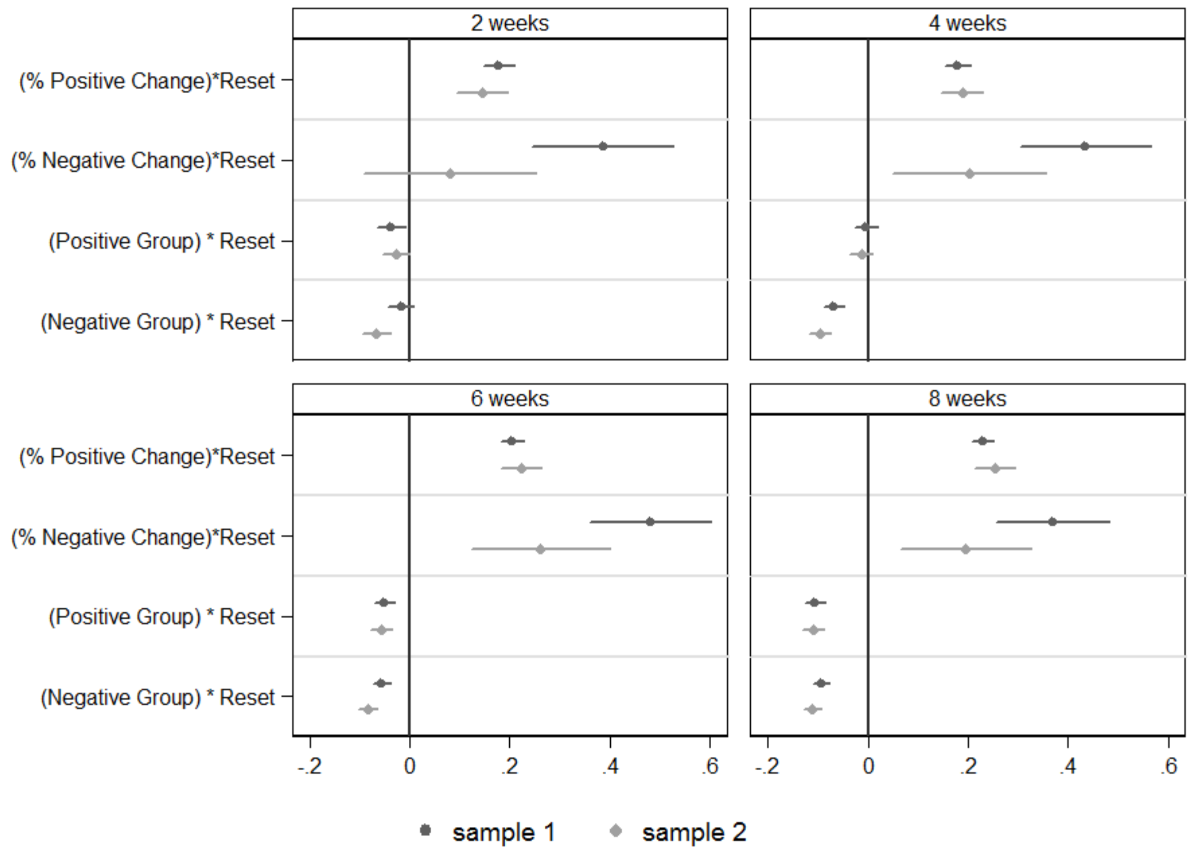
Table 1.24: Robustness Test of PCC with Lags, Sample 2

	(1)		(2)		(3)	
	b	se	b	se	b	se
Reset	0.017	(0.014)	0.018	(0.014)	0.018	(0.014)
Reset, lag 2 weeks	0.061*	(0.024)	0.061*	(0.025)	0.061*	(0.025)
Reset, lag 4 weeks			0.097**	(0.036)	0.097**	(0.036)
Reset, lag 6 weeks					0.155***	(0.046)
(% Positive Change) * Reset	0.147***	(0.026)	0.147***	(0.026)	0.149***	(0.026)
(% Positive Change) * Reset, lag 2 weeks	0.115***	(0.027)	0.115***	(0.027)	0.115***	(0.027)
(% Positive Change) * Reset, lag 4 weeks			0.192***	(0.033)	0.193***	(0.033)
(% Positive Change) * Reset, lag 6 weeks					0.196***	(0.031)
(Positive Group) * Reset	-0.030*	(0.013)	-0.032*	(0.013)	-0.034**	(0.012)
(Positive Group) * Reset, lag 2 weeks	-0.050**	(0.015)	-0.054***	(0.015)	-0.056***	(0.015)
(Positive Group) * Reset, lag 4 weeks			-0.107***	(0.016)	-0.110***	(0.016)
(Positive Group) * Reset, lag 6 weeks					-0.105***	(0.016)
% Positive Change	-0.228***	(0.034)	-0.234***	(0.034)	-0.242***	(0.034)
Positive Group	0.038**	(0.012)	0.041***	(0.012)	0.046***	(0.012)
(% Negative Change) * Reset	0.075	(0.085)	0.064	(0.084)	0.053	(0.082)
(% Negative Change) * Reset, lag 2 weeks	-0.105	(0.075)	-0.110	(0.073)	-0.117	(0.072)
(% Negative Change) * Reset, lag 4 weeks			-0.168*	(0.066)	-0.169**	(0.064)
(% Negative Change) * Reset, lag 6 weeks					-0.368***	(0.065)
(Negative Group) * Reset	-0.068***	(0.014)	-0.070***	(0.014)	-0.072***	(0.014)
(Negative Group) * Reset, lag 2 weeks	-0.132***	(0.012)	-0.134***	(0.012)	-0.135***	(0.012)
(Negative Group) * Reset, lag 4 weeks			-0.093***	(0.011)	-0.094***	(0.011)
(Negative Group) * Reset, lag 6 weeks					-0.210***	(0.011)
% Negative Change	0.261***	(0.047)	0.268***	(0.047)	0.285***	(0.047)
Negative Group	-0.005	(0.012)	0.002	(0.012)	0.009	(0.012)
% Change in Store AGs	-0.000	(0.000)	-0.000	(0.000)	-0.000	(0.000)
(% Change in Store AGs)*Reset	0.002	(0.002)	0.002	(0.002)	0.002	(0.002)
Facings wide, % of total ice cream	0.044**	(0.016)	0.046**	(0.016)	0.049**	(0.015)
Facings deep, % of total ice cream	0.087*	(0.041)	0.091*	(0.041)	0.094*	(0.040)
Facings high, % of total ice cream	0.143***	(0.013)	0.142***	(0.013)	0.141***	(0.013)
Avg. temperature (deg F, 00s)	0.416**	(0.154)	0.408**	(0.156)	0.391*	(0.157)
No. of UPCs in store (00s)	0.373***	(0.014)	0.376***	(0.014)	0.378***	(0.014)
Constant	1.494***	(0.139)	1.478***	(0.139)	1.473***	(0.140)
Adjusted R^2	0.713		0.713		0.711	
Observations	958,568		1,016,593		1,074,631	
Division FEs	Y		Y		Y	
Two-week FEs	Y		Y		Y	
(Division FE)*(Period FE)	Y		Y		Y	
AG FEs	Y		Y		Y	

Standard errors in parentheses

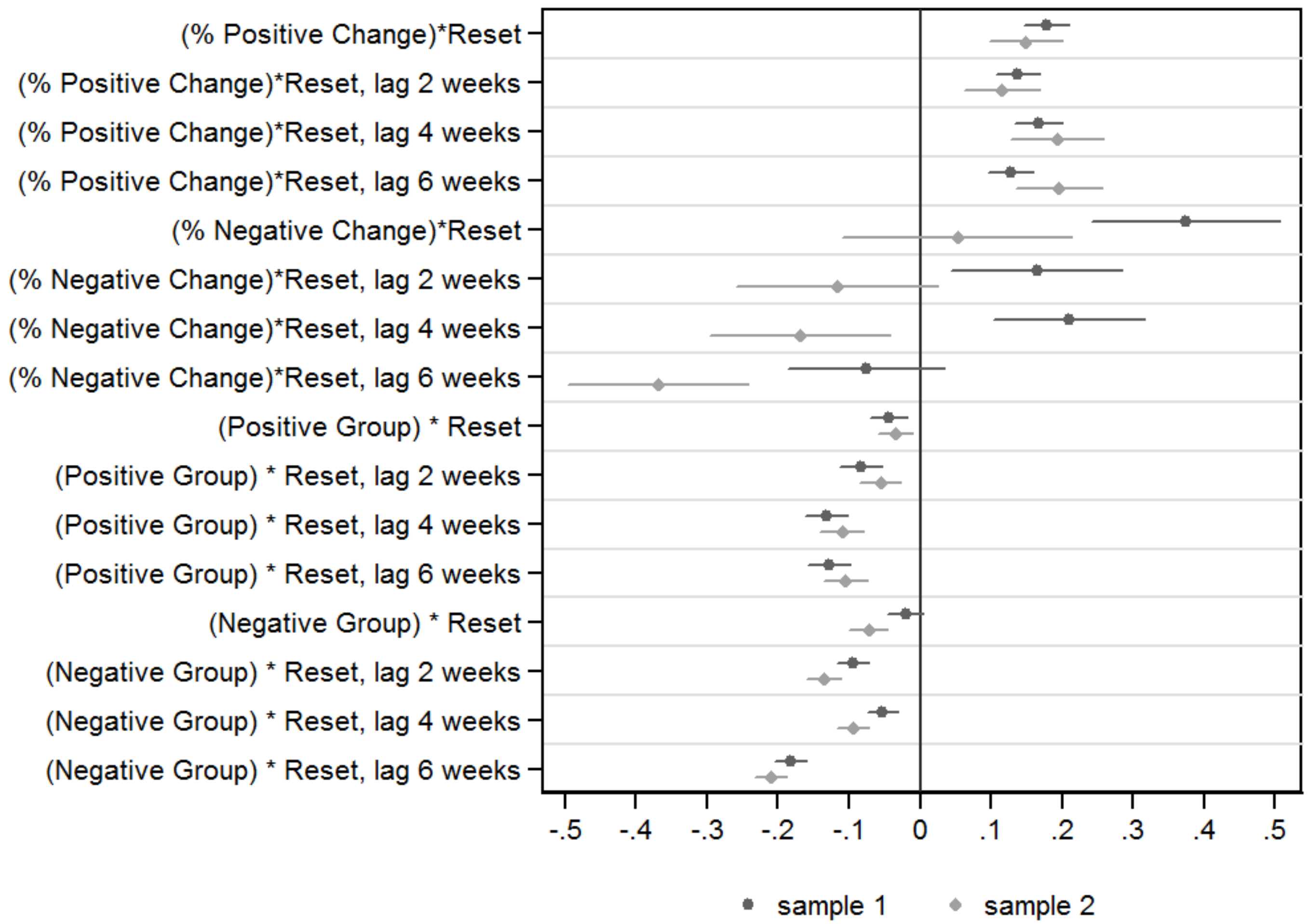
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Figure 1.8: Robustness Test Estimates of change in variety variables



Point estimates and 95% confidence intervals (not exponentiated) of the change in variety variables from specification (1.8) are plotted for each definition of the treatment period. Results are strikingly similar to those from specification (1.6).

Figure 1.9: Robustness Test Estimates of Change in Variety Variables with Lags



Point estimates and 95% confidence intervals (not exponentiated) of the change in variety variables from specification (1.9) are plotted for each definition of the treatment period. Results are strikingly similar to those from specification (1.7) comparing Figures (1.8) and (1.9) to Figures (1.3) and (1.5). While there are minor variations in point estimates, the signs and significance levels for these variables are equivalent between the models. Because this specification uses less precise division fixed effects in place of store fixed effects, the estimates of the temperature and number of UPC control variables change notably. However, because the interaction between the division and period fixed effects capture price changes (see discussion in Section (1.3)), the consistency of the change of variety estimates indicates that our analysis is not biased due to changes in price.

1.5 Conclusion

In this paper, we use a detailed dataset of product-level shelf schematics and transactions from a national grocer to measure the quantity response from changes in ice cream product variety. Our empirical setting incorporates a natural experiment in which stores adopted variety changes gradually over two months, allowing us to apply a difference in differences strategy to identify causal effects. We find that on average, increases in variety within a product group led to higher sales and decreases in variety caused sales losses in the short run. However, these average effects became more negative over time for both proliferations and reductions in variety, to the extent that product groups with increases in variety actually experienced losses in sales six weeks after the treatment period.

The results from our analysis predict that larger increases in variety increase sales more, and larger decreases in variety reduce sales more, with an equivalent magnitude decrease having a stronger effect than an increase. Further, larger increases in variety sustained improved sales for at least six to eight weeks, but smaller increases in variety experienced sales losses in the longer term. Finally, we find that initial sales losses from large variety reductions were followed by more moderate losses, so that, in terms of percent sales losses, large variety reductions performed better than both small increases and decreases in variety in later periods.

While our research addresses the own effects of a variety increase on a product group, future research might further investigate cross effects, both within brand and between brands, to determine whether variety changes in a particular product group affect sales of product groups with no variety changes. In addition, sales responses to changes in variety may result from consumers switching purchases within a product category, substitution of purchases over time, or the entrance and exit of consumers into a product category. Sloot, Fok, and Verhoef (2006) explore such consumer responses in the context of laundry detergent product removals, but a similar analysis in our setting could extend our understanding of consumer responses to product introductions. Finally, we evaluate purchase quantities but not profits in response to assortment changes. Future research could incorporate the costs of introducing or maintaining products on shelf, to evaluate profit maximizing strategies for retailers.

Chapter 2

The Effect of Quality Disclosure on Nearby Firms

2.1 Introduction

Do firms respond to quality disclosure by nearby competitors? This paper demonstrates that the Los Angeles County restaurant hygiene disclosure law had a spillover effect on neighboring, unregulated areas. This experimental design overcomes identification issues that plagued earlier literatures on product differentiation and agglomeration by using exogenous variation in consumer information as a result of the Los Angeles County hygiene policy.

On January 16, 1998, when the L.A. County grade card policy came into effect, the county began to provide every restaurant with a hygiene grade card at all inspections. Over time, cities also individually adopted policies requiring restaurants to post the grade card prominently at their front entrances. In an important paper, Jin and Leslie (2003), hereafter JL, confirm that the L.A. County grade card policy improved average restaurant hygiene. Following JL, the term "mandatory restaurant" refers to a restaurant in any city that has already implemented the required posting policy. "Voluntary restaurants" refers to restaurants in the period after the county began providing grade cards, but in cities that have not yet adopted required posting.[1] Finally, the term "mandatory posting restaurants" refers to mandatory restaurants that are physically posting their grade cards (e.g. they have been inspected since the adoption of the grade card policy and therefore have a grade card to post).

This paper demonstrates that voluntary restaurants improved hygiene scores more when near mandatory posting restaurants, implying the mandatory posting policy positively affected voluntary restaurant hygiene. To investigate spatial relationships between restaurant

[1] The data do not include individual posting decisions of firms in voluntary posting areas and can thus only examine the restaurant hygiene scores. Thus, the terms "voluntary" and "voluntarily posting" refer to restaurants or regions where grade cards are provided but restaurants are not required to post them.

hygiene and quality disclosure policies, a unique dataset was constructed by combining inspection data from the Los Angeles County Department of Health Services (DHS) with high resolution geographic data.[2] In addition to restaurant hygiene scores on a 100-point scale, the analysis here uses the distance to the nearest mandatory posting restaurant, for every inspection of a voluntary restaurant. Prior empirical analyses on the role of inter-firm distance on firm behavior have suffered from an identification challenge due to endogenous location choice (see Davis (2006) for a discussion). As further discussed below, the rollout of the mandatory policy is rapid and unrelated to the hygiene of nearby restaurants, precluding such identification problems. For voluntary restaurants, the average marginal effect of being 4.0 miles closer to a mandatory posting restaurant was a one point increase in improved hygiene, compared to an overall average increase of three points for all voluntary restaurants.

The focus here on market geography confirms that mandatory posting policies on nearby restaurants affected the hygiene choices in voluntary restaurants. In their paper, JL discuss (but do not test) two possible explanations for average hygiene scores to increase in voluntary areas, but neither explanation is inherently spatial. Specifically, the authors reference the informational unraveling literature, which suggests the credible disclosure of hygiene through grade cards may be sufficient to improve hygiene. Also, JL suggest owners and managers of voluntary restaurants may improve sanitation practices in anticipation of mandatory posting adoption by their city. Neither informational unraveling nor manager anticipation necessarily suggest a different hygiene response by restaurants within the same city, and thus neither can fully explain the spatial results provided in this paper.

Why might voluntary restaurants improve hygiene more when near mandatory posting restaurants? The existing economic literature explores at least two important roles of distance in firm behavior. First, on the demand side, a rich product differentiation literature explores markets where consumers have preferences over product attributes such as firm location or product quality.[3] In this framework, producing differentiated goods mitigates price competition between firms.[4] Both the theoretical and the empirical literatures on product differentiation tend to focus on the relationship between price and distance, especially since quality and other non-price attributes may be difficult to measure in practice. An exception is Datta and Sudhir (2010), who find an empirical relationship between reduced geographic differentiation due to zoning and increased diversity in retail formats (e.g. food retailers may diversify into supermarkets, convenience stores, and mass-merchandising, etc.). In addition, Neven and Thisse (1990) investigate a theoretical model in which firms choose both loca-

[2]After the time period in the dataset, the DHS title was changed to the "Department of Public Health," but the acronym "DHS" in maintained in this paper.

[3]Most theoretical analyses in this field explore differentiation along a single attribute, either location (see the literature stemming from Hotelling (1929); Salop (1979); Samuelson (1952)) or quality (Gabszewicz and Thisse (1979); Shaked and Sutton (1982), and subsequent works), and thus do not inform how restaurant location may affect hygiene choices.

[4]Empirical support for this theory is provided by Davis (2006), who applies discrete choice methods as in Berry, Levinsohn, and Pakes (1995) and finds price increases tend to affect revenues of competing theaters less when they are further away.

tion and quality. The authors find that firms located close together ease price competition by maximally differentiating in terms of quality. Because restaurants under the mandatory grade card policy tended to increase hygiene scores, the framework in Neven and Thisse (1990) might have predicted the opposite result from this paper—that, in order to ease price competition, voluntary firms would improve hygiene less when located closer to mandatory firms. However, Neven and Thisse (1990) note their results are sensitive to assumptions about the distribution of consumers, the transportation cost function, and the dimensionality of the product space. As a result, I do not interpret the results here as suggesting the role of distance on the demand side is necessarily minor.

Second, on the supply side, proximity may enhance cooperative interactions between firms (Fujita and Thisse, 2002). For example, if restaurant employees or owners communicate with each other, restaurants in voluntary areas may replicate hygiene practices adopted by nearby restaurants under mandatory posting. This "spillover" argument has been studied extensively to emphasize the positive role of face-to-face communication on Silicon Valley's emergence as a center for innovative activity (Saxenian, 1994). Although they do not consider heterogeneity in product quality, Mai and Peng (1999) demonstrate theoretically that inclusion of a spillover effect may reduce the competitive pressure firms face when located near each other. Thus, to the extent that price competition drives restaurants near each other to differentiate in terms of hygiene, cooperative opportunities might cause voluntary restaurants to improve quality when near mandatory restaurants.

In summary, this paper contributes to the existing literature by describing the spatially-driven heterogeneity in voluntary restaurant hygiene choices due to the Los Angeles County grade card system. Previous work suggests that distance may enhance both competitive pressures and cooperative opportunities between firms, and it is not clear ex ante how distance to a mandatory restaurant will affect the hygiene choices of voluntary restaurants. The analysis here uses exogenous variation from the implementation of the grade card system to identify the role of distance in inter-firm responses, and the empirical evidence demonstrates firms improve quality more when they are nearer to restaurants under a policy of mandatory hygiene disclosure.

The rest of this paper proceeds as follows. In Section 2, background information regarding the grade card system, a description of the data, and preliminary results are provided. Section 3 demonstrates that voluntary restaurants improve hygiene more when they are closer to a mandatory restaurant, and Section 4 concludes.

2.2 The grade card system

The DHS performs randomly timed routine inspections of virtually all food establishments in Los Angeles county, with restaurants receiving an average of 2.4 inspections per year during the period covered by this analysis.[5] Violations of the California Retail Food

[5]The DHS inspects every restaurant in the county, except those in the cities of Long Beach, Pasadena, and Vernon, who perform their own inspections and have been excluded from this analysis since they do not

Code are deducted from a maximum possible grade of 100 points, and facilities are issued a grade card at the time of inspection.[6] During 1998, all inspected restaurants received grade cards. "A" grade cards were given to restaurants scoring 90 to 100 points, "B" for 80 to 89 points, and "C" for 70 to 79 points. Scores of 69 and below received a grade card denoting the numeric score. In mandatory areas, the DHS required restaurants to post their grades until the next inspection, at which time the inspector provided a new grade card reflecting the new score. An inspection could result in restaurant closure if the inspector identified an "imminent health hazard," no matter what the overall inspection score. For areas requiring grade card posting, a Notice of Closure would be posted.

As noted in JL, the grade card system arose as a rapid and unanticipated reaction to hidden camera news reports of poor hygiene in restaurants in November 1997. The resulting county ordinance went into effect at the county level on January 16, 1998, providing standard format grade cards to all restaurants inspected in the county, requiring prominent front-door posting in all unincorporated areas, and allowing cities to individually adopt mandatory front door posting. Between this date and June 5, 2005, 77 of the 88 cities in the county adopted the mandatory posting policy, with 22 cities (28.6% of adopters) adopting in the first six months of 1998. By the end of 1998, 49 cities (63.6% of adopters) adopted mandatory posting, on 43 unique days. Prior to the implementation of the grade card system, customers concerned about hygiene practices had to ask restaurants to view inspection reports. JL report anecdotal evidence suggesting customers rarely made such requests, and members of the DHS have noted that consumers were largely unaware of inspection results. It is therefore reasonable to regard restaurant hygiene as largely unknown to consumers until the introduction of grade cards.

Average hygiene scores, the number of inspections, and the number of restaurants inspected over time are provided in Table 2.1. The increase in scores after the second quarter in 1997 reflects a change in inspection criteria on July 1st of that year. Score increases in 1998 are a combination of the grade card policy and the second change in inspection criteria on March 18th.[7] Both of the changes in the inspection process led to systematic changes in restaurant scores and are controlled for in the analysis.

JL estimated an average hygiene increase of 4.40 points (5.3%) for mandatory restaurants and 3.25 points (3.9%) for voluntary restaurants, and find that the magnitude of the effects are statistically different from each other. Using inspection data combined with tax revenues and statewide hospital admissions, they argue: (1) good hygiene grades are as-

---

employ the grade card system during the time period covered by the data.

[6]For example, a one point deduction is issued for detection of vermin without a major infestation, for toilets that are dirty or in disrepair, for hair found in food, and for certain improper food thawing practices. Four points are deducted for certain improper food storage practices, for ready-to-eat food stored in contact with raw meat, for preparing food on the floor, and for reserving foods such as tortilla chips and salsa. Six points are deducted for major vermin infestations, for employees not washing hands after toilet use or after handling raw chicken, for lack of operable toilets, and for food employees having open wounds on hands.

[7]The first change in inspection criteria excluded a previously used subjective element in the inspection scoring, in which inspectors could deduct points (up to 40 points) for less than excellent overall restaurant hygiene. The second change in scoring criteria added a few new possible violations.

Table 2.1: Hygiene scores over time, all inspections

Quarter		Hygiene Score Mean	SD	No. of Inspections	No. Restaurants Inspected†
1995	Q3	74.37	15.55	9298	8929
	Q4	75.76	14.97	9967	9564
1996	Q1	75.63	14.90	11969	11449
	Q2	75.53	14.64	11754	11155
	Q3	75.09	14.76	12516	11976
	Q4	75.28	14.99	11131	10617
1997	Q1	75.91	14.54	12040	11562
	Q2	75.20	14.87	14188	13481
	Q3	84.66	10.44	9921	9490
	Q4	82.46	12.12	10645	10202
1998	Q1	87.09	10.10	12146	11570
	Q2	90.71	7.36	17353	15862
	Q3	89.64	7.50	9859	9497
	Q4	90.08	6.81	11103	10737
All Dates		80.74	14.25	163890	37533

† *No. Restaurants Inspected* provides the number of unique restaurants inspected in the given time period

sociated with higher restaurant revenue only after grade cards are introduced, and (2) the grade card policy led to reduced hospitalizations for digestive disorders, implying increased hygiene scores at least partly reflect actual hygiene improvements.

The panel dataset in this paper includes every inspection of restaurants and retail food markets performed by the Los Angeles County DHS between July 1995 through December 1998.[8] For each inspection, the DHS data contain the inspection date, numeric hygiene score, any violations found, and the inspected restaurant's name, address, and general size. This paper constructs a new dataset containing the geographic coordinates of every site inspected and, from these coordinates, distances (in miles) to other restaurants. In particular, the analysis here uses the minimum crow-flies distance to a mandatory posting restaurant. For restaurants inspected between 1995-1999, the data also include the type of food served by a restaurant (i.e. Chinese versus Mexican).[9] The dataset is limited in that it includes neither firm prices nor revenues.

2.2.1 Geographic data

The geographic coordinates of every restaurant derive from the restaurant address information and the inspection data with ESRI StreetMaps geographic data and searches on GoogleMaps. For each inspection of a voluntary restaurant, measure "MD1" refers to the minimum distance to a mandatory restaurant that was posting a card at the time of the voluntary restaurant's inspection. To better understand MD1, consider an observation in which voluntary restaurant i is inspected on date t. MD1 is then the distance from restaurant i to the nearest mandatory restaurant that existed and was posting on date t.[10] That is, MD1 is not determined from the distance to any restaurant that closed before or opened after date t. In addition, mandatory restaurants will not have a grade card (and therefore cannot post) until they undergo their first inspection after January 16th, 1998. In determining MD1, distances to mandatory restaurants that do not have a grade card by date t are not considered. MD1 is then the minimum distance to any of the remaining mandatory restaurants, and thus provides the minimum distance to a mandatory restaurant posting a grade card at the time of the voluntary restaurant's inspection. Because the timing of inspections is random and the implementation of mandatory posting is exogenous to city restaurant characteristics, whether a mandatory restaurant has a grade card to post is a source of exogenous variation.

The evolution of MD1 for inspections of voluntary restaurants over time are displayed in Table 2.2. The decrease in the mean distance reflects the implementation of mandatory posting by cities throughout the year. For all inspections of voluntary restaurants, Table 2.3 provides the percent of inspections in one-mile categories of MD1. In over half of the total

[8] At the end of 1998, the DHS ceased to provide grade cards in cities without the mandatory posting policy so as to encourage further adoption of mandatory grade card posting at the city level.

[9] Food type is determined using the restaurant's website, or other websites such as OpenTable, Urbanspoon, and Menuism.

[10] A restaurant is assumed to exist on date t if it is inspected within 4 months of date t, or if it is inspected both before and after date t.

inspections, voluntary restaurants are within 2 miles of a mandatory posting restaurant. In over 95% of the inspections of voluntary restaurants, the nearest mandatory posting restaurant was less than 6 miles away.

Exogenous variation in the minimum distance measure is provided by the random timing of mandatory adoption with respect to city-level restaurant characteristics. Following JL, analysis of a survival model demonstrates that the timing of city adoption is unrelated to the hygiene of restaurants within the city and to the hygiene of restaurants outside the city but within a one-half mile of a city border:

$$\begin{aligned} ln\ f_j(t) = \alpha(t) + \underline{\gamma}(demog\ vars)_j + \beta_1(avg\ grade)_i + \beta_2(prop\ \text{“}A\text{”})_j \\ + \beta_3(no.\ rests)_j + \beta_4(avg\ grade,\ outside)_j \\ + \beta_5(prop\ \text{“}A\text{”},\ outside)_j + \beta_6(no\ rest\ outside)_j + \epsilon_j \end{aligned} \tag{2.1}$$

where $f_j(t)$ is the probability of city i adopting the mandatory policy at time t, given that it hasn't yet done so. Variables $(avg\ grade)_j$ and $(avg\ grade, outside)_j$ provide the average hygiene grade of restaurants within city j and of restaurants within 0.5 miles of the city, respectively. Similarly, $(prop\ \text{“}A\text{”})_j$ and $(prop\ \text{“}A\text{”}, outside)_j$ provide the proportion of restaurants with “A” grades within and one mile outside of city j, respectively. Controls are included for the number of restaurants within the city and the number within one mile outside of the city, respectively $(no.\ rests)_j$ and $(no.\ rests\ outside)_j$, and for within-city demographic variables $(demo\ var)_j$. Table 2.4 provides the results of the survival model and demonstrates that none of the restaurant hygiene variables are related to the timing of city adoption.

Figure 2.1 provides the distribution of the average change in hygiene scores before versus after a restaurant becomes voluntary, related to MD1. The data are aggregated into one-mile bins. The negative trend between improvements in hygiene scores and MD1 implies mandatory posting does have an effect on nearby voluntary restaurants, and this result is formalized in the next section. The trend is particularly evident for distances between zero and six miles, where over 96% of the data lie. Score percentiles become erratic and distributions narrow for values above ten miles in both plots, where the number of observations decreases dramatically, as implied in Table 2.3. It is likely that the small sample size for these distances causes the volatile pattern, but there may also be omitted factors at longer distances.

2.3 The Effect of Proximity to Mandatory Posting Restaurants on Hygiene Scores

To determine whether the hygiene scores of voluntary restaurants differ based on proximity to mandatory posting restaurants, this paper estimates the following regression:

$$h_{it} = \alpha_i + q_t + \beta_1 mandatory_{it} + \beta_2 voluntary_{it} + \beta_3 (vol)_{it} * MD1_{it} + \gamma_2 c_{2t} + \epsilon_{it} \tag{2.2}$$

Table 2.2: Minimum Distance to a Mandatory Posting Restaurant, for Inspections of Voluntary Restaurants in Quarters of 1998

Quarter		No. of Inspections	Hygiene Score Mean	Hygiene Score SD	Min Distance (miles) to Mandatory Posting Restaurant (MD1†) Mean	Min Distance (miles) to Mandatory Posting Restaurant (MD1†) SD
1998	Q1	9459	87.94	9.30	3.516	2.729
	Q2	6276	91.16	7.43	1.566	1.756
	Q3	2921	90.04	7.88	1.265	0.990
	Q4	2430	90.21	6.94	1.067	0.946
	Total	21086	89.45	8.45	2.342	2.375

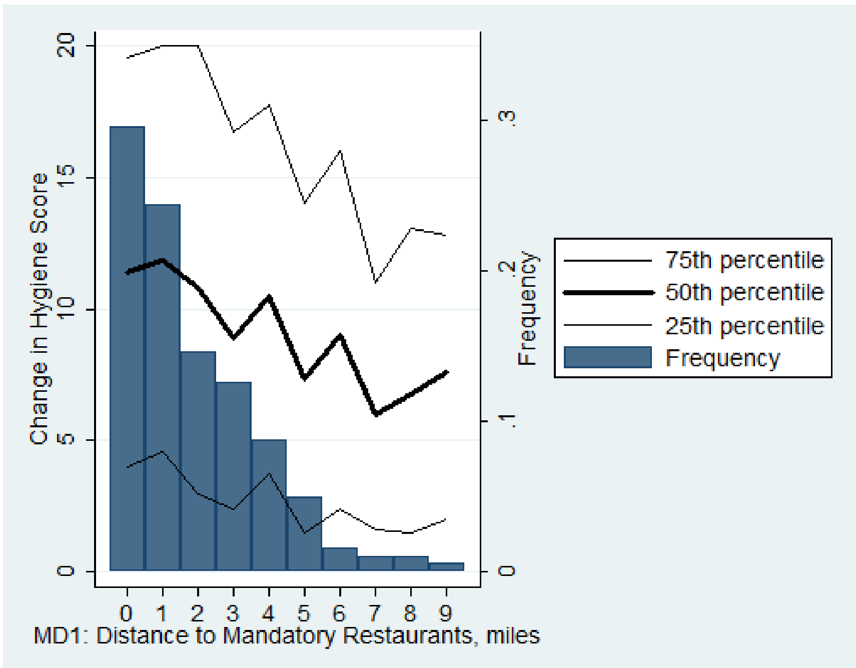
Table 2.3: Distribution of Inspections and Inspection Rates of Voluntary Restaurants by Minimum Distance to Mandatory Posting Restaurants (MD1)

MD1 (mi)	Inspections No.	Inspections %	Restaurants	Insp per Rest Mean	Insp per Rest SD
0-1	5346	29.6	4419	1.29	0.54
1-2	4415	24.4	3679	1.20	0.47
2-3	2644	14.6	2367	1.12	0.37
3-4	2283	12.6	2189	1.04	0.21
4-5	1584	8.8	1546	1.02	0.16
5-6	896	5.0	875	1.02	0.15
6-7	283	1.6	271	1.04	0.21
7-8	193	1.1	189	1.02	0.14
8-9	192	1.1	188	1.02	0.14
9-10	117	0.6	116	1.01	0.09
10+	138	0.8	138	1.00	0.00

All statistics for inspections of voluntary restaurants only

† For voluntary restaurant i being inspected on date t, "MD1" is the minimum distance in miles to a restaurant that on t is both in a mandatory posting jurisdiction and has a a grade card to post (i.e. the mandatory restaurant was inspected between January 16th, 1998 and t).

Figure 2.1: Distribution of Average Change in Scores for Inspections of Voluntary Restaurants over Minimum Distance to Mandatory Posting Restaurants (MD1)



Inspections data aggregated in bins of one mile. The vertical axis provides the change in score for MD1, the minimum distance to a mandatory restaurant that is posting (for further description of this measure, see text or Tables 2.2 and 2.3, where the change in score is equal to the average score for a restaurant when it is voluntary, less its average score before it becomes voluntary. "Frequency" is the percent of voluntary restaurants with MD1 values that fall in the given one-mile bin, where MD1 is averaged if a restaurant receives multiple inspections while voluntary. Over 95% of restaurants have MD1 values of less than six miles.

Table 2.4: Proportional Hazards Estimation of City Adoption

	Hazard Ratio	(SE)
Average hygiene score§	0.987	(0.110)
Proportion of "A" grades§	2.676	(9.179)
Total Restaurants§	0.996	(0.006)
Average hygiene score, outside§	1.00	(0.0843)
Proportion of "A" grades, outside§	5.215	(12.458)
Total Restaurants, outside§	0.998†	(0.001)
Demographic variables§§	Yes	
Observations	70	
Log-likelihood	-214.15	

Standard errors in parentheses (not corrected for spatial dependence).

† $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

§Calculated using data for the date of the news report on restaurant hygiene (November 18, 1997).
§§Demographic variables are the average of the 1990 and 2000 census variables (total population, total male population, median age, total households, population by ethnic groups, average household size, median income, poverty rate, etc.)

In the above equation, an observation is an inspection at a given restaurant i, on a given day t, where h_{it} is the numeric hygiene score out of 100 points. The variables *mandatory* and *voluntary* are dummies indicating the grade card regime of the city the inspected restaurant is located in (these variables do not refer to the actual posting behavior of the inspected restaurant). Before grade cards are introduced on January 16, 1998, every inspection receives zero for both of these dummies, whereas after this date, all inspections fall under either *mandatory* or *voluntary*. Restaurant inspections with a value of one for *mandatory* occurred in cities requiring posting at the time of inspection, while inspections with a value of one for *voluntary* were performed at restaurants that had a grade card, but were not required to post them. The coefficient of interest is on the interaction between the voluntary dummy and MD1, the minimum distance to a mandatory restaurant that is posting at time t. In equation (2.2), β_3 captures the effect of proximity to mandatory posting restaurants on the hygiene scores of voluntary restaurants.

Equation (2.2) includes restaurant fixed effects α_i to control for any possible time-invariant effects and quarter of the year time dummies, q_t. While the first change in inspection criteria is captured by the quarterly fixed effects, the additional dummy c_{2t} controls for the second change in inspection criteria discussed in the Background section. The c_{2t} variable has a value of one for all dates beginning March 18, 1998.

Restaurant hygiene outcomes may be correlated over space for reasons unrelated to the introduction of grade cards. For example, pest infestations or access to outside suppliers of hygiene maintenance services may be locally determined. To eliminate the bias associated with such spatial correlation, this paper reports robust standard errors corrected for spatial dependence as in Conley (1999). In this specification, errors may be more correlated based on their proximity in terms of both geographic location and time of inspection. Although not reported here, both standard errors clustered at the city level and two-way standard errors clustered by city and by quarter of the year (as suggested in Cameron, Gelbach, and Miller (2006)) were similar to, and often smaller than, standard errors corrected for spatial dependence.

Table 2.5 provides the main result of this paper in Column (6). Because the specification and dataset are slightly different than JL, the table includes columns (1) to (5) to demonstrate the effect of these differences. Columns (1) and (2) of Table 2.5 provide results using JL's specification on their data versus the data in this paper, respectively. Qualitatively, the estimates are the same between the two datasets, and the mandatory and voluntary policies lead to significant increases in hygiene scores in both datasets. However, in this paper, the point estimates for the effects of the mandatory and voluntary policies are slightly smaller than those of JL. There are several sources of this discrepancy. First, this analysis uses the entire sample of inspections while JL only use restaurants for which they have confidential tax revenue data, leading them to drop about 43% of the restaurants from their original sample. Second, a comparison of JL's raw DHS data with data directly from the DHS for the dates July-December 1998 demonstrates that the DHS provided slightly different datasets each time. This analysis combines both datasets into one. Finally, the dataset here uses cleaned city data based on the geographical location of the restaurant, correcting the assignment of grade card policy for a small set of observations.

Comparing Column (3) with Column (2) of Table 2.5 demonstrates that the addition of time fixed effects slightly decreases the estimates and significance of the *mandatory* and *voluntary* variables, but again the results are qualitatively similar to the case without time fixed effects. When using these time fixed effects, the effect of the first inspection criteria change (on July 1, 1997) is captured by a quarter dummy. Columns (4) and (5) replicate Columns (3) and (2) respectively, but utilize standard errors corrected for spatial dependence as discussed above, leading to only minor decreases in significance for the regression using quarter fixed effects.

The principal result from estimation (2.2) is provided in Column (6) of Table 2.5, where the variable $vol * (MD1)$ interacts the voluntary dummy with the minimum distance to a mandatory posting restaurant, as described in Section (2.2.1). The negative and significant estimate on the MD1 variable for voluntary restaurants supports mandatory posting having a spatially attenuated effect on the hygiene of voluntary restaurants. Thus, while the average effect of the voluntary policy was an increase of 3.07 points, implementing the mandatory posting policy on restaurants one mile further away reduced the hygiene improvement in response to the voluntary policy by an average of 0.244 points. This marginal effect would move the median voluntary restaurant hygiene score to the 43rd percentile, and is equivalent to 2.9% the standard deviation in voluntary restaurant scores. Given the average distance over the time period of this sample, the average magnitude of the effect constituted a 0.58 point decrease.

Interestingly, including a quadratic term for minimum distance did not explain additional variation in the hygiene scores. Although not shown here, when including a squared

Table 2.5: Effect of Proximity to Mandatory Posting Restaurants on Hygiene Scores

	Jin & Leslie (2003)	This paper, all dates				
	(1)	(2)	(3)	(4)	(5)	(6)
mandatory	4.396***	3.069***	2.638**	3.069***	2.638**	3.069***
	(1.405)	(0.886)	(1.143)	(0.743)	(0.886)	(0.880)
voluntary (average effect)	3.253***	2.372***	2.199***	2.372***	2.199**	3.043***
	(0.355)	(0.323)	(0.481)	(0.652)	(0.701)	(0.744)
vol*(MD1)‡						-0.244*
						(0.102)
inspection criteria II§	8.089***	8.922***		8.922***		
	(0.991)	(0.606)		(0.464)		
inspection criteria III§§	10.416***	11.550***	1.609***	11.550***	1.609**	1.478**
	(1.354)	(1.032)	(0.495)	(0.715)	(0.530)	(0.532)
restaurant FEs	Yes	Yes	Yes	Yes	Yes	Yes
time FEs	No	No	Yes	No	Yes	Yes
Observations	69,991	163,890	163,890	163,890	163,890	163,890
Percent of total observations	42.3%	100%	100%	100%	100%	100%
adj.R^2	0.587	0.506	0.506	0.506	0.506	0.506
	Data only for restaurants paired to revenue data	All restaurants				

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$
Standard errors in parentheses; clustered on the city level for (1), (2), and (3), and corrected for spatial dependence (Conley (1999)) in (4)-(6).

‡ $vol * (MD1)$ interacts voluntary dummy with the minimum distance in miles to a mandatory restaurant that is posting on date t (MD1). See text for further details.
§ *inspection criteria II* is a dummy with value 1 for all inspections between July 1, 1997 and March 18, 1998.
§§ *inspection criteria III* is a dummy with value 1 for all inspections after March 18, 1998.

Column (1) displays results from Jin and Leslie (2003), who use a subsample of the DHS data matched with revenue data, while (2) presents the same estimation using this paper's fuller inspections dataset (discussion of differences in Results section). Column (3) repeats the analysis in column (2), but adds quarterly fixed effects. Columns (4) and (5) replicate columns (2) and (3), respectively, but use standard errors corrected for spatial dependence. Column (6) provides the main result of this paper: voluntary restaurants further from mandatory posting restaurants improve hygiene scores less.

term, the regression estimates on the spatial variables are jointly significant (not individually significant), but the point estimates indicate that the nonlinearity only becomes meaningful outside of the relevant distance range (i.e. above 75 miles).

Note that the results in Table 2.5 Column (6) directly indicate that restaurants respond to the conditions of neighbors, and cannot be completely explained by firms improving grades in anticipation of their cities adopting mandatory posting. It is possible that through witnessing the example of their neighbors, owners of restaurants located nearer to mandatory posting restaurants were more inclined to believe their own city would adopt mandatory posting and thus increased scores more. However, if anticipation levels of restaurant owners are not affected by the conditions of neighbors, then we would not expect distance to be significantly correlated with hygiene scores, since adoption occurs at the city level. In short, consideration of the spatial distribution of the posting policies is necessary to explain the results of the analysis in this paper.

2.3.1 Robustness

Table 2.6 provides several results supporting the robustness of the distance effect on hygiene. First, using the same specification as in equation (2.2) but replacing MD1 accordingly, these results demonstrate that distance to a mandatory posting restaurant significantly affects voluntary restaurant hygiene scores under several refinements of distance. For example, the analysis leverages data on the food type a restaurant serves and the actual grade being posted by mandatory posting restaurants to construct alternative distance measures. Second, estimates of equation (2.2) with added controls for market density confirm that the effect of distance to a mandatory posting restaurant on voluntary restaurants is not simply a product of market density.

Column (1) in Table 2.6 reproduces the main result from Table 2.5. All regressions in this section use standard errors corrected for spatial dependence, as discussed above. Columns (2)-(4) of Table 2.6 each provide a different specification of the minimum distance to a mandatory posting restaurant, and demonstrate that the significance of the proximity effect is robust and the estimate of the effect is relatively stable. Instead of MD1, Column (2) uses the minimum distance to a mandatory restaurant that is posting on the latest date that a city adopted mandatory posting before the date of the observed inspection ("posting at last treatment"). Relative to MD1, the distance variable in Column (2) will not include restaurant openings and closures since the last adoption of mandatory by a city. Column (3) instead uses the minimum distance to a mandatory posting restaurant that serves the same food type as the restaurant in the observed inspection. The coefficient on the distance variable in Column (3) is smaller, perhaps because restaurants serving the same food type tend to provide the same level of hygiene. Finally, Column (4) uses the minimum distance to a mandatory posting restaurant that is posting an A grade. The estimate on distance to mandatory posting restaurants posting an "A" is larger than the estimate on the distance to any mandatory posting restaurant, implying that being closer to a mandatory restaurant posting an "A" leads voluntary restaurants to improve grades slightly more. Note that the correlation between MD1 and minimum distance to a mandatory posting restaurant with an "A" grade is about 96%, making it difficult to determine whether simply posting or the grade itself induces the effect.

The data do support the argument that posting, and not simply the policy alone, is important for the distance effect, as demonstrated in Column (5). This regression uses the minimum distance to any restaurant in any city that has already adopted the mandatory policy. Unlike previous regressions which use distances mandatory restaurants actually posting at the time of the observation, this measure includes restaurants in mandatory areas that have not been inspected since the adoption of the grade card policy, and therefore do not have a grade card to display. While the point estimate of this variable is consistent with previous distance measures, increased variation leads this variable lose significance. This result suggests that voluntary restaurants are more consistently sensitive to mandatory restaurants that are posting.

Columns (6) and (7) of Table 2.6 demonstrate that MD1 is not simply a proxy for market density. For each restaurant inspected, column (6) controls for the minimum distance to any other restaurant (regardless of policy), separately before and after the grade card policy. Column (7) controls for the number of restaurants within one mile of the inspected restaurant (rescaled by 0.01), again for inspected restaurants under any policy. The significance of the minimum distance variable is robust and the point estimates are relatively stable after controlling for market density in both Columns (6) and (7). Perhaps as expected, the number of restaurants within one mile is significantly related to hygiene improvements for inspections of both mandatory and voluntary restaurants, while it is not related to hygiene before grade cards are introduced. The estimates for this market density variable are not significantly different between inspections of mandatory versus voluntary restaurants. Note that the market density seems to explain quite a bit of variation in the hygiene responses to the grade card policy, since the *mandatory* and *voluntary* dummies lose significance after these controls are included.

2.4 Conclusion

This paper explores spillover effects of a policy requiring restaurants to disclose quality information to consumers. In particular, the analysis here examines the introduction of the Los Angeles County hygiene grade card policy, under which a subset of restaurants were required to clearly display grade cards, while other restaurants were given grade cards but not required to display them. Previous research demonstrated that restaurants under the mandatory posting policy improved hygiene on average (Jin and Leslie (2003)). To determine the effect of the mandatory posting policy on other restaurants, this paper utilizes a panel dataset combining restaurant hygiene and new geographic data. The key result is that the mandatory hygiene disclosure policy had a spillover effect on the hygiene of nearby restaurants under a voluntary posting policy.

This analysis contributes to the existing literature by providing direct evidence that re-

Table 2.6: Additional Analyses of Effect of Proximity to Mandatory on Hygiene Scores

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
mandatory	3.069***	3.783***	3.437***	3.145***	2.878**	3.108***	0.442
	(0.880)	(0.916)	(0.892)	(0.878)	(0.883)	(0.883)	(0.831)
voluntary	3.043***	3.807***	3.577***	3.174***	2.743***	3.080***	1.269
	(0.744)	(0.819)	(0.768)	(0.749)	(0.796)	(0.746)	(0.804)
vol*MD1	-0.244*					-0.245*	-0.295**
	(0.102)					(0.102)	(0.091)
vol*(min to mand, posting at last treatment)		-0.223***					
		(0.0526)					
vol*(min to mand, same type, posting)			-0.196***				
			(0.0383)				
vol*(min to mand, posting "A")				-0.264**			
				(0.098)			
vol*(min to mand)					-0.204		
					(0.200)		
min to rest, after grade cards						-0.478	
						(0.725)	
min to rest, before grade cards						0.853	
						(0.572)	
vol*(no. rests w/in 1 mi)/100							0.652*
							(0.271)
mand*(no. rests w/in 1 mi)/100							0.752**
							(0.261)
invol*(no. rests w/in 1 mi)/100							-0.383
							(0.072)
Observations	163890	163890	163890	163890	163890	163890	163890
adj. R^2	0.506	0.507	0.507	0.506	0.506	0.506	0.510

Standard errors corrected for spatial dependence in parentheses.

All regressions include restaurant, time FEs, and dummy for second change in inspection criteria.

† $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Column (1) replicates Column (4) of Table 2.5. Results show the effect of proximity is robust to using various different minimum distance measures (Columns (2)-(4)). Column (5) shows the minimum distance to any restaurant within a mandatory area (including those that have not gotten grade cards yet) is not significant, implying posting matters. The significance of MD1 is robust to controls for the minimum distance to any posting restaurant and the minimum distance to any restaurant, both before and after grade cards (Column (6)); and to controls for the number of restaurants within one mile (Columns (7)).

gional quality-disclosure regulations affect non-targeted firms. The spillover result expands upon the analysis of JL, who noted that voluntary restaurants might improve scores due to either informational unraveling or anticipation of a change to mandatory. It is possible that the spatial spillover is simply a mechanism by which the unraveling or anticipation occurred. For example, observing the mandatory policy nearby may have caused restaurant managers to place a higher probability on an impending change to mandatory. On the other hand, the spillover may be evidence of alternative mechanisms driving voluntary restaurant hygiene improvement. As discussed in the Introduction, previous work suggests that distance may enhance both competitive pressures and cooperative opportunities between firms. Future research measuring the importance of each possible mechanism would provide a useful understanding of inter-firm behavior.

Regardless of the specific mechanism, increases in voluntary restaurant hygiene scores were at least partially achieved in response to the mandatory policy. In particular, mandatory posting for a given restaurant caused nearby restaurants under voluntary posting to improve hygiene, and the magnitude of this hygiene spillover effect declines for restaurants located further away. Simple robustness tests demonstrate (1) voluntary restaurants responded significantly to restaurants in mandatory areas that were actually posting at the time of the voluntary restaurant's inspection, and (2) the spatial effect is not explained by market density.

Bibliography

Berry, S., J. Levinsohn, and A. Pakes, 1995, "Automobile Prices in Market Equilibrium," *Econometrica*, 63(4), 841–890.

Boatwright, P., and J. C. Nunes, 2001, "Reducing assortment: An attribute-based approach," *The Journal of Marketing*, pp. 50–63.

Borle, S., P. Boatwright, J. B. Kadane, J. C. Nunes, and S. Galit, 2005, "The effect of product assortment changes on customer retention," *Marketing science*, 24(4), 616–622.

Cameron, A. C., J. B. Gelbach, and D. L. Miller, 2006, "Robust Inference with Multi-way Clustering," .

Chernev, A., 2003a, "Product assortment and individual decision processes.," *Journal of Personality and Social Psychology*, 85(1), 151.

——— , 2003b, "When more is less and less is more: The role of ideal point availability and assortment in consumer choice," *Journal of Consumer Research*, 30(2), 170–183.

Conley, T. G., 1999, "GMM estimation with cross sectional dependence," *Journal of Econometrics*, 92(1), 1–45.

Datta, S., and K. Sudhir, 2010, "Does Reducing Spatial Differentiation Increase Product Differentiation? Effects of Zoning on Retail Entry and Format Variety," *SSRN eLibrary*.

Davis, P., 2006, "Spatial competition in retail markets: movie theaters," *The RAND Journal of Economics*, 37(4), 964–982.

Dhar, R., 1997, "Consumer preference for a no-choice option," *Journal of Consumer Research*, 24(2), 215–231.

Draganska, M., and D. C. Jain, 2005, "Product-Line Length as a Competitive Tool," *Journal of Economics & Management Strategy*, 14(1), 1–28.

Fujita, M., and J.-F. Thisse, 2002, *Economics of Agglomeration*. Cambridge University Press.

Gabszewicz, J. J., and J.-F. Thisse, 1979, "Price competition, quality and income disparities," *Journal of Economic Theory*, 20(3), 340 – 359.

Hotelling, H., 1929, "Stability in Competition," *The Economic Journal*, 39(153), pp. 41–57.

Iyengar, S. S., and M. R. Lepper, 2000, "When choice is demotivating: Can one desire too much of a good thing?," *Journal of personality and social psychology*, 79(6), 995.

Jin, G. Z., and P. Leslie, 2003, "The Effect of Information on Product Quality: Evidence from Restaurant Hygiene Grade Cards," *The Quarterly Journal of Economics*, 118(2), 409–451.

Kahn, B. E., 1998, "Dynamic relationships with customers: high-variety strategies," *Journal* *of the Academy of Marketing Science*, 26(1), 45–53.

Kennedy, P., 1983, "PRACTITIONER'S CORNER*: Logarithmic Dependent Variables and Prediction Bias," *Oxford Bulletin of Economics and Statistics*, 45(4), 389–392.

Kuksov, D., and J. M. Villas-Boas, 2010, "When more alternatives lead to less choice," *Marketing Science*, 29(3), 507–524.

Mai, C. C., and S. K. Peng, 1999, "Cooperation vs. competition in a spatial model," *Regional Science and Urban Economics*, 29(4), 463 – 472.

Mogilner, C., T. Rudnick, and S. S. Iyengar, 2008, "The mere categorization effect: How the presence of categories increases choosers perceptions of assortment variety and outcome satisfaction," *Journal of Consumer Research*, 35(2), 202–215.

Neven, D., and J. F. Thisse, 1990, "On quality and variety competition," in *Economic Decision-Making: Games, Econometrics, and Optimisation; Contributions in Honour of Jacques H. Dreze*, ed. by J. J. Gabszewicz, J. F. Richard, and L. A. Wolsey. Elsevier Science Publishers, North Holland, Amseterdam, pp. 175–199.

Reutskaja, E., and R. M. Hogarth, 2009, "Satisfaction in choice as a function of the number of alternatives: When goods satiate," *Psychology & Marketing*, 26(3), 197–203.

Salop, S. C., 1979, "Monopolistic Competition with Outside Goods," *The Bell Journal of Economics*, 10(1), pp. 141–156.

Samuelson, P. A., 1952, "Spatial Price Equilibrium and Linear Programming," *The American Economic Review*, 42(3), pp. 283–303.

Sarver, T., 2008, "Anticipating regret: Why fewer options may be better," *Econometrica*, 76(2), 263–305.

Saxenian, A., 1994, *Regional Advantage*. Harvard University Press, Cambridge, UK.

Scheibehenne, B., 2008, "The effect of having too much choice," *unpublished dissertation, Humboldt University, Berlin, Germany.*

Scheibehenne, B., R. Greifeneder, and P. M. Todd, 2010, "Can There Ever Be Too Many Options? A Meta-Analytic Review of Choice Overload," *Journal of Consumer Research*, 37(3), 409–425.

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

Shah, A. M., and G. Wolford, 2007, "Buying behavior as a function of parametric variation of number of choices," *Psychological Science*, 18(5), 369–370.

Shaked, A., and J. Sutton, 1982, "Relaxing Price Competition Through Product Differentiation," *The Review of Economic Studies*, 49(1), pp. 3–13.

Sloot, L. M., D. Fok, and P. C. Verhoef, 2006, "The short-and long-term impact of an assortment reduction on category sales," *Journal of Marketing Research*, pp. 536–548.

Zeelenberg, M., W. W. Van Dijk, A. S. Manstead, and J. vanr de Pligt, 2000, "On bad decisions and disconfirmed expectancies: The psychology of regret and disappointment," *Cognition & Emotion*, 14(4), 521–541.