How Does Consumer Voice Respond to Antitrust Policy? Evidence from Supermarket Divestitures*

Daniel Hosken
Federal Trade Commission
d_hosken@yahoo.com

Frank Pinter
Federal Trade Commission
fpinter@ftc.gov

Devesh Raval
Federal Trade Commission
draval@ftc.gov

July 8, 2025

Abstract

Little is known about how consumer voice can inform antitrust policy. We focus on the antitrust remedy of divestitures intended to mitigate harm from horizontal mergers. We examine how Yelp reviews, our measure of voice, react to the divestiture of 249 grocery stores in two large supermarket mergers. Consumer reviews increase dramatically following divestitures. The share of negative reviews for divested stores rises, driven by reviews concerning prices, but user composition remains unchanged. Finally, negative reviews do not increase after voluntary asset sales, pointing to conflicts of interest between the government and merging parties as an explanation for our findings.

Keywords: divestitures, antitrust, reviews, mergers

*The views expressed in this article are those of the authors. They do not necessarily represent those of the Federal Trade Commission or any of its Commissioners. We thank Carl Bialik and Luther Lowe for providing us our dataset of Yelp reviews, as well as Angelina Tomseth, Joy Bhattacharya, Yara Chami, and Arushi Ghosh for excellent research assistance. We also thank Paul Ellickson, Ron Goettler, Gloria Sheu, Mo Xiao, and Brett Wendling for their comments, and Daniel Chaves, Oscar Jara, Zach Levin, and Ben Leyden for their discussions of this paper.

1 Introduction

Hirschman (1972) argued that consumers have two approaches when confronted with problems in the marketplace: voice or exit. Antitrust economics has traditionally focused on the exit margin. For example, the U.S. Department of Justice (DOJ)-Federal Trade Commission (FTC) Horizontal Merger Guidelines highlight using the diversion ratio – a measure of the share of consumers that switch to a competitor after a small price increase – to evaluate a merger between potential rivals. An important weakness of this approach, however, is that it does not directly measure how competition changes affect inframarginal consumers, including through adjustments in quality.

Policymakers are thus interested in complementing existing tools by using consumer voice to learn about competition problems and assess competition policy.[1] For example, in April 2024, the DOJ and FTC launched `HealthyCompetition.gov`, a portal for public reporting of anticompetitive practices in the health care sector, in an analogue to existing consumer protection complaint portals. In October 2024, the FTC launched a separate complaint portal for complaints about mergers and antitrust issues more generally.[2] The FTC has also begun a series of Open Commission Meetings where the public can listen to deliberations on antitrust policy and voice their complaints directly to the Commissioners deciding on such policy.[3]

We examine whether consumer voice can inform competition policy in the context of divestitures. Divestitures are a major antitrust remedy that allow firms to merge if they agree to sell off assets in areas with significant overlap between the merging parties. In the grocery industry, the FTC has required over 600 grocery stores divested to assuage competition concerns since 1990. In our analysis, we focus on two supermarket mergers, the 2015 Albertsons/Safeway and 2016 Ahold/Delhaize transactions, which the FTC declined to block after requiring that 249 grocery stores were divested.

[1]Consumer voice in the form of complaints was a much more important source of information for antitrust enforcers before the Hart Scott Rodino (HSR) Act of 1976 required firms to pre-notify the agencies of impending mergers. See, for example, Weaver (1977) for how antitrust investigations worked before the HSR Act.

[2]The FTC's antitrust complaint portal is located at https://www.ftc.gov/enforcement/submit-merger-antitrust-comment. The following press releases detail the launch of HealthyCompetition.gov and the FTC merger and antitrust complaint portals: https://www.ftc.gov/news-events/news/press-releases/2024/04/federal-agencies-launch-portal-public-reporting-anticompetitive-practices-health-care-sector and https://www.ftc.gov/news-events/news/press-releases/2024/10/ftc-finalizes-changes-premerger-notification-form.

[3]See https://www.ftc.gov/news-events/events/open-meetings for videos of previous meetings; the FTC states that "Open Commission Meetings are intended to open the work of the Commission to the public."

Several high-profile failures have raised questions on how such well divestitures work; our data includes one such failure in the 146 stores divested to Haggen in the Albertsons/Safeway transaction. The Haggen chain went bankrupt several months after the divestiture, after which all stores divested to Haggen closed temporarily and some permanently. We can thus compare the strength of consumer's voice in a major, clear failure to other divestiture events – stores divested to other buyers in Albertsons/Safeway (22 stores) and stores divested in Ahold/Delhaize (81 stores).

We measure consumer voice using the universe of Yelp supermarket reviews, which include the review rating, review text, and characteristics of the reviewing consumer. The review data allows us to examine whether consumers are more likely to exercise their voice through reviews after a divestiture, whether post-divestiture reviews were positive or negative, and whether consumers were concerned about prices or quality after the divestiture.

Our empirical strategy is a difference in differences regression. The divested stores are our treated group, which we compare to control stores located in the same state, but which did not belong to the merging parties and were not located in the same geographic region. Thus, our control group should not include any stores directly affected by the merger, or any stores competing with the divested stores.

Consumers do express their voice in response to divestitures. We find massive increases in consumer reviews post-divestiture, with a 280% increase within four quarters of the divestiture. We continue to find 70% more reviews relative to baseline on these stores one to two years post-divestiture. Despite clear differences in long-run outcomes – the Haggen divestiture led to store failures, while the other divestitures generally did not – we find similar increases in reviews for both sets of stores.

This outpouring of consumer voice is mostly due to negative reviews. We measure negative reviews based both on the star rating as well as machine learning sentiment scores of the review text, and find that the share of negative reviews increases by 40% within four quarters of the divestiture. For the Haggen stores, reviews improve after the Haggen stores are resold after Haggen's bankruptcy. For stores not sold to Haggen, we find a persistent increase in negative reviews over the two years post-divestiture. Several reviews mention the FTC specifically when criticizing the divestiture, indicating that some consumers were aware that the divestitures were due to competition policy.

To get a better sense of consumer concerns, we use machine learning zero-shot classification methods to assess the topics that consumers address in their reviews. Zero shot classification uses a large language model to predict whether a consumer is referencing a specific topic, and has been found to be very effective in other contexts (e.g., Agarwal et al. (2024) on radiology diagnosis). We examine three topics – prices, products, and customer service – and allow reviews to touch on multiple topics. We find that the large increase in negative reviews is driven by an 103% increase in complaints about prices as a share of negative reviews within four quarters of the divestiture.

We next examine several measures of differences in the composition of users, including whether the reviewer is a "new" supermarket reviewer, whether the reviewer has ever been Yelp Elite (a frequent Yelp reviewer) in our sample period, and whether the review was flagged by Yelp as not recommended, which is an indicator of whether the review is fake (Luca and Zervas, 2016). Surprisingly, we find no differences in the composition of reviewers for divested stores post-divestiture across all of these measures. In addition, we find little change in either the length or readability of reviews post-divestiture.

With government mandated divestitures, unlike with most voluntary asset sales, the seller of the assets will compete with the purchaser of the divested assets. This anticipated competition may result in conflicts of interest where the merged firm has incentives to take actions to diminish the competitive significance of the divested assets. For example, it may attempt to sell divested assets to a weak firm, handicap the buyer when transitioning stores, or only sell weak assets. Such strategic motives have been alleged in court filings by Haggen in its lawsuit against Albertsons after its failure and Albertsons in its lawsuit against Kroger after the Albertsons/Kroger merger was blocked.

We evaluate whether strategic motives may have been a factor in the increased negative reviews following divestitures by analyzing how reviews responded following two voluntary asset sales in our data. First, we examine how reviews changed at the stores acquired in the Albertsons/Safeway and Ahold/Delhaize mergers. In these transactions, the merging firms faced a management challenge from adapting to a significant increase in scale, but consumers could continue shopping at stores operating the same retail banners. We find little evidence of changes in reviews for these acquired stores, suggesting that the voluntary change of assets did not materially affect consumers in the short run.

Second, we examine stores sold as part of the bankruptcy of the A&P chain. While supermarket assets were transferred to other firms operating different retail banners, A&P did not have incentives to lessen the competitive performance of these stores as it was exiting the industry. Following the store sales, we find increases in reviews, although they are smaller than for the divestiture events, and decreases in the share of negative reviews post-sale. Here, the sold supermarkets were likely operated by better owners than the bankrupt A&P under different supermarket banners, which consumers noticed and liked.

Our work adds to the literature that examines the efficacy of merger policy by estimating the competitive effects of consummated mergers (Ashenfelter, Hosken and Weinberg, 2014; Asker and Nocke, 2021).[4] The overwhelming majority of studies in this literature estimate the effect of mergers on price or output (e.g., Ashenfelter, Hosken and Weinberg (2013) for home appliances and Ashenfelter, Hosken and Weinberg (2015) for beer).

Our paper makes a unique contribution to this literature by explicitly measuring consumers' assessment of consummated mergers. Within the merger retrospective literature, we add to a small but growing set of papers evaluating the effectiveness of divestitures as an antitrust remedy. Both government agencies (Federal Trade Commission, 1999, 2017; General Accountability Office, 2002) and researchers (Argentesi et al., 2021; Brown, Eckert and Shaffer, 2023; Chen et al., 2022; Friberg and Romahn, 2015; Lagos, 2018; Osinski and Sandford, 2021; Soetevent, Haan and Heijnen, 2014; Tenn and Yun, 2011; Wang et al., 2023) have examined divestitures in several contexts.

In addition, scholars in economics and marketing have long focused on how consumer voice affects markets. In his seminal work, Hirschman (1972) thought that consumers were more likely to respond with voice in concentrated markets. Both Beard, Macher and Mayo (2015) and Gans, Goldfarb and Lederman (2021) assess this prediction empirically and find greater voice in more concentrated markets. Our paper contrasts with this work as it examines a change in market structure from competition policy rather than pre-existing differences in market structure.

More broadly, scholars have repeatedly shown that consumer voice affects demand (Anderson and Magruder, 2012; Luca, 2011; Lewis and Zervas, 2020). Measures of voice such as reviews and complaints have potential biases (He, Hollenbeck and Proserpio, 2022; Luca and Zervas, 2016;

[4]Beyond the cited literature reviews, the FTC maintains a bibliography of published merger retrospectives on its website: `https://www.ftc.gov/policy/studies/merger-retrospective-program/bibliography`.

Mayzlin, Dover and Chevalier, 2014; Raval, 2020), although there are ways to amplify consumers' voice and correct potential biases (Fradkin, Grewal and Holtz, 2021; Fradkin and Holtz, 2023; Grosz and Raval, 2025; Nosko and Tadelis, 2015). Finally, the disclosure of quality can have supply side effects on several dimensions, including incentivising firms to improve quality (Jin and Leslie, 2003; Kolstad, 2013) or "game" the system (Dranove et al., 2003), increasing supplier congestion (Chartock, 2023), and altering firm advertising and branding strategy (Hollenbeck, 2018; Hollenbeck, Moorthy and Proserpio, 2019).

2 Background

2.1 Divestiture Policy

The primary goal of prospective merger policy is to ensure that mergers do not harm competition. However, enforcers often have to trade-off potential efficiencies from a merger in some markets against competitive harms in other markets (Williamson, 1968). One way to balance this tradeoff is through a settlement or remedy. As FTC Chairman Ferguson remarks in a statement on a recent mandated divestiture[5]:

> If, for example, a merger has anticompetitive and procompetitive features, a lawsuit blocking the entire merger would protect the public from the merger's anticompetitive effects but would also deny the public the benefit of the procompetitive effects. A settlement that successfully prevents the merger's anticompetitive features can strike a balance that permits the procompetitive aspects to proceed. Assuming the settlement would in fact prevent the merger's anticompetitive effects, the settlement would fully protect the competitive process while also promoting the innovation and growth that the remainder of the merger might foment.

The most common such remedy is a divestiture. With a divestiture remedy, an antitrust agency requires some assets of the merging parties sold to third parties to maintain competition in markets with potential competitive harm. Divestiture remedies require the merging firms to identify a set of assets to divest and a buyer for those assets. In addition, for such a remedy to succeed in

[5] See https://www.ftc.gov/system/files/ftc_gov/pdf/synopsys-ansys-ferguson-statement-joined-by-holyoak-meador.pdf.

restoring competition, the new buyer would have to become an effective competitor and prevent price increases or quality declines. Chairman Ferguson provides the following considerations for a divestiture to be successful in his statement:

> Nor should the Commission ordinarily accept a structural remedy unless it involves the sale of a standalone or discrete business, or something very close to it, along with all tangible and intangible assets necessary (1) to make that line of business viable, (2) to give the divestiture buyer the incentive and ability to compete vigorously against the merged firm, and (3) to eliminate to the to the extent possible any ongoing entanglements between the divested business and the merged firm. The Commission must also be confident that the divestiture buyer has the resources and experience necessary to make that standalone business competitive in the market.

Identifying the set of assets and the buyer for a divestiture is challenging. First, the merging firms have an inherent agency problem constructing the divestiture. The antitrust agency wants the merging firms to divest valuable assets to a strong competitor to maintain competition, while the merging parties would prefer to divest low quality assets to a weak competitor. Second, the agency has an informational disadvantage relative to the merging firms. The merging firms have deep knowledge of the industry and the quality of the assets they choose to divest and hold. Agency staff must rely on third parties – potential buyers and industry experts – with potentially different interests than the agency to evaluate the quality of the divestiture package.

The evidence is mixed on how successful divestitures have been in maintaining competition. While the FTC's 1999 and 2017 Divestiture studies (Federal Trade Commission, 1999, 2017) concluded that the FTC's divestures were successful overall, both reports also noted that a significant fraction of divestitures failed.[6] For example, the 1999 study found that the buyer was no longer actively participating in the market in about 25% of divestitures examined. The 2017 study concluded that more than 80% of orders across 50 examined divestitures maintained or restored competition, suggesting a significant fraction of divestitures failed to maintain competition.

Researchers have begun to evaluate the efficacy of divestitures and, like the FTC studies, find

[6]The 1999 FTC Divestiture study states, "The Study supports the view that divestitures have been successful remedies for anticompetitive mergers." (page 8). The 2017 FTC Divestiture Study states, "The Study supports the view that divestitures have been successful remedies for anticompetitive mergers." (page 1).

decidedly mixed results. Three studies examine divestitures in retail markets broadly similar to those in our study. Argentesi et al. (2021) studies a merger of Dutch supermarket chains and finds that divestitures were successful in maintaining competition post-merger. Lagos (2018) examines divestitures of gasoline stations in Chile and concludes that divestures were only effective where the divested stations were located within one kilometer of one of the merging parties' stations. Finally, Osinski and Sandford (2021) study a divestiture following a casino merger and find that the divested casino, but not the merged firm's casino, performed worse post-merger.

2.2 Albertsons/Safeway and Ahold/Delhaize Mergers

In this article, we examine two large supermarket mergers – Albertsons/Safeway as well as Ahold/Delhaize – in which the government mandated a set of divestitures. Pre-merger, Albertsons and Safeway operated 630 and 1,332 stores, and Ahold and Delhaize operated 710 and 1,291 stores. While these chains were large, they did not directly compete in most of the markets in which they operated. The FTC required 168 stores divested in the Albertsons/Safeway merger and 81 stores in the Ahold/Delhaize merger to preserve competition in a large number of relatively narrow geographic markets.[7]

In the divestitures for these mergers, the FTC did not require the merging firms to sell off an entire business unit, e.g., all of the stores in the business unit servicing the area where the firms competed, the distribution centers servicing those stores, and the use of one of the retail brand names.[8] Instead, the divestitures included stores owned by both of the merging firms even in adjacent markets. For example, of the 83 stores divested in California, 28 were previously operated by Safeway and 55 by Albertsons.[9]

[7] For example, in its aid for public comment, the FTC stated that "The relevant geographic markets in which to analyze the effects of the Acquisition are areas that range from a two- to ten-mile radius around each of the Respondents' supermarkets, depending on factors such as population density, traffic patterns, and unique characteristics of each market." See https://www.ftc.gov/system/files/documents/cases/150127cereberusfrn.pdf. Some of these markets were very close to one another. For example, the FTC identified one market as Santa Barbara, California and another as Santa Barbara/Goleta, California, two regions about 9 miles apart, among 130 distinct geographic markets that would be adversely affected by the merger.

[8] The decision to accept this divestiture deviated from the FTC's stated preference to prefer divestitures of standalone businesses, as in Chairman Ferguson's statement above. For example, the FTC's 2017 Divestiture study stated "The study found that all divestitures of ongoing businesses succeeded, whether the divestiture was to an upfront buyer or a post-order buyer. This finding reinforces what the Commission and staff have long known: divestiture of an ongoing business, which includes all assets necessary for the buyer to begin operations immediately, maximizes the chances that the market will maintain the same level of competition post-divestiture." See page 21.

[9] See the FTC's Albertsons/Safeway Aid to Public Comment at: https://www.ftc.gov/system/files/documents/cases/150127cereberusfrn.pdf. For instance, the divesture required two Safeway stores to be divested

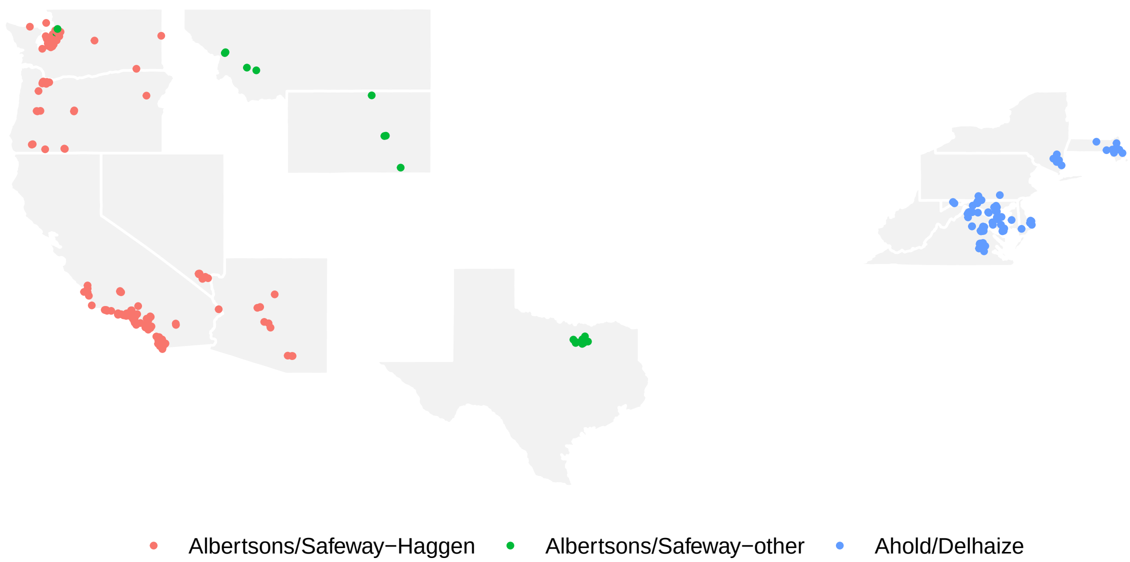

Figure 1: Map of divested stores in the Albertsons/Safeway and Ahold/Delhaize mergers.

Most of the firms purchasing divested stores were much smaller supermarket chains than the merging firms.[10] Empirical work examining productivity growth (Foster, Haltiwanger and Krizan, 2006) finds that large firms are much more likely to have high labor productivity in the retail trade sector, suggesting that small retailers obtaining divested stores may face difficulties in effectively competing with larger chain retailers. In addition, the divestiture did not require the merging firms to divest a retail brand. As a result, the owner of the divested asset had to enter the retail market with a new, unknown retail product that likely increased the riskiness of the divestiture.

One of the divestitures – stores in the Albertsons/Safeway transaction sold to the retailer Haggen (see Figure 1) – became a notorious failure. The divestiture transformed Haggen from a small chain that operated 18 stores in Washington and Oregon pre-merger to a regional West Coast chain by adding 146 stores from Albertsons/Safeway located in Arizona, California, Nevada, Washington,

in the Santa Barbara/Goleta market and two Albertsons stores in the nearby Santa Barbara market.

[10] While two of the divestiture buyers in the Ahold/Delhaize case were large firms, 70 of the divested stores were purchased by much smaller retailers operating less than 200 stores prior to the divestiture. Albertsons (with 2,200 stores) and Publix (with 1,100 stores) purchased 1 and 10 of the divested stores. Weis and Topps were regional supermarket chains that operated 165 and 163 stores prior to the divestiture, and purchased 38 and 6 stores respectively. The remaining purchasers were much smaller, operating 61 or fewer stores; Saubel's Market (with 3 stores) purchased one store, Big Y (with 61 stores) 7 stores, and SuperValu (a wholesaler operating some corporate-owned stores) 18 stores. See FTC Aid to Public Comment: https://www.ftc.gov/system/files/documents/cases/160722koninklijkeanalysis.pdf.

and Oregon. Thus, the transaction required Haggen to establish a retail presence in a large number of disparate regions far from its traditional markets, and quickly develop a network to supply these stores.

Haggen took over the divested stores between March and June of 2015, but very quickly began to have severe problems effectively operating its stores. It ultimately filed for bankruptcy in September of 2015 and began the process of closing all of its California stores, many of which were subsequently purchased by Albertsons. Because of the potential significance of the failed Haggen divestiture, we examine our findings on both samples limited to and excluding the Haggen outlets.

3 Data

Our primary data set is the universe of U.S. Yelp supermarket reviews written between October 2004 and July 2021. The data contains each review's star rating (1-5), text, characteristics of the reviewing consumer, and the address and brand affiliation of the reviewed supermarket. We have 1.9 million reviews associated with 77,347 grocery store and supermarket establishments across the United States. Of these reviews, 4% are updates, when a user revisits a prior review, and the remaining 96% are original reviews.

A key challenge for our study is constructing a unique store identifier. While the Yelp data contains a separate store identifier for each establishment, identified by an address and store name, Yelp sometimes creates a new store identifier for an establishment when a retail outlet changes names, such as due to a merger. Because the goal of our study is to measure the change in the quality of establishments before and after ownership changes, we create a time invariant identifier corresponding to each "store" defined by a unique retail address. After grouping the establishments by physical location, our sample consists of 71,268 unique stores.[11]

For our analyses, we define the unit of observation as a store/quarter; for example, when studying the count of reviews, we aggregate the number of reviews a store receives to the quarter. This aggregation makes sense for two reasons. First, consumers review supermarkets at a relatively

[11] In the Yelp data, multiple establishments can be observed at what is effectively the same address over time for a number of reasons. First, there is sometimes simple duplication (common in user-generated data), user-generated distinctions (for example, separating pharmacy and food departments at a supermarket), or the turnover of stores and banners over time. To address these issues we simply collapse all establishments at an address into one unit. To overcome inconsistencies in address strings, we use a combination of text cleaning with R and geocoding with ArcGIS to map establishments in the raw Yelp data to stores.

low frequency, so many store-quarter combinations have no or a small number of reviews. Second, we do not know the exact dates divested stores reopened with their new ownership and banner.[12]

3.1 Summary statistics

Time trends. As shown in Table 1, reviews grew in the early 2010s before plateauing. The number of reviews plateaued around 280,800 in 2015; the number of stores reviewed each year plateaued around 35,000 in 2018. Measuring negative reviews as those with 1 or 2 stars, ratings generally became more negative over time. This decline could indicate compositional changes in the makeup of Yelp reviewers or changing norms about the purpose of Yelp reviews.

Table 1: Time trends in review counts and ratings, 2013–2019.

	2013	2014	2015	2016	2017	2018	2019
# stores w/ reviews	27,559	29,885	32,392	33,762	35,070	36,857	36,344
# reviews	199,570	240,851	280,800	273,684	280,242	292,762	273,946
Negative share	24%	27%	30%	33%	36%	37%	41%
Mean star rating	3.54	3.50	3.42	3.37	3.30	3.27	3.16

Note: Reviews are assigned to stores by unique address. A small number of locations are dropped as invalid or outside the US. Negative reviews are 1- or 2-star reviews.

Number of reviews per store, 2013–19. Among stores that received any reviews, the median number of reviews per store is 5; the 80th percentile is 21 or slightly more than one review per quarter, and the mean is 24. The top 1% of stores account for 16% of the reviews; the top 20% of stores account for 73% of the reviews.

Regional variation in Yelp review density. Figure 2 shows the average number of reviews per location (2014–18) in each of the 48 contiguous states. California, Nevada, and Hawaii have the most (50, 53, and 60 reviews per location, respectively); conveniently, many of the stores divested to Haggen were in California. West Virginia, Arkansas, and Mississippi have the fewest (4, 5.2, and 5.3 reviews per location, respectively).

Divestiture specific reviews. We identified several Yelp reviews as specifically critiquing the FTC's divestiture policy. We excerpt several examples below:

[12]Because the divestitures associated with a given merger usually occurred in a single quarter, or across two consecutive quarters, this level of aggregation also mitigates issues due to staggered treatment.

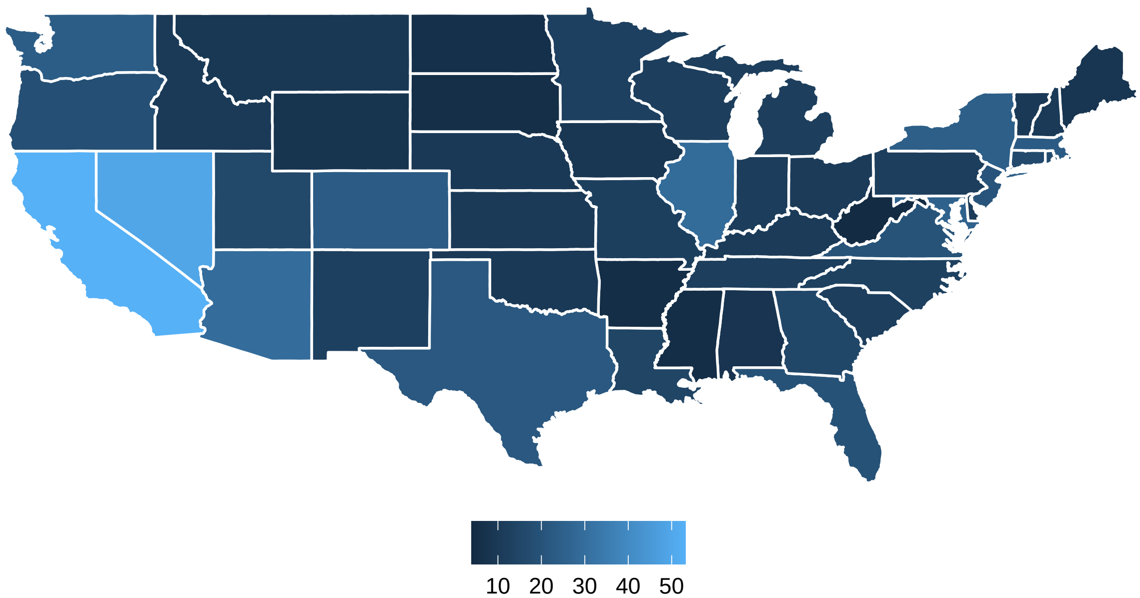

Figure 2: Average number of reviews per location, 2014–18, across states.

> The FTC took Vons away from us, a store that had a great organic brand, good specialty items, AND competitive prices.

> Thanks FTC (Obama) for making them sell our local stores to an upscale expensive grocer. Is there anything not getting worse under this administration?

> The morons at the FTC made Albertson's divest itself of the two Von's in the area out of fear that it would hold a monopoly. Von's was replaced by the overpriced Haggen's which screwed the workers and closed in six months. Now, we have Gelson's in its place and I fear it will meet the same fate.

4 Empirical approach

Our main empirical strategy is a difference in differences regression. We consider the divested stores as treated and compare divested stores to control stores that were not directly affected by the merger or competing with the divested stores. To interpret our results as causal effects, we assume that had the merger not occurred, trends in the review patterns of the control group would have been parallel to trends in the review patterns of the treatment group.

Table 2: Summary statistics on reviews, by store group (control/excluded/treatment) and time (pre/post merger).

		Control		Excluded		Treatment	
		Pre	Post	Pre	Post	Pre	Post
Ahold/Delhaize	Reviews	70,787	104,459	24,520	40,518	251	827
	Stores represented	6,406	8,028	2,571	3,045	49	67
	ZIP3s represented	149	153	123	129	25	29
	Avg. rating	3.48	3.25	3.33	3.13	3.16	2.74
	% negative	28%	37%	32%	41%	34%	51%
Albertsons/Safeway	Reviews	59,842	229,736	90,609	363,598	1,280	9,107
	Stores represented	5,870	9,210	7,116	10,036	155	165
	ZIP3s represented	112	112	132	147	54	56
	Avg. rating	3.62	3.35	3.49	3.18	3.06	2.82
	% negative	24%	35%	27%	40%	38%	50%

Note: Compiled using reviews from January 2013 to July 2021. In this table only, review dates are assigned to pre/post merger using a cutoff of January 1, 2015 for Albertsons/Safeway and July 1, 2016 for Ahold/Delhaize, shortly before the mergers closed; actual divestitures occurred in the months following. Reviews are assigned to stores by unique address. Stores are assigned to treatment, control, and excluded groups as described in the text.

Our treatment group includes stores divested to Haggen (146 stores) in the Albertsons/Safeway transaction, stores not divested to Haggen in Albertsons/Safeway (22 stores), and all stores divested in the Ahold/Delhaize transaction (81 stores). In sum, our treatment group includes 11,465 reviews across 242 unique addresses.[13]

Our excluded group contains stores that were located in the same geographic region (three-digit zip code) as any divested store, or that belonged to any of the merging parties. These stores may have been affected by the mergers or divestitures through changes in local competition or firm linkages.

Our control group includes stores located in the same state as a divested store, but not located in the same geographic region or belonging to the merging parties. The control group thus includes grocery stores and supermarkets of various sizes and formats in urban, suburban, and rural areas. In sum, our control group accounts for 459,899 reviews across 19,099 unique addresses. In Table 2, we provide the number of reviews, average rating, and stores for the treatment group, control group, and excluded group both before and after divestiture.

Our difference in differences regressions have the standard two-way fixed effects form. For

[13]We were unable to match 7 divested stores to the Yelp data.

store i in quarter t, we model the conditional expectation of a nonnegative outcome variable y_{it} (Wooldridge, 2023) as:

$$E[y_{it} \mid i, t] = \exp\left(\gamma_i^y + \delta_{t,\text{state}(i)}^y + \sum_{s \neq -1} \beta_s^y \cdot I(i \in \text{treated}) \cdot I(t - t_0(i) = s)\right), \tag{1}$$

where treated is the set of treated stores, $t_0(i)$ is the quarter of treatment, γ is a vector of unit fixed effects, and δ is a vector of state-quarter fixed effects. The treatment effect in percentage units is $\exp(\beta^y) - 1$. Our specifications vary in the choice of outcome variable y. For example, when we measure how frequently stores are reviewed in a quarter, y_{it} is the count of reviews for store i in quarter t.

In addition, for many outcome variables, we examine how the share of that variable changes post-divestiture. To do so, we report the difference in treatment effects for that variable relative to a baseline (such as relative to the number of reviews). For example, when we examine changes in the negative review share of overall reviews, we estimate the effect on negative reviews $\hat{\beta}^n$ and the effect on total reviews $\hat{\beta}^y$, and then compute the change in the share of negative reviews in percentage terms as $\exp(\hat{\beta}^n - \hat{\beta}^y)$.[14]

Throughout, we report standard errors clustered by the ZIP3 of the store, which allows for correlations between stores in the same local area due to local market shocks.[15]

5 Main Effects

In this section, we examine how the divestiture affected the number of reviews and the sentiment of those reviews. Our difference-in-differences estimates are stark: divested stores immediately experience multi-fold increases in reviews. The increase in review activity is concentrated in the quarters immediately following the divestitures, suggesting the consumers are exercising their voice in response to a change in product offerings. Most of the additional reviews post-divestiture are negative. Finally, we do not find changes in the composition of users reviewing supermarkets post-divestiture.

[14]Standard errors on the difference in $\hat{\beta}$ across specifications requires calculating the covariance across specifications. To do so, we stack the regressions with different outcome variables and estimate jointly.

[15]We examine alternative clustering methods in Appendix B.3 and find broadly similar standard errors. See Section 7 for a discussion.

Table 3: Summary of difference-in-differences estimates, aggregated to four-quarter time periods.

Time Period	No. Reviews	Share Negative Reviews
Quarters 0-3	275%	44%
	[225%, 333%]	[23%, 70%]
Quarters 4-7	69%	-27%
	[38%, 108%]	[-40%, -10%]
Pre-mean	0.88	0.45

Note: This table shows the results of the difference-in-differences regressions. Each column represents a different dependent variable; the table presents the estimated effect on the number of reviews and the effect on the *share* of negative reviews, computed as the difference between estimated regression coefficients. The Quarters 0-3 and 4-7 rows show the estimated dynamic treatment effects, both point estimates and 95% confidence intervals, transformed to percentage units. The 'Pre-mean' row shows either the mean of the dependent variable in the treatment group in the quarter before divestiture, or the ratio of means (mean dependent variable divided by mean of 'as share of' variable).

5.1 Review Frequency

We begin by examining how the Albertsons/Safeway and Ahold/Delhaize divestitures affected the frequency of reviews. Figure 3 shows the estimated percentage effect of divestiture on reviews, where all estimates are relative to the quarter prior to the divestiture. For the event study estimated using the entire sample (overall, in red), divestitures cause consumers to write more reviews, although this effect decays over time. The total effect over the first year (the quarter of divestiture and the three quarters following) is a massive 275% increase in reviews. In the second year, the effect is smaller but still positive at a 69% increase. For comparison, the average divested store had 0.88 reviews in the quarter before divestiture, as shown in the first column of Table 3.

Because so many of the divested stores in our sample went to a single buyer, Haggen, that failed in the year following the divestiture, we also examine how our estimates change with different cuts of the divestiture sample. We estimate the difference-in-difference model using only the stores divested to Haggen (blue), only the stores divested to Haggen that continued as supermarkets after Haggen's exit (green), and the non-Haggen sample (all of the Ahold/Delhaize stores, and the other Albertsons/Safeway divested stores, shown in purple).[16] The results of the event study are remarkably similar for all of the samples: consumers are much more likely to review stores in the year following the divestiture, with a smaller effect in the following year. While the confidence

[16] Following Haggen's failure, some of the formerly Haggen store locations were sold to firms operating outside of the supermarket industry. To explore if our findings are driven by relatively poor retail locations, we estimated the model limited to stores that remained supermarkets.

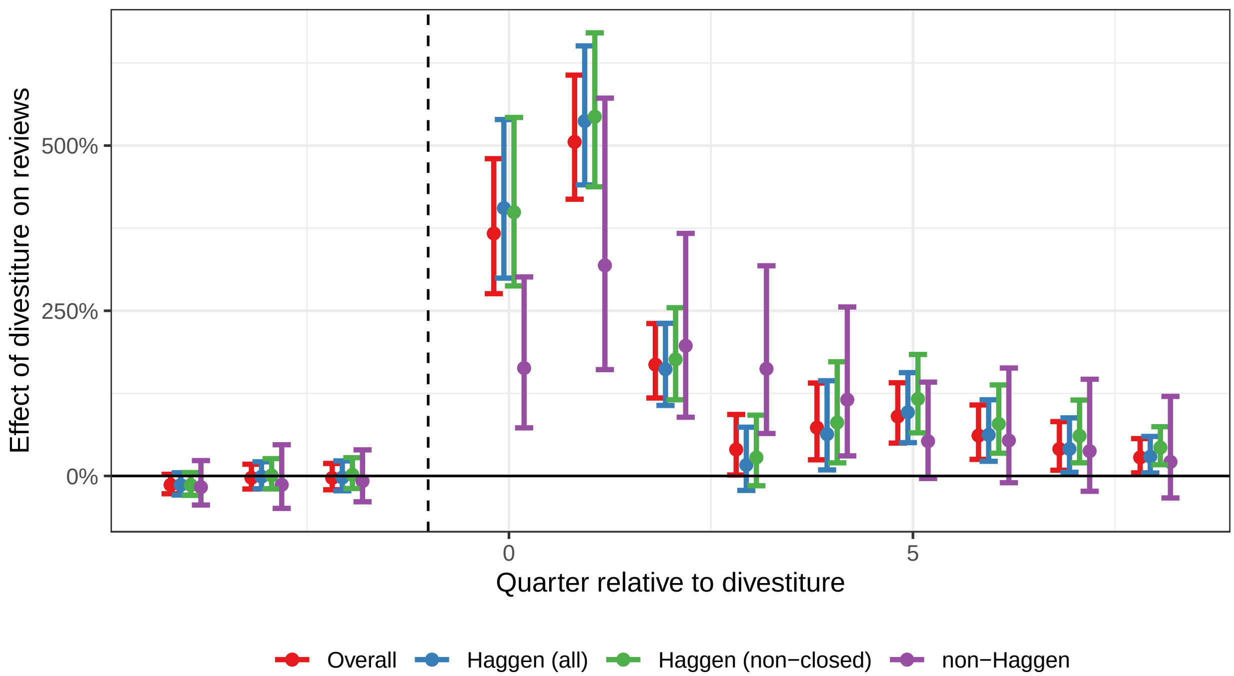

Figure 3: Effect of divestiture on number of reviews (event study plot)

intervals are much larger for the non-Haggen stores due to the lower baseline level of reviewing at these stores, the pattern and estimated magnitudes are very similar. Overall, as documented in Table 3, we find that the number of reviews increases by 275% in the first year and 69% in the second year.[17]

5.2 Review Sentiment

We next examine how the sentiment of reviews changed following the divestiture. We measure review sentiment using the star ratings accompanying reviews and consider 1- or 2- star reviews as negative reviews. Since the number of reviews increased so much, as documented above, we examine how the share of negative reviews changed following the divestiture.

To do so, we estimate $\hat{\beta}^n$, the difference in difference effect for negative reviews, and $\hat{\beta}^y$, the difference in difference effect for all reviews, using equation (1). We then report $\exp(\hat{\beta}^n - \hat{\beta}^y)$, the causal effect of the divestiture on the negative review share of overall reviews in percentage terms.

Figure 4 shows the percentage effect of divestiture on the negative share of reviews for each of

[17] Appendix C provides estimates for the first and second year for only the stores divested to Haggen, only the stores divested to Haggen that continued as supermarkets after Haggen's exit, and the non-Haggen sample.

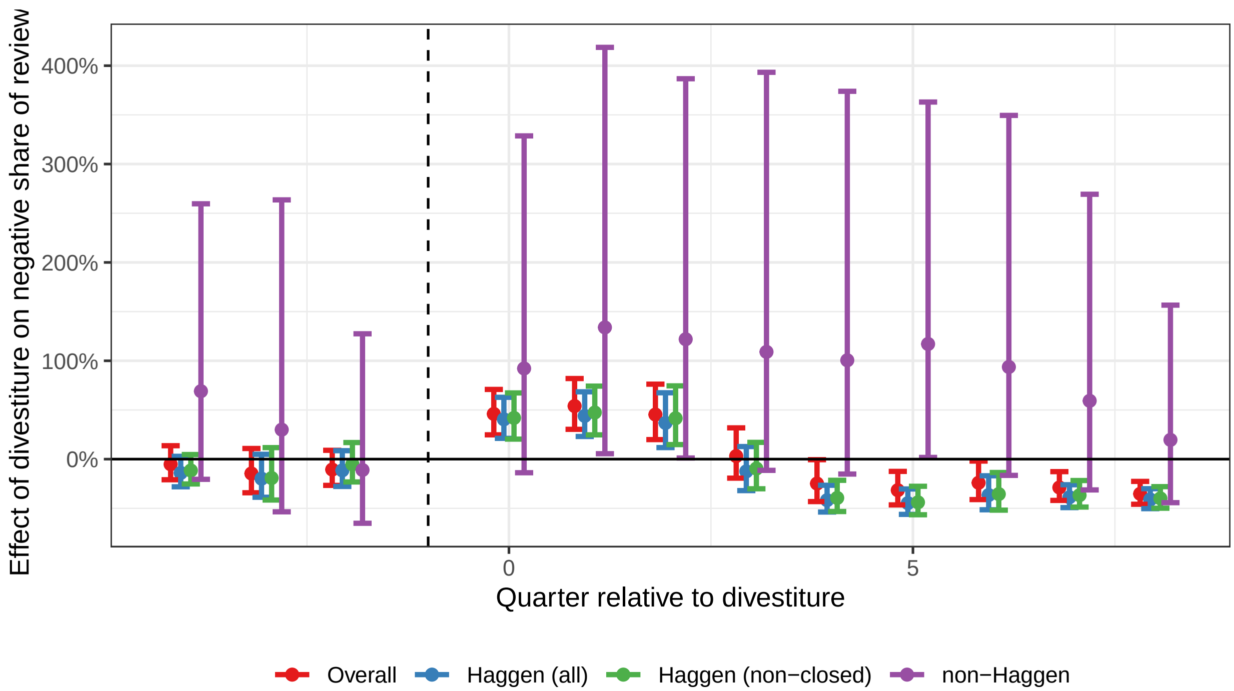

Figure 4: Effect of divestiture on negative share of reviews (event study plot)

the four estimation samples. Using the overall sample, the share of negative reviews increases 44% within the first year, as shown in the second column of Table 3.[18] We find a increase in the share of negative reviews for both the Haggen stores (37%) and non-Haggen stores (114%) in the first year post-divestiture.

In the second year post-divestiture, the share of negative reviews *decreases* by 27% relative to the counterfactual. However, we see a very different pattern in the second year for the Haggen and non-Haggen divested stores. For the Haggen stores, the fraction of negative reviews declines relative to the baseline by 40% in the second year. This decline suggests that the new non-Haggen operators of stores originally divested to Haggen likely improved the quality of these stores, or that consumers were happy to have an operating grocery store again. However, for the non-Haggen stores, which were still operated by the original divestiture buyers, the fraction of negative reviews increases in the second year by a similar magnitude, at 93%, to the first year following the divestiture. Reviewers of these stores continue to be more likely to leave negative reviews, likely because the decline in quality post-divestiture persists.

[18] For comparison, the divested stores' reviews were 45% negative, on average, in the quarter before divestiture; a 44% increase on this baseline would be a jump of 20 percentage points, to 65%.

5.3 Composition of Reviewers

One explanation for the large surge in reviews after stores are divested is that the types of reviewers who post reviews in Yelp also changes post-divestiture. In that case, the rise in the share of negative reviews might reflect different types of reviewers rather than a decline in store quality.

In order to assess potential changes in reviewer composition, we examine multiple measures of consumer characteristics in the Yelp data. For each measure, we report the effect on the share of that variable by differencing the effect for the measure from the effect for the count of all reviews, just as we did for the share of negative reviews in the previous section.

Our first measure is the number of reviewers that post their first supermarket review on Yelp, which can capture changes in users who do not normally review on Yelp. Second, the number of reviews flagged as "not recommended" (Luca and Zervas, 2016) allows us to examine whether more uncommon reviewers (such as first-time reviewers) whose reviews Yelp might flag as fake review more post-divestiture. Third, we use the number of reviewers that were never "Yelp Elite" to measure changes in reviews from frequent Yelp reviewers.

Surprisingly, we find small and insignificant effects of the divestiture on the share of reviews for all of these measures of user composition. As Table 4 shows, in the first year post-divestiture we estimate a rise of 7% for the share of new supermarket reviewers, 0% for the share of never Yelp Elite reviewers, and -13% for the share of reviews flagged as not recommended.

We continue to find insignificant changes to user composition if we use several alternative measures of user types: reviews by users who made no other Yelp reviews ("Only Yelp Review"), who were outside the top 10% of users by review count ("Not a Top 10% Reviewer"), or who made no other grocery store reviews ("Only Grocery Review").

We next examine the characteristics of the reviews themselves, and continue to find only small changes post-divestiture. We examine the length of the review and the readability of the review as measured by the Flesch reading ease score.[19] In the first year post-divestiture, we estimate an 8% increase in the share of 'readable' reviews as measured by reviews with an above-median Flesch

[19] The Flesch Reading Ease score assigns a text's readability a number between 1 (hardest) and 100 (easiest). The Flesch Reading Ease measure is defined as

$$206.835 - 1.015\left(\frac{words}{sentences}\right) - 84.6\left(\frac{syllables}{words}\right). \qquad (2)$$

score. Examining the length of reviews, we find a 15% increase in the share of reviews with an above-median number of characters and a 9% increase in the share of reviews with more than 280 characters, which is the maximum length of a tweet.

Table 4: User composition effects of divestitures, difference-in-differences estimates.

	Effect		
Type	Quarters 0-3	Quarters 4-7	Pre-mean
New Grocery Reviewer	7%	-1%	0.58
	[-5%, 20%]	[-11%, 10%]	
Not-Recommended Flag	-13%	-16%	0.16
	[-36%, 19%]	[-40%, 18%]	
Never Yelp Elite	0%	-4%	0.83
	[-5%, 5%]	[-10%, 2%]	
Only Yelp Review	50%	22%	0.03
	[-34%, 240%]	[-45%, 168%]	
Not a Top 10% Reviewer	1%	-4%	0.72
	[-8%, 10%]	[-12%, 5%]	
Only Grocery Review	11%	7%	0.41
	[-6%, 31%]	[-9%, 26%]	
Long Review (median)	15%	9%	0.50
	[1%, 32%]	[-8%, 28%]	
Long Review (tweet)	9%	5%	0.68
	[1%, 18%]	[-3%, 13%]	
Readable Review (median)	8%	9%	0.50
	[-4%, 21%]	[-2%, 22%]	

Note: This table shows the effect of divestiture on the dependent variable as a share of the count of reviews, computed as the difference between estimated regression coefficients. The Quarters 0-3 and 4-7 columns show the estimated dynamic treatment effects, both point estimates and 95% confidence intervals, transformed to percentage units. The 'Pre-mean' column shows the ratio of the mean of the dependent variable to the mean count of reviews, in the treatment group in the quarter before divestiture.

6 Review Topics

While the above analysis showed that consumers respond to divestitures through changes in reviews, it did not shed light on what consumers' concerns were. Thus, in this section, we use large language models to get a better sense of what consumers are reviewing by predicting the probability that any review is on specific topics.

We use machine learning zero-shot classification methods to predict the topics that consumers address in their reviews. We use the `bart-large-mnli` large language model downloaded from

Hugging Face to assign a vector of topics to each review.[20] The model assigns each review a score, between 0 and 1, for each topic; the scores are independent and need not add to one.

In an initial manual examination of reviews, we observed that consumers often left positive and negative reviews discussing the prices changed at stores, customer service, and products. Reviews on products often discussed the variety of products available in store as well as product quality. For this reason, we examine three topics – prices, products, and customer service – and allow reviews to touch on multiple topics. Figure 5 depicts histograms of our three topic scores across all the reviews relevant to our divestitures analysis. The scores span the full range from zero to one, giving us useful variation in the probability of each topic across reviews.

We further assess the performance of the LLM topic model by comparing it to a manual review of a random sample of reviews with three human coders.[21] We find that the coders' views generally accord with the topic scores from zero-shot classification. The coders' aggregated views are highly correlated with the topic scores, with correlations ranging from 0.53 to 0.79 across topics and the highest correlation of 0.79 for the prices topic. Correlations between the topic scores and individual reviewers are between 0.47 and 0.76, compared to between 0.71 and 0.85 between the individual reviewers, and are also the highest for prices. In addition, the coders are more likely to disagree amongst themselves or be unsure about the review when the LLM itself is uncertain – that is, when the topic scores are in the intermediate range close to 0.5. We provide further details of this analysis in Appendix D.

We first examine how the probabilities of each topic change post-divestiture. We observe a large increase in the number of consumers discussing prices: the share of reviews about prices increases 68% in the first year and 27% in the second year. Table 5 provides the estimate for the first year and second year for each topic. Without knowing if these reviews are complaints or praise, however, we do not know if reviews about prices explain the negative trend. Therefore, we will restrict attention to negative reviews.

We then examine how the probabilities of each topic change post-divestiture for unhappy consumers. To do so, we estimate equation (1) for the store-quarter sum of the probability of each topic for negative reviews. We then compare that estimate to the overall change in negative reviews

[20] See `https://huggingface.co/facebook/bart-large-mnli`.

[21] All three coders were undergraduate summer interns at the FTC.

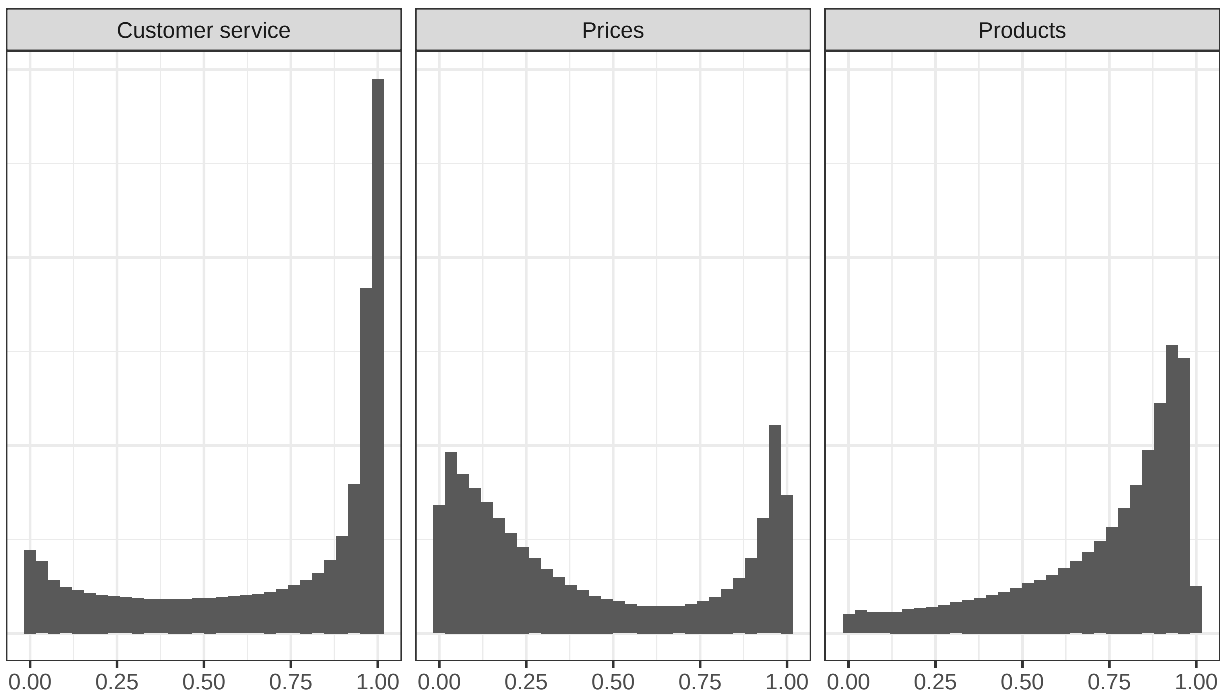

Figure 5: Histograms of topic scores across all reviews in states with divestitures.

Note: Topic scores computed using zero-shot encoding on all reviews from the following states and time periods: AZ, CA, MT, NV, OR, TX, WA, WY (2013–17); DE, MA, MD, NY, PA, VA, WV (2014–19).

by reporting estimates of the change in share of each topic relative to all negative reviews. Figure 6 depict the event study estimates, while Table 5 provides the estimate for the first year and second year for each topic. Since these three topics are not mutually exclusive, it is possible for the share of all three topics in consumers' reviews to increase or decrease.

We observe a large increase in the number of consumers discussing prices in their negative reviews post-divestiture, with an increase of 104% in the first year and 41% in the second year. The estimated effect is largest for the Haggen divested stores. In the first and second year following the divestitures, the share of consumers commenting on prices in negative reviews increases by about 100% and 40%, respectively. This finding is consistent with allegations made by Haggen that Albertsons/Safeway did not share pre-divestiture store level pricing information. Because it lacked this pricing information, Haggen claimed that it charged prices that were much higher than consumers expected when it took over the divested stores.[22] Interestingly, consumers were still more likely to mention pricing in negative reviews after the Haggen stores had been sold to other buyers in the second year following the divestiture.

For the non-Haggen stores, the increase in the fraction of negative reviews discussing prices is smaller in the first year. While the estimates are much less precise, we estimate that the fraction of negative reviews mentioning price increased by 70% and 45% in the first and second year following the divestiture relative to the pre-divestiture period.

We also observe differences in the change in the mentions of products and customer service between the models estimated using the Haggen and non-Haggen divested stores. For the non-Haggen stores, we estimate an increase in the fraction of negative reviews mentioning products (37%) and customer service (5%). Although these estimates are imprecise, they are suggestive of lower quality on product selection post-divestiture. By contrast, for the Haggen sample, we observe a smaller increase in the fraction of reviews mentioning products, at 16% in the year following the divestiture, and a decrease in the fraction of reviews mentioning customer service, at -25% in the year following the divestiture. Overall, we find that negative reviews about products increase by

[22] In its legal complaint against Albertsons after the divestiture, Haggen claimed that Albertsons/Safeway did not share its pre-merger pricing data (See paragraphs 35 and 36 of the complaint). Haggen's plan had been to use the same general pricing approach as Albertsons/Safeway when launching the new stores. Instead, Haggen was forced to uses its existing pricing strategy for its stores in the Pacific Northwest that were often much higher than the prices previously charged at the Southern California stores. This allegation is also consistent with some of the complaints we have seen in the text of Haggen's Yelp reviews.

17% in the first year and 11% in the second year, and negative reviews about customer service decline by 22% in the first year and 10% in the second year.

Table 5: Effect of divestiture on topics' share of reviews.

	Share of All Reviews			Share of Negative Reviews		
Topic	Quarters 0-3	Quarters 4-7	Pre-mean	Quarters 0-3	Quarters 4-7	Pre-mean
Prices	68%	27%	0.43	104%	41%	0.39
	[48%, 90%]	[13%, 43%]		[68%, 149%]	[17%, 70%]	
Products	13%	11%	0.70	17%	11%	0.66
	[6%, 20%]	[4%, 18%]		[7%, 29%]	[0%, 23%]	
Customer Service	-20%	-7%	0.75	-22%	-10%	0.73
	[-24%, -15%]	[-12%, -2%]		[-28%, -14%]	[-19%, 0%]	

Note: This table shows the results of the difference-in-differences regressions for topics as a share of all reviews ('Share of All Reviews') and as a share of negative reviews ('Share of Negative Reviews'). The Quarters 0-3 and 4-7 columns show the estimated dynamic treatment effects, both point estimates and 95% confidence intervals, transformed to percentage units. The 'Pre-mean' column shows the ratio of means.

7 Robustness

In this section, we examine the robustness of our estimates to several threats to identification, including a violation of the parallel trends assumption and alternative control groups. We also examine alternative ways of clustering the standard errors and an alternative LLM based approach to measuring sentiment. Our results are generally robust to these alternative approaches.

Parallel Trends Our difference in difference estimates rest on an assumption of parallel trends: that shocks unrelated to the divestiture will affect treatment and control stores symmetrically. We examine this assumption through the econometric approach of Rambachan and Roth (2023). Rambachan and Roth (2023) assumes an upper bound on any violation of parallel trends based upon a fixed multiple of the largest observed violation prior to treatment, and then computes confidence intervals by varying this multiple.

In Appendix B.1, we depict the confidence intervals for our main estimates based upon this approach, examining both the effect for the first quarter and the average effect for the first year across a wide set of potential upper bounds. We assess robustness by the multiple of the largest pre-treatment violation at which our estimate lose significance; our baseline confidence intervals

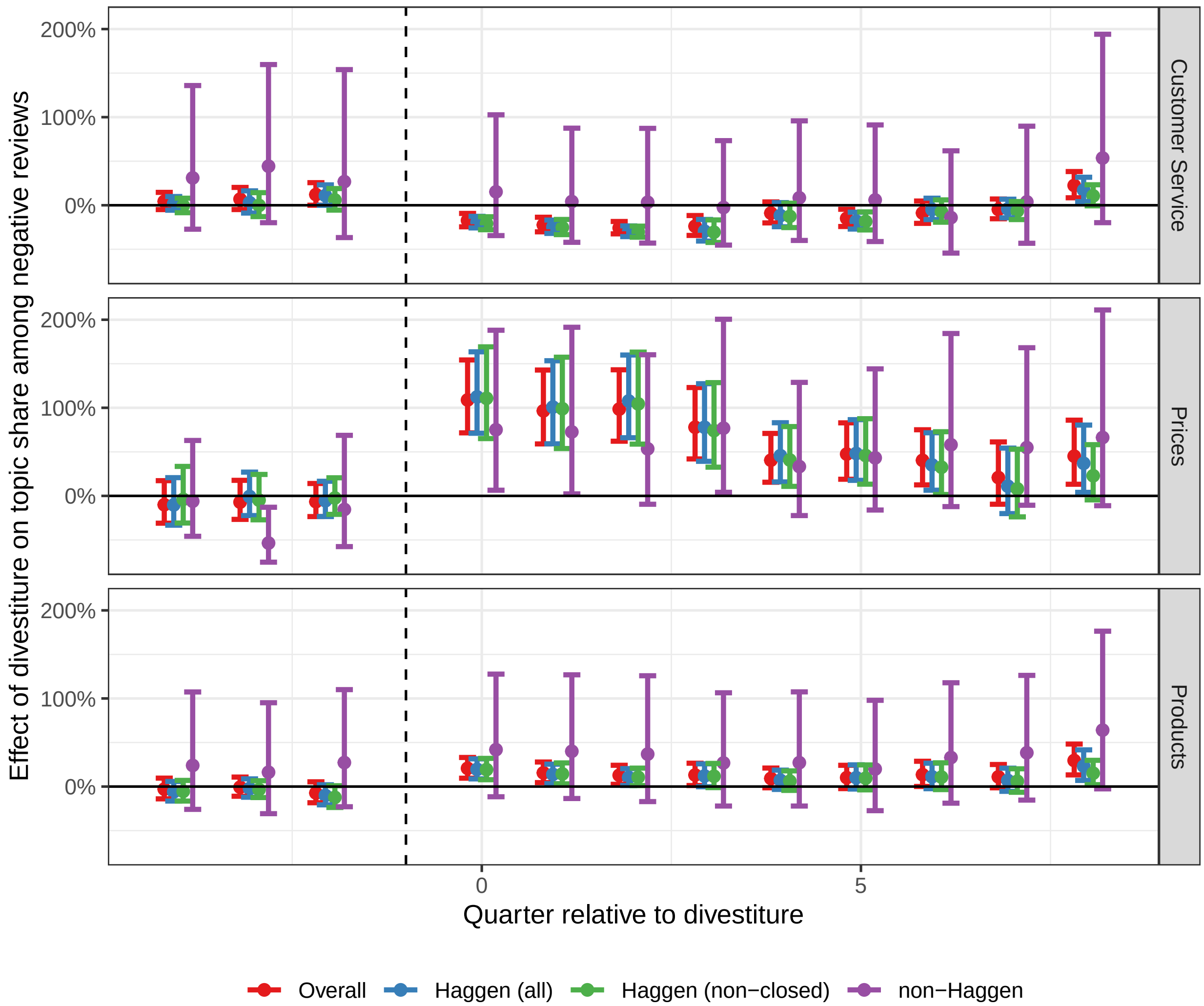

Figure 6: Effect of divestiture on each topic's share among negative reviews (event study plot)

assume a multiple of 0, while Rambachan and Roth (2023) consider a multiple of 1 as a natural benchmark.

Our estimates are quite robust to violations of parallel trends under this methodology. For the increase in the number of reviews, our estimates remain significant at a multiple of 2 for the effect in the first quarter, and lose significance at a multiple of 2 for the effect in the first year. For the increase in the share of negative reviews on prices, our estimates remain significant at a multiple of 2 for the effect in the first quarter, and lose significance at a multiple of 1 for the effect in the first year. Our estimates are less robust for the increase in the share of negative reviews; for this measure, our estimates lose significance at a multiple of 1 for the effect in the first quarter, and at a multiple of 0.4 for the effect in the first year.

Alternative Control Groups The key to any difference in difference analysis lies in the definition of the control group. Our baseline estimates used the control group of stores in the same state as a divested store and not located in ZIP3 regions with a divested store. In this section, we discuss broader and narrower alternative control groups.

For a broader control group, we consider all stores in the same state as a divested store, and so also include stores in the same ZIP3 as the divested stores. This control group is larger, and so will yield slightly more precise estimates, but risks contamination due to competition between divested stores and control-group stores. For a narrower control group, we consider only stores in our baseline control group that are "chain" stores. For each state, we define a chain store are stores with a name that appears five or more times in the state. These stores are more likely to be similar to the divested stores, which were large supermarkets owned by major chains.

In Appendix B.2, we report difference in difference estimates for our main specifications using both alternative control groups, and find that our results are almost unchanged using either group.

Alternative Clustering Approaches Our baseline estimates cluster our standard errors at the ZIP3 level, which allows for correlations between stores due to local market shocks. In Appendix B.3, we examine several forms of clustering, including clustering by location (narrower than our baseline), by supermarket chain to allow for shocks to chains that affect all chain stores, and two-way clustering by local area (ZIP3) and chain, and find broadly similar standard errors.

Alternative Sentiment Scores Our baseline estimates identify negative reviews based on low star ratings. To validate our use of star ratings, we also use LLMs to assign a sentiment score to each review based on the review text, based upon the `twitter-roberta-base-sentiment-latest` model downloaded from Hugging Face.[23] In Appendix B.4, we show that the LLM sentiment scores are highly correlated with star ratings, and that our difference in difference estimates on the share of negative reviews are similar if we use the continuous negativity sentiment score to measure negative reviews.

8 Mechanisms

Absent competition concerns, voluntary asset sales should improve total welfare by transferring assets to better owners. However, the divestitures we study were mandated by antitrust authorities and so not voluntary. In particular, because the merged firm will compete with the divested stores post-divestiture, the inherent conflict of interest could lead the merged firm to harm the relative competitive position of the new owner, choose to sell assets to weak buyers that are less likely to restore the previous level of competition post-divestiture, or choose to sell weak assets.

Multiple lawsuits provide details of such strategic behavior. For example, as described earlier, Haggen sued Albertsons following the Albertsons/Safeway divestiture, alleging that Albertsons took a number of actions in violation of the purchase agreement that handicapped the entry of Haggen, including refusing to share pricing information on the divested stores. Recently, Albertsons sued Kroger after the Kroger/Albertsons merger was blocked by multiple courts, claiming that Kroger failed to exercise its best efforts to provide the FTC with a viable divestiture package. Albersons alleged that Kroger rejected stronger divestiture buyers than its preferred candidate of C&S Wholesale Grocers, which operated primarily as a wholesaler and had a history of acquiring and closing retail grocery stores. In addition, Albersons claimed that Kroger "cherry-picked" under-performing and less valuable stores to divest, rather than divesting stores that would be the most likely to remedy competition concerns from the deal.[24]

Supermarket divestitures also lead to other simultaneous changes that affect consumers' shop-

[23] See `https://huggingface.co/cardiffnlp/twitter-roberta-base-sentiment-latest`.

[24] See `https://www.grocerydive.com/news/albertsons-kroger-lawsuit-merger-ftc/735780/` and `https://www.albertsonscompanies.com/newsroom/press-releases/news-details/2024/Albertsons-Files-Lawsuit-Against-Kroger-for-Breach-of-Merger-Agreement/default.aspx` for more details.

ping experience. First, the divestitures resulted in a change in ownership of the stores, which could disrupt store operations. Second, all of the divested stores changed banners, with most of the new banners unfamiliar to consumers living near the stores. For example, all of the divested California stores were subsequently operated as Haggen stores, a banner not previously operated in California. It might be difficult to convince consumers to visit an unfamiliar chain, or to ensure that the new banner matches local consumer tastes on price and quality dimensions.

To isolate these alternative potential explanations, we conduct two additional empirical exercises using the same statistical model and identification strategy as in Section 4. First, we examine changes in stores' reviews for those stores that changed ownership in the Albertsons/Safeway and Ahold/Delhaize deals. For these stores, the supermarket banner did not change with the merger. Second, we examine the sale of grocery stores after A&P's bankruptcy in 2015 that resulted in A&P's exit from grocery retailing in the U.S. Here, the bankruptcy sale leads to an ownership change and banner change. However, because A&P left the market, it did not have conflicts of interest that might lead it to sell the stores to anyone but the highest bidder for them, or to sabotage them before the sale.

These results of these analyses are consistent with conflicts of interest of the merged firm as the likely explanation for our findings. The acquired stores of the merged Albertsons/Safeway or Ahold/Delhaize see little change in reviews after the merger. While these acquisitions may have led to management challenges, the retail brands available to consumers did not change. For the A&P bankruptcy, we also find increases in reviews, albeit smaller than for divested stores, which is consistent with consumers noticing and commenting on the new banner. However, reviews improve post-sale, which suggests that consumers were made better off with the voluntary asset sales.

8.1 Acquired Firm Stores

In this analysis, we consider all non-divested stores owned by the acquired firms in the two transactions (Safeway or Delhaize) that were in a state with divestitures as treated by the merger. Our control group is stores owned by grocery retailers unaffected by the merger operating in the same states as stores with divestitures, but different 3-digit zip codes than the acquired stores.[25]

In Figure 7, Figure 8, and Table 6, we report the estimates from this analysis. We find small

[25] Appendix A.2 describes the sample.

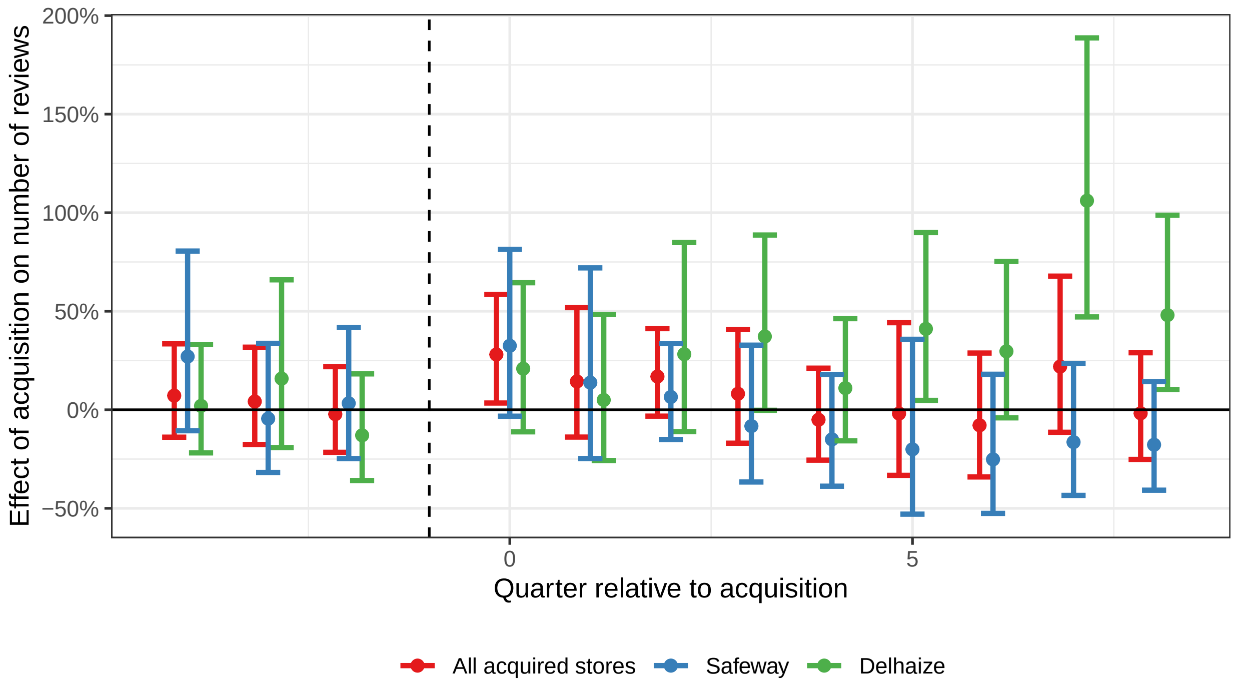

Figure 7: Effect of acquisition on number of reviews (event study plot)

and statistically insignificant increases in reviews post-merger for the acquired firms' stores at 17% in the first year and 1% in the second year. We also find a statistically significant increase in the proportion of negative reviews in the second year after the merger (37%). As with the divested stores, we observe an estimated increase in the number of negative reviews mentioning prices, although we do not find a statistically significant change in the other topics discussed in negative reviews.

Since Safeway operated stores in over 90% of 3-digit zip-codes in California before and after the merger, the estimates in Table 6 are primarily identified using information for the Safeway stores not in California and the stores owned by Delhaize in states with divestitures. We thus also estimate these models using a broader control group, containing all grocery stores not owned by the merging firm or a divestiture buyer located in the states with divestitures, as we discuss in Section 7.[26] These results are shown in Appendix B.5 (see Table 18); with much smaller standard errors, we find effectively no change in the number of reviews, share of negative reviews, and in the topics mentioned in negative reviews.

[26] As we further discuss in Appendix B.2, a potential weakness of the all-state control group is that it may lead to attenuated estimated treatment effects.

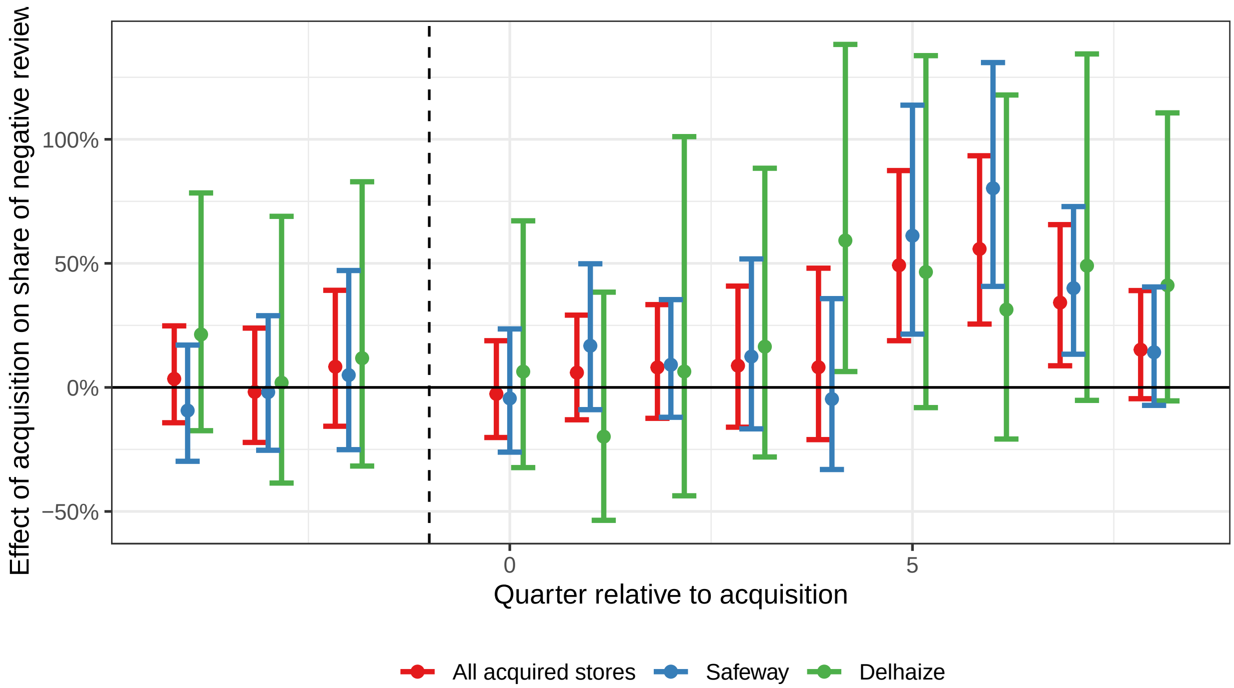

Figure 8: Effect of acquisition on share of negative reviews (event study plot)

Table 6: Summary of difference-in-differences estimates for acquired stores.

			Topic Share of Negative Reviews		
Time Period	No. Reviews	Share Negative Reviews	Prices	Products	Customer Service
Quarters 0-3	17%	5%	22%	2%	2%
	[-4%, 42%]	[-11%, 24%]	[0%, 48%]	[-8%, 14%]	[-9%, 14%]
Quarters 4-7	1%	37%	28%	0%	7%
	[-24%, 35%]	[14%, 64%]	[3%, 58%]	[-11%, 12%]	[-6%, 21%]
Pre-mean	1.40	0.51	0.34	0.61	0.74

Note: This table shows the results of the difference-in-differences regressions for the acquired store treatment. Each column represents a different dependent variable; the table presents the estimated effect on either the dependent variable itself, or on the *share* of the dependent variable within a broader category, computed as the difference between estimated regression coefficients. The Quarters 0-3 and 4-7 rows show the estimated dynamic treatment effects, both point estimates and 95% confidence intervals, transformed to percentage units. The 'Pre-mean' row shows either the mean of the dependent variable in the treatment group in the quarter before the treatment, or the ratio of means (mean dependent variable divided by mean of 'as share of' variable).

We also examine several additional treatment groups and find generally null effects. First, we examine effects for all merging firms' stores, combining acquired and acquiring firm's stores, in Appendix B.6, and find similar null effects in Table 19. Overall, we conclude that the stores owned and operated by the merging firms did not experience an important change in the number or content of reviews.

Second, we examine effects on local competitor stores, defined as chain stores in the same local market (ZIP3) as a divested store, but not involved in the merger, in Appendix B.7. We do not find detectable changes in reviews for such stores following a divestiture.

8.2 Former A&P Stores

A&P was arguably the first grocery retailer to successfully exploit the scale economies associated with chains in the early 20th century, including through the introduction of private label products and improvements in distribution (Ellickson, 2016). However, by the early 2000s, A&P's significance in the grocery industry had declined to a regional firm with primary operations in the New York City metropolitan area. In July 2015, A&P filed for bankruptcy and began the process of closing and selling off all of its remaining 281 grocery stores, with the last A&P store closing on November 25th, 2015.

A&P's store locations were sold to a number of different grocery and non-grocery retailers. Of the 226 A&P stores that appeared in the Yelp reviews data, 154 stores remained operating as grocery stores in 2016 following A&P's exit and sale of store locations.[27]

A large proportion (60%) of the former A&P stores that remained in the supermarket industry were sold to three relatively large grocery retailers with a major presence in the New York City metro area: Albertsons (owner of Acme), Ahold/Delhaize (owner of Stop & Shop), and ShopRite. The remainder were sold to smaller firms.[28]

The sale of A&P's stores to other grocery retailers represents both a change in the retail banner and in ownership. However, in selling the stores to other food retailers, A&P did not have an incentive in handicapping the stores or selling to a weaker buyer as it no longer remained a grocery

[27] The remaining 72 former A&P stores were operated as non-grocery retail outlets, such as operating as liquor stores.

[28] Appendix A.3 presents some descriptive statistics about the treated stores of former A&P stores and the control group of grocery stores owned by firms not involved in the A&P transactions located in different 3-digit zip codes, but the same states, as the A&P stores.

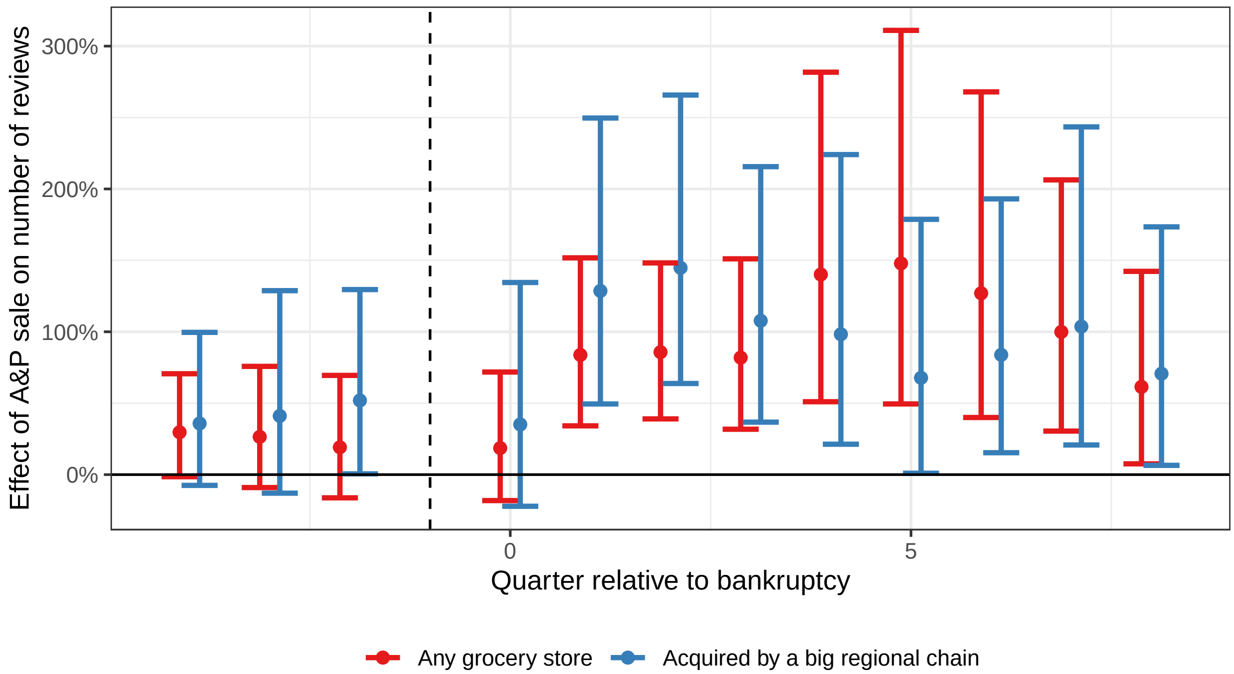

Figure 9: Effect of A&P sale on number of reviews (event study plot)

retailer post-sale.

We show the primary results of our differences in differences analysis in Figure 9, Figure 10, and Table 7. The sale of the former A&P stores to non-A&P retailers caused consumer reviews to increase significantly, with increases in the number of reviews of 66% over the first year post-sale and 129% over the second year. These change are statistically significant. The increased consumer reviews are much less likely to be negative: the share of negative reviews falls by 29% and 40%, respectively, in the first first year and second year following the sale. Again, these results are statistically different than zero. We also find an increase in the share of positive reviews on customer service at a 51% increase in the first year and a 43% in the second year. We do not, however, observe any statistically significant change in the content of positive reviews for the prices and products topics.

In addition, we restrict the analysis to the stores sold to three large regional chains (Albertsons, Ahold/Delhaize, and ShopRite) in Appendix B.8. We again find that consumers increase their use of voice following the sale of stores to other regional chains, that the increase in consumer voice is significantly less likely to be negative, and that the topic of customer service increases as a share

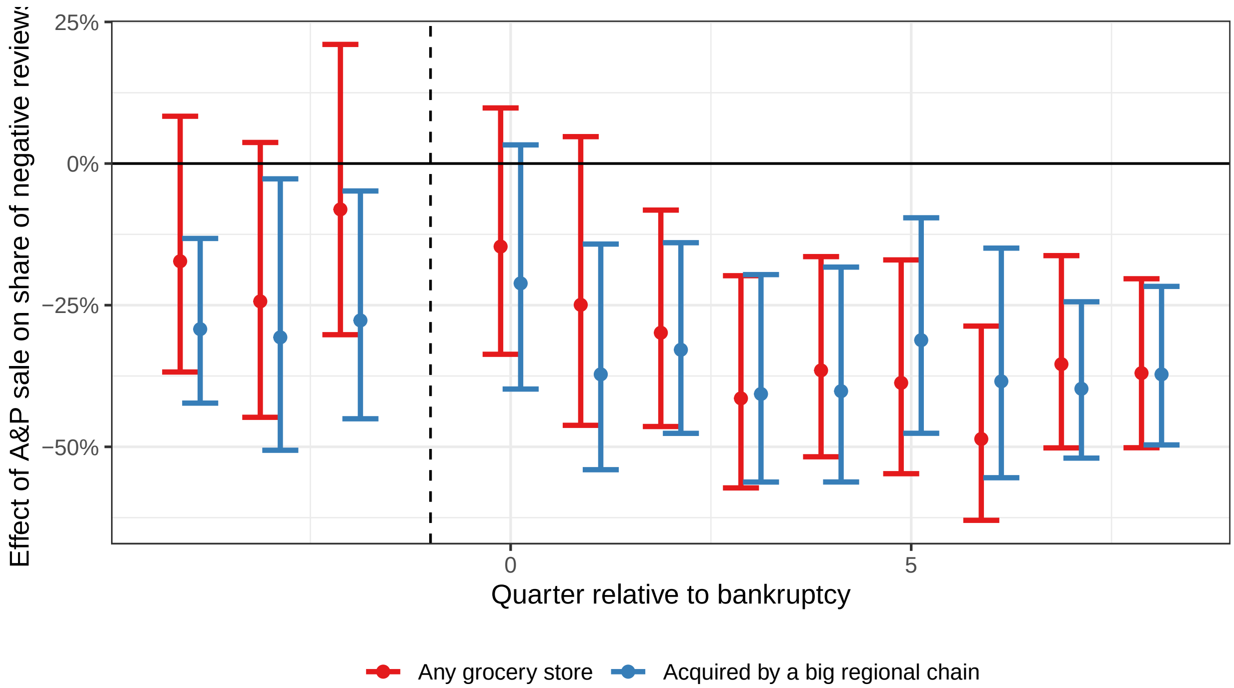

Figure 10: Effect of A&P sale on share of negative reviews (event study plot)

Table 7: Summary of difference-in-differences estimates for post-A&P stores.

			Topic Share of Positive Reviews		
Time Period	No. Reviews	Share Negative Reviews	Prices	Products	Customer Service
Quarters 0-3	66%	-29%	10%	14%	51%
	[26%, 121%]	[-44%, -10%]	[-40%, 101%]	[-28%, 81%]	[1%, 127%]
Quarters 4-7	129%	-40%	2%	11%	43%
	[47%, 257%]	[-53%, -22%]	[-53%, 121%]	[-38%, 96%]	[-9%, 125%]
Pre-mean	0.60	0.67	0.32	0.43	0.29

Note: This table shows the results of the difference-in-differences regressions for the A&P store treatment. Each column represents a different dependent variable; the table presents the estimated effect on either the dependent variable itself, or on the *share* of the dependent variable within a broader category, computed as the difference between estimated regression coefficients. The Quarters 0-3 and 4-7 rows show the estimated dynamic treatment effects, both point estimates and 95% confidence intervals, transformed to percentage units. The 'Pre-mean' row shows either the mean of the dependent variable in the treatment group in the quarter before the treatment, or the ratio of means (mean dependent variable divided by mean of 'as share of' variable).

of positive reviews.

9 Discussion and Conclusion

In this article, we have examined how consumers responded to a major antitrust remedy – government divestitures – in the supermarket industry; we studied the divestiture of 249 stores in the Albertsons/Safeway and Ahold/Delhaize transactions. Using consumer reviews on Yelp, we found that consumer reviews rose massively after the divestiture for divested stores. These additional reviews are mostly negative. We then used a large language model to classify the topic of these reviews, and found that negative reviews about prices drive the rise in negative reviews after the divestiture. Our estimates on changes in consumer voice are similar for the stores divested to Haggen in the Albertsons/Safeway transaction, which went bankrupt soon after the divestiture, and other stores.

While our work has shown that consumer voice responds to divestitures, the underlying mechanisms behind the consumer dissatisfaction that we document are less clear. A major difference between government mandated divestitures due to antitrust concerns and voluntary asset sales is that the merging parties will compete against the divestiture buyer and so have potential conflicts of interest in structuring the divestiture. To determine which mechanisms are most likely to be the source of increased divestitures, we also examine two voluntary asset sales that should not have such conflicts. First, we examine how reviews changed as stores that were acquired as the result of these mergers: the Safeway stores acquired by Albertsons and the Delhaize stores acquired by Ahold. We find little change in the reviews of the acquired stores after the ownership change.

We next examine the fate of stores sold in the bankruptcy of A&P. These stores experience a change in ownership and retail banner, but A&P did not have an incentive to harm the stores as it exited grocery retailing. We find smaller increases in reviews than the divested stores and decreases in the share of negative reviews. Overall, these additional analyses point to conflicts of interest as a likely explanation for the finding that divestitures dramatically increased reviews. To further decompose the different effects of a divestiture, future work could examine other changes such as a banner change without an ownership change.

Our work provides guidance for antitrust authorities considering whether divestitures could remedy competition problems for proposed mergers. First, our analysis found increases in negative reviews discussing prices post-divestiture, and so reinforces the traditional antitrust concern with effects on price from changes in competition. Second, because reviews are public, competition authorities could use reviews to provide real-time feedback in the success of divestitures, and intervene early for divestitures that appear to be failing. Finally, "divestiture retrospectives" examining previous divestitures, similar to those conducted in this article, could provide agencies with evidence on what types of divestitures are more or less likely to succeed across a set of industries. Such evidence could allow them to improve the form of divestitures and ensure that they eliminate the potential harms to competition from mergers of competing firms.

Our results are also relevant for policymakers seeking to enhance the effectiveness of regulation. Industry regulators often exercise their powers to mandate specific private actions, such as improvements or asset sales, despite having incomplete information about quality outcomes (Laffont and Tirole, 1993). In addition, when prices are set by regulators, providers may find ways to reduce quality, as in Eliason et al. (2020) for dialysis. In such markets, consumer voice – for example, about patient care at a hospital or customer service issues at a regulated utility – provides an alternative to disclosures on quality mandated by regulators, or to intermediaries such as certifiers (Dranove and Jin, 2010). Consumer voice may also allow the regulator to intervene quickly to address unexpected problems and, in the long run, improve outcomes.

References

Agarwal, Nikhil, Ray Huang, Alex Moehring, Pranav Rajpurkar, Tobias Salz, and Feiyang Yu. 2024. "Comparative Advantage of Humans versus AI in the Long Tail." Vol. 114, 618–622, American Economic Association 2014 Broadway, Suite 305, Nashville, TN 37203.

Anderson, Michael, and Jeremy Magruder. 2012. "Learning from the crowd: Regression discontinuity estimates of the effects of an online review database." *The Economic Journal*, 122(563): 957–989.

Argentesi, Elena, Paolo Buccirossi, Roberto Cervone, Tomaso Duso, and Alessia Marrazzo. 2021. "The effect of mergers on variety in grocery retailing." *International Journal of Industrial Organization*, 79: 102789.

Ashenfelter, Orley C, Daniel S Hosken, and Matthew C Weinberg. 2013. "The price effects of a large merger of manufacturers: A case study of Maytag-Whirlpool." *American Economic Journal: Economic Policy*, 5(1): 239–261.

Ashenfelter, Orley C, Daniel S Hosken, and Matthew C Weinberg. 2015. "Efficiencies brewed: pricing and consolidation in the US beer industry." *The RAND Journal of Economics*, 46(2): 328–361.

Ashenfelter, Orley, Daniel Hosken, and Matthew Weinberg. 2014. "Did Robert Bork understate the competitive impact of mergers? Evidence from consummated mergers." *The Journal of Law and Economics*, 57(S3): S67–S100.

Asker, John, and Volker Nocke. 2021. "Collusion, mergers, and related antitrust issues." In *Handbook of industrial organization.* Vol. 5, 177–279. Elsevier.

Beard, T Randolph, Jeffrey T Macher, and John W Mayo. 2015. ""Can you hear me now?" Exit, voice, and loyalty under increasing competition." *The Journal of Law and Economics*, 58(3): 717–745.

Brown, David P, Andrew Eckert, and Blake Shaffer. 2023. "Evaluating the impact of divestitures on competition: Evidence from Alberta's wholesale electricity market." *International Journal of Industrial Organization*, 89: 102953.

Cameron, A. Colin, Jonah B. Gelbach, and Douglas L. Miller. 2011. "Robust Inference With Multiway Clustering." *Journal of Business & Economic Statistics*, 29(2): 238–249.

Chartock, Benjamin L. 2023. "Quality Disclosure, Demand, and Congestion: Evidence from Physician Ratings."

Chen, Viola, Christopher Garmon, Kenneth Rios, and David Schmidt. 2022. "The competitive efficacy of divestitures: an empirical analysis of generic drug markets." *Available at SSRN 4036607.*

Dranove, David, and Ginger Zhe Jin. 2010. "Quality disclosure and certification: Theory and practice." *Journal of economic literature*, 48(4): 935–963.

Dranove, David, Daniel Kessler, Mark McClellan, and Mark Satterthwaite. 2003. "Is more information better? The effects of "report cards" on health care providers." *Journal of political Economy*, 111(3): 555–588.

Eliason, Paul J, Benjamin Heebsh, Ryan C McDevitt, and James W Roberts. 2020. "How acquisitions affect firm behavior and performance: Evidence from the dialysis industry." *The Quarterly Journal of Economics*, 135(1): 221–267.

Ellickson, Paul B. 2016. "The evolution of the supermarket industry: from A & P to Walmart." In *Handbook on the Economics of Retailing and Distribution.* 368–391. Edward Elgar Publishing.

Federal Trade Commission. 1999. "A study of the Commission's Divestiture Process."

Federal Trade Commission. 2017. "The FTC's Merger Remedies 2006-2012: A Report of the Bureaus of Competition and Economics."

Foster, Lucia, John Haltiwanger, and Cornell J Krizan. 2006. "Market selection, reallocation, and restructuring in the US retail trade sector in the 1990s." *The Review of Economics and Statistics*, 88(4): 748–758.

Fradkin, Andrey, and David Holtz. 2023. "Do incentives to review help the market? evidence from a field experiment on airbnb." *Marketing Science.*

Fradkin, Andrey, Elena Grewal, and David Holtz. 2021. "Reciprocity and Unveiling in Two-sided Reputation Systems: Evidence from an Experiment on Airbnb." *Marketing Science*, 40(6).

Friberg, Richard, and André Romahn. 2015. "Divestiture requirements as a tool for competition policy: A case from the Swedish beer market." *International journal of industrial organization*, 42: 1–18.

Gans, Joshua S, Avi Goldfarb, and Mara Lederman. 2021. "Exit, tweets, and loyalty." *American Economic Journal: Microeconomics*, 13(2): 68–112.

General Accountability Office. 2002. "Study Needed to Assess the Effects of Recent Divestitures on Competition in Retail Markets."

Grosz, Michel, and Devesh Raval. 2025. "Amplifying Consumers' Voice: The FTC's Report-Fraud Website Redesign." *Marketing Science*, 44(3): 525–545.

He, Sherry, Brett Hollenbeck, and Davide Proserpio. 2022. "The market for fake reviews." *Marketing Science*, 41(5): 896–921.

Hirschman, Albert O. 1972. *Exit, voice, and loyalty: Responses to decline in firms, organizations, and states.* Harvard university press.

Hollenbeck, Brett. 2018. "Online reputation mechanisms and the decreasing value of chain affiliation." *Journal of Marketing Research*, 55(5): 636–654.

Hollenbeck, Brett, Sridhar Moorthy, and Davide Proserpio. 2019. "Advertising strategy in the presence of reviews: An empirical analysis." *Marketing Science*, 38(5): 793–811.

Jin, Ginger Zhe, and Phillip Leslie. 2003. "The effect of information on product quality: Evidence from restaurant hygiene grade cards." *The Quarterly Journal of Economics*, 118(2): 409–451.

Kolstad, Jonathan T. 2013. "Information and quality when motivation is intrinsic: Evidence from surgeon report cards." *American Economic Review*, 103(7): 2875–2910.

Laffont, Jean-Jacques, and Jean Tirole. 1993. *A theory of incentives in procurement and regulation.* MIT press.

Lagos, Vicente. 2018. "Effectiveness of Merger remedies: Evidence from the retail gasoline industry." *The Journal of Industrial Economics*, 66(4): 942–979.

Lewis, Gregory, and Georgios Zervas. 2020. "The Supply and Demand Effects of Review Platforms." *Unpublished manuscript.*

Luca, Michael. 2011. "Reviews, Reputation, and Revenue: The Case of Yelp. com."

Luca, Michael, and Georgios Zervas. 2016. "Fake It Till You Make It: Reputation, Competition, and Yelp Review Fraud." *Management Science*, 62(12): 3412–3427.

Mayzlin, Dina, Yaniv Dover, and Judith A. Chevalier. 2014. "Promotional Reviews: An Empirical Investigation of Online Review Manipulation." *The American Economic Review*, 104(8): 2421–2455.

Nosko, Chris, and Steven Tadelis. 2015. "The Limits of Reputation in Platform Markets: An Empirical Analysis and Field Experiment." National Bureau of Economic Research.

Osinski, F David, and Jeremy A Sandford. 2021. "Evaluating mergers and divestitures: A casino case study." *The Journal of Law, Economics, and Organization*, 37(2): 239–277.

Rambachan, Ashesh, and Jonathan Roth. 2023. "A More Credible Approach to Parallel Trends." *The Review of Economic Studies*, 90(5): 2555–2591.

Raval, Devesh. 2020. "Whose voice do we hear in the marketplace? Evidence from consumer complaining behavior." *Marketing Science*, 39(1): 168–187.

Soetevent, Adriaan R, Marco A Haan, and Pim Heijnen. 2014. "Do auctions and forced divestitures increase competition? Evidence for retail gasoline markets." *The Journal of Industrial Economics*, 62(3): 467–502.

Tenn, Steven, and John M Yun. 2011. "The success of divestitures in merger enforcement: Evidence from the J&J–Pfizer transaction." *International journal of industrial organization*, 29(2): 273–282.

Wang, Xiangrui, Ron C Mittelhammer, Thomas L Marsh, and Jill J McCluskey. 2023. "Is divestiture effective as a merger remedy in the US beer industry?" *Review of Industrial Organization*, 62(1): 1–18.

Weaver, Suzanne. 1977. *Decision to prosecute: Organization and public policy in the Antitrust Division.* MIT press Cambridge, MA.

Williamson, Oliver E. 1968. "Economies as an Antitrust Defense: The Welfare Tradeoffs." *The American Economic Review*, 58(1): 18–36.

Wooldridge, Jeffrey M. 2023. "Simple approaches to nonlinear difference-in-differences with panel data." *The Econometrics Journal*, 26(3): C31–C66.

A Data Construction Details

A.1 Divestiture Timing

As described in Section 4, the treatment date in our difference-in-differences regressions is the quarter of divestiture, which may differ from the quarter the merger was finalized. In both transactions, about 35% of divestitures occurred in the same quarter the merger was finalized, as shown in Table 8 and Table 9.[29] The rest occurred the following quarter (except for ten Martin's stores near Richmond, Virginia, which were divested to Publix a year after the merger).

Table 8: Timing of divestitures in the Albertsons/Safeway transaction.

State	2015Q1	2015Q2
CA	18	65
MT	4	
OR	4	16
TX	12	
WA	19	9
WY	4	
AZ		10
NV		7

Note: Counts of divested stores by state and quarter of divestiture. Quarter of divestiture is used as the treatment date in regression specifications.

Table 9: Timing of divestitures in the Ahold/Delhaize transaction.

State	2016Q3	2016Q4	2017Q3
MA	8		
MD	4	21	
NY	5		
PA	3		
VA	6	13	10
WV	1		
DE		4	

Note: Counts of divested stores by state and quarter of divestiture. Quarter of divestiture is used as the treatment date in regression specifications.

[29] The Albertsons/Safeway merger closed on January 30, 2015, and the Ahold/Delhaize merger closed on July 23, 2016.

A.2 Acquired Firm Stores

This section details the data construction steps for the acquired store analysis in Section 8.

We first compile a list of Albertsons, Safeway, Ahold, and Delhaize brand-state combinations from a variety of online sources, and then match those brand-states to the user-generated store labels (strings) in Yelp. (The list is reproduced in Table 10, along with the count of matched locations.) We define an 'acquired' store as a location that *ever* matched a Safeway or Delhaize brand-state combination, was located in a state that also contained a divested store (in the same transaction), and was not itself divested.

Table 10: Count of stores matched to acquirer/acquired brands in Yelp data.

Owner	Brand	Locations
Ahold	Giant	426
	Martin's	51
	Stop & Shop	553
Albertsons	Acme	265
	Albertsons	1,231
	Amigos	20
	Jewel-Osco	239
	Lucky	127
	Market Street	21
	Shaws	237
	Star Market	38
	United Supermarkets	66
Delhaize	Bottom Dollar	54
	Food Lion	1,198
	Hannaford	219
Safeway	Carr's	18
	Pak 'N Save	8
	Pavilions	45
	Randalls	81
	Safeway	1,314
	Tom Thumb	216
	Vons	360

Note: Generated by string matching Yelp data to a list of brand-states compiled by the authors from online sources. The Locations column counts unique locations (by address and/or geocode).

Summary statistics on the resulting treatment, control, and excluded groups are shown in Table 11. Excluded stores are either in the same ZIP3 as an acquired store, were a divested store, or were owned at any time by the acquiring firm (Albertsons or Delhaize).

Table 11: Summary statistics on reviews in the acquired firm analysis.

		Control		Excluded		Treatment	
		Pre	Post	Pre	Post	Pre	Post
Delhaize	Reviews	53,401	77,878	41,207	65,966	950	1,960
	Stores represented	4,440	5,532	4,284	5,233	302	375
	ZIP3s represented	94	98	131	132	73	79
	Avg. rating	3.45	3.22	3.43	3.20	3.38	3.14
	% negative	28%	38%	30%	39%	26%	38%
Safeway	Reviews	1,767	10,675	135,266	530,233	14,698	61,533
	Stores represented	586	1,341	11,556	16,989	999	1,081
	ZIP3s represented	52	52	135	149	109	112
	Avg. rating	3.36	3.21	3.63	3.33	2.75	2.46
	% negative	32%	39%	23%	36%	47%	59%

Note: Compiled using reviews from January 2013 to July 2021. Review dates are assigned to pre/post merger using a cutoff of January 1, 2015 for Albertsons/Safeway and July 1, 2016 for Ahold/Delhaize, shortly before the mergers closed. Reviews are assigned to stores by unique address. Stores are assigned to treatment, control, and excluded groups as described in the text.

A.3 Former A&P Stores

In the A&P analysis in Section 8, the treated group contains 154 former A&P stores that remained grocery stores following the 2015 bankruptcy.[30] As in the main analysis, our unit of analysis is a store location, which we track across ownership and banner changes. Control stores were in the same state, but different ZIP3, as treated stores.

Table 12: Summary statistics on reviews in the A&P stores analysis.

	Control		Excluded		Treatment	
	Pre	Post	Pre	Post	Pre	Post
Reviews	19,067	58,013	37,316	101,804	905	3,755
Stores represented	3,220	5,067	3,786	5,551	121	142
ZIP3s represented	103	103	42	42	36	40
Avg. rating	3.46	3.28	3.45	3.22	2.38	2.56
% negative	28%	36%	27%	38%	58%	55%

Note: Compiled using reviews from January 2013 to July 2021. Review dates are assigned to pre/post using a cutoff of July 1, 2015, shortly before A&P filed for bankruptcy; store sales followed in the months afterward. Reviews are assigned to stores by unique address. Stores are assigned to treatment, control, and excluded groups as described in the text.

[30] Of the 281 A&P stores open at the time of the bankruptcy, 226 appear in our Yelp data, of which 72 locations did not remain grocery stores.

B Robustness Checks

B.1 Robustness to violations of parallel trends

To interpret our two-way fixed effects regression estimates as causal, we make a parallel trends assumption on the treatment and control groups. Our justification is that shocks unrelated to the divestiture (such as statewide macroeconomic shocks or changes in the userbase of Yelp) will affect both groups of stores in the same way.

In this section, we test the robustness of our key results to the parallel trends assumption, using the 'relative magnitudes' approach of Rambachan and Roth (2023). Rather than assuming parallel trends hold exactly, this approach assumes an upper bound on any violation of parallel trends (as a fixed multiple $\bar{M}$ of the largest observed parallel trends violation prior to treatment) and computes confidence intervals given that bound. The standard parallel trends assumption corresponds to a bound of $\bar{M} = 0$, while Rambachan and Roth (2023) write that $\bar{M} = 1$ may be a "natural benchmark."[31]

In Figure 11, we show that our reviews finding (from Figure 3), that reviews increase substantially upon divestiture, is highly robust. The effect in the first quarter ($t = 0$) remains significant (and large) even under a severe parallel trends violation of $\bar{M} = 2$. The average effect in the first four quarters (weighted equally across quarters) loses significance around $\bar{M} \approx 2$.

In Figure 12, we show that our negative share finding (from Figure 4), that the negative share of reviews increases upon divestiture, is more sensitive to parallel trends. The effect in the first quarter loses significance around $\bar{M} \approx 1$, and the average effect in the first four quarters loses significance at $\bar{M} \approx 0.4$. Our interpretation is that our estimates of its initial sign and magnitude are fairly reliable, but our estimates of the persistence of the negative effect may not be.

In Figure 13, we show that our prices topic finding (from Figure 6), that the share of negative reviews about prices increases upon divestiture, is highly robust in the short run. The effect in the first quarter remains significant even under a violation of $\bar{M} = 2$, but the average effect in the first four quarters loses significance around $\bar{M} \approx 1$.

[31] The relative magnitudes approach fits most naturally with our quarter-level regressions, which we present in event study figures, rather than our aggregated four-quarter regressions, which we present in tables. Accordingly, we approximate longer-run effects using equal-weighted averages of quarter-level estimates.

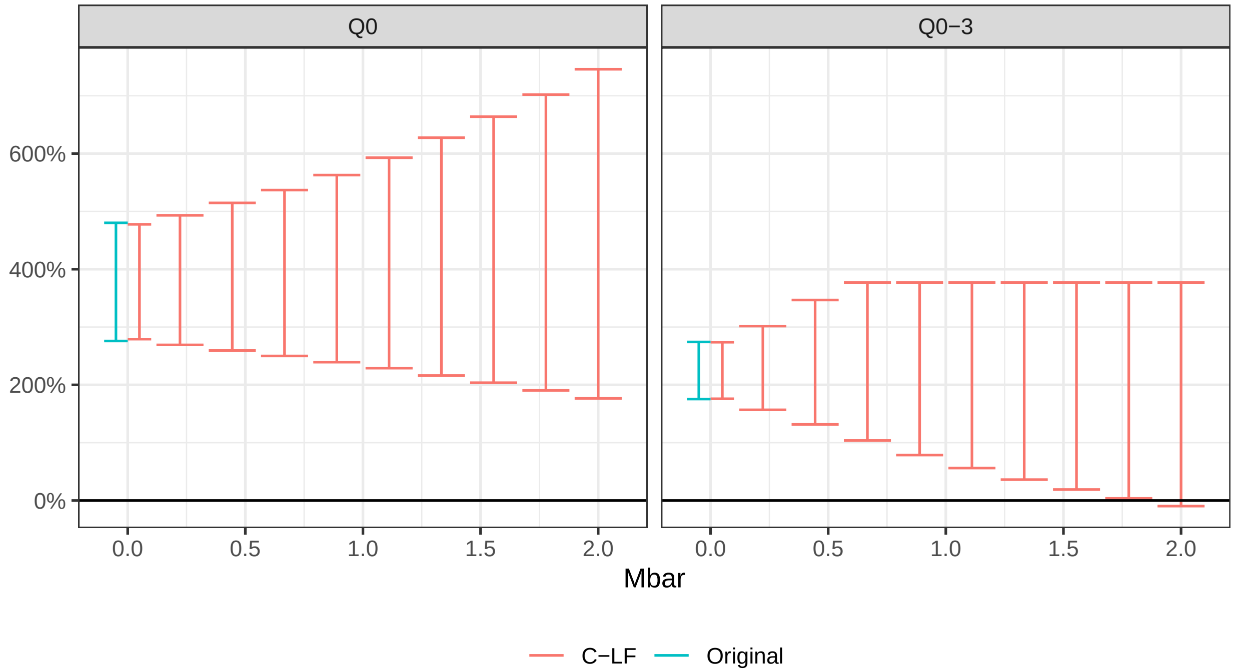

Figure 11: Robustness of the effect of divestiture on number of reviews.

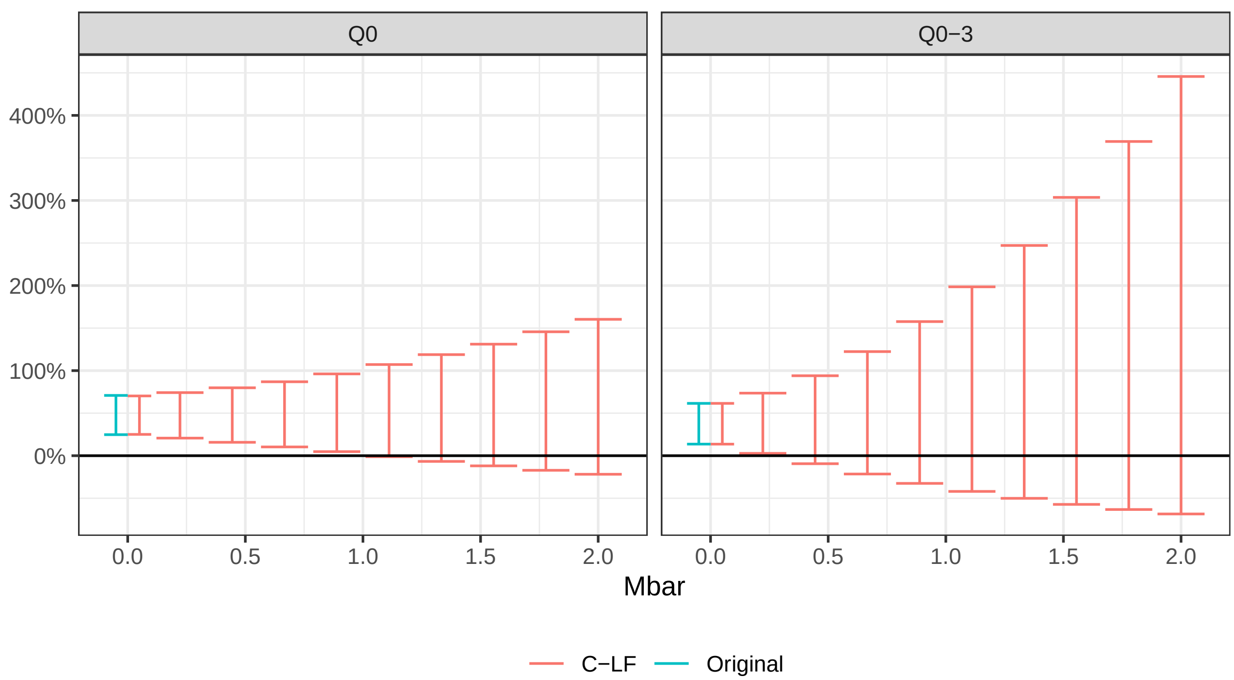

Figure 12: Robustness of the effect of divestiture on negative share of reviews.

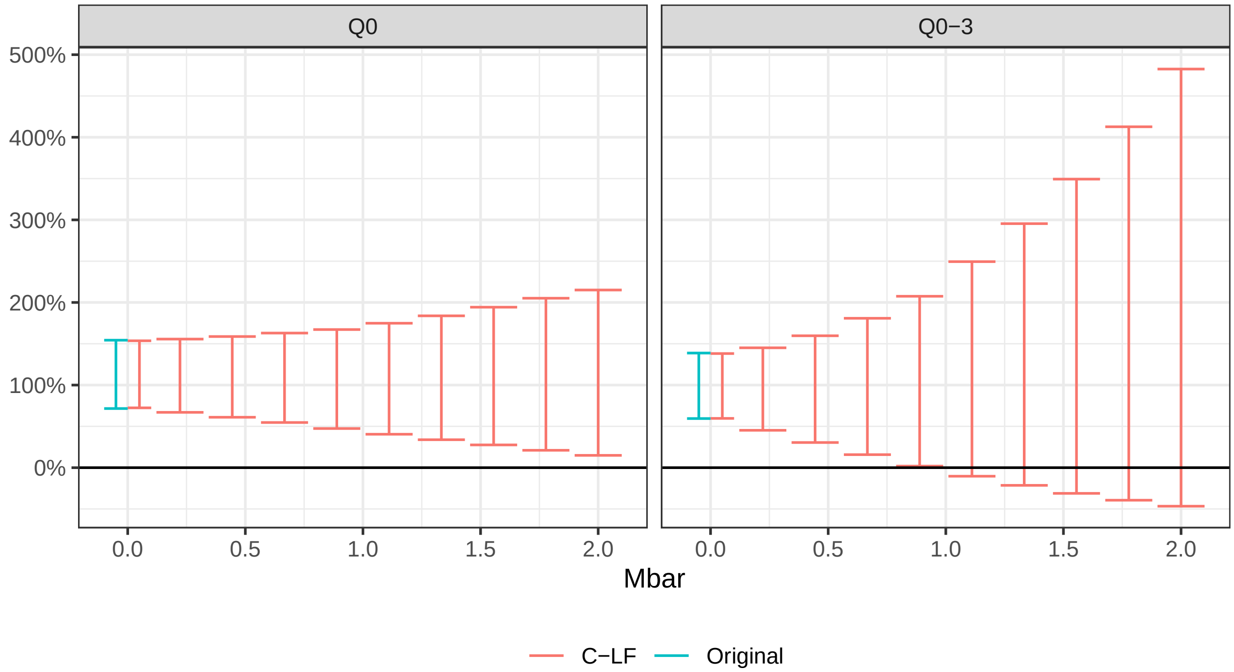

Figure 13: Robustness of the effect of divestiture on the price topic's share of negative reviews.

B.2 Alternative control groups

We examine two alternative control groups. First, we examine a wider control group of all non-merging stores in the same state as a divested store, rather than only those in ZIP3 regions without a divested store. Second, we examine a narrower control group that only contains "chain" stores, which we define as stores with a name that appears five or more times in the state. These stores are more likely to be similar to the divested stores, which were large supermarkets owned by major chains, and should exclude small ethnic grocery stores that are less comparable to the divested stores.

The disadvantage of the all-state control group is that it includes stores that are competing with the divested stores. These stores might change their behavior (reduce quality, increase prices) in response to actions taken by the divested stores. In principle, this control group could lead to attenuated treatment effects. However, including stores in the same ZIP3 may better capture localized demand or cost shocks.

We report estimates with the all-state control group, and only chain control group, in Table 13. We find almost identical estimates and confidence intervals for these control groups as in our baseline estimates in Table 3. For example, in our baseline estimates we find a 275% increase in reviews and 44% increase in the share of negative reviews, compared to 279% and 41% for the expanded control group and 280% and 42% for the narrowed control group. Our estimates for review topics are also robust to changes in the control group, and are reported in Table 14.

Table 13: Estimated divestiture effects with alternative control groups.

Control Group	Time Period	No. Reviews	Share Negative Reviews
All-state	Quarters 0-3	279%	41%
		[230%, 335%]	[21%, 65%]
	Quarters 4-7	67%	-30%
		[37%, 102%]	[-43%, -15%]
Chain stores	Quarters 0-3	280%	42%
		[230%, 338%]	[21%, 67%]
	Quarters 4-7	70%	-28%
		[38%, 110%]	[-42%, -12%]

Note: This table shows the results of the difference-in-differences regressions where the control group (1) contains all stores in the same state as a divested store, or (2) is restricted to chain stores. Each column represents a different dependent variable. Each column represents a different dependent variable; the table presents the estimated effect on the count of reviews, and the *share* of negative reviews, computed as the difference between estimated regression coefficients. The Quarters 0-3 and 4-7 rows show the estimated dynamic treatment effects, both point estimates and 95% confidence intervals, transformed to percentage units.

Table 14: Topics estimates with alternative control groups.

		Topic Share of Negative Reviews		
Control Group	Time Period	Prices	Products	Customer Service
All-state	Quarters 0-3	104%	19%	-21%
		[68%, 148%]	[8%, 30%]	[-28%, -14%]
	Quarters 4-7	40%	11%	-8%
		[17%, 68%]	[0%, 22%]	[-17%, 1%]
Chain stores	Quarters 0-3	103%	17%	-22%
		[66%, 147%]	[6%, 29%]	[-29%, -14%]
	Quarters 4-7	41%	11%	-10%
		[17%, 70%]	[0%, 23%]	[-19%, 0%]

Note: This table shows the results of the difference-in-differences regressions where the control group (1) contains all stores in the same state as a divested store, or (2) is restricted to chain stores. Each column represents a different dependent variable; the table presents the estimated effect on the *share* of the dependent variable within a broader category, computed as the difference between estimated regression coefficients. The Quarters 0-3 and 4-7 rows show the estimated dynamic treatment effects, both point estimates and 95% confidence intervals, transformed to percentage units.

B.3 Alternative clustering methods

All the confidence intervals we present are based on clustering by ZIP3, which is a proxy for the local geographical markets to which the treatments are assigned. In Table 15, we show that the confidence intervals are similar if we instead cluster by store location, by the location's chain,[32] or two-way (Cameron, Gelbach and Miller, 2011) by ZIP3 and chain.

Table 15: Confidence intervals under alternative clustering methods

	Clustering by			
Outcome	ZIP3	Chain	ZIP3 + Chain	Location
Review Count (Q0-3)	[225%, 333%]	[226%, 332%]	[230%, 327%]	[219%, 342%]
Review Count (Q4-7)	[38%, 108%]	[25%, 128%]	[29%, 122%]	[31%, 118%]
Share Negative Reviews (Q0-3)	[23%, 70%]	[16%, 80%]	[17%, 78%]	[21%, 73%]
Share Negative Reviews (Q4-7)	[-40%, -10%]	[-41%, -9%]	[-42%, -8%]	[-40%, -11%]
Prices (Share Neg Rev) (Q0-3)	[68%, 149%]	[86%, 124%]	[85%, 125%]	[67%, 150%]
Prices (Share Neg Rev) (Q4-7)	[17%, 70%]	[24%, 60%]	*	[11%, 79%]
Products (Share Neg Rev) (Q0-3)	[7%, 29%]	[9%, 26%]	[9%, 26%]	[7%, 29%]
Products (Share Neg Rev) (Q4-7)	[0%, 23%]	[3%, 20%]	[4%, 19%]	[1%, 23%]
Cust. Service (Share Neg Rev) (Q0-3)	[-28%, -14%]	[-26%, -18%]	[-25%, -18%]	[-29%, -14%]
Cust. Service (Share Neg Rev) (Q4-7)	[-19%, 0%]	[-17%, -2%]	[-17%, -2%]	[-18%, 0%]

Note: This table shows how the 95% confidence intervals around the central results of the divestiture treatment vary when standard errors are clustered by different methods. From left to right, the columns compare clustering by ZIP3; chain (approximated using string matching); two-way by ZIP3 and chain; and individual store location. For two-way clustering, * indicates the standard error is undefined due to a failure of positive-definiteness.

B.4 Assigning sentiment using review text

In our main analyses, we classify 1- and 2-star reviews as negative, and 3-, 4-, and 5-star reviews as non-negative. To validate our use of star ratings, we also use the `twitter-roberta-base-sentiment-latest` model with Hugging Face to assign a sentiment score to each review.[33] In this section, we instead use the "negativity" sentiment score between 0 and 1 based on review text in place of the star ratings. The negativity score is highly correlated with star rating, as shown in Table 16.

Accordingly, the difference-in-differences results (shown in Table 17) are extremely similar. In these analyses, since the negativity score is continuous, the outcome variable in the regressions is

[32] A location is assigned a chain based on string matching on the business name, if the name appears five or more times within a state. If we find multiple businesses at one location, we use the business with the the closest review to January 1, 2015.

[33] See `https://huggingface.co/cardiffnlp/twitter-roberta-base-sentiment-latest`.

Table 16: Review negativity score by star rating

Star rating	25th percentile	50th percentile	75th percentile	Mean
1.0	0.73	0.87	0.92	0.77
2.0	0.43	0.76	0.88	0.64
3.0	0.02	0.16	0.62	0.31
4.0	0.00	0.01	0.03	0.07
5.0	0.00	0.01	0.01	0.03

Note: This table shows the relationship between star rating and the negativity score (0–1), which is assigned solely using review text. Columns show the 25th, 50th, and 75th percentiles and mean of the negativity score among reviews with each star rating.

the negativity score (summed by store-quarter) rather than the count of 1 and 2 star reviews.

Table 17: Summary of difference-in-differences estimates, using negative sentiment scores.

		Topic Share of Negative Reviews		
Time Period	Share Negative Reviews	Prices	Products	Customer Service
Quarters 0-3	34%	108%	19%	-22%
	[14%, 57%]	[75%, 146%]	[8%, 30%]	[-29%, -14%]
Quarters 4-7	-24%	44%	12%	-10%
	[-37%, -8%]	[22%, 69%]	[1%, 23%]	[-20%, 0%]
Pre-mean	0.38	0.37	0.66	0.73

Note: This table shows the results of the difference-in-differences regressions using a continuous negative sentiment metric in place of star rating. Each column represents a different dependent variable; the table presents the estimated effect on the *share* of the dependent variable within a broader category, computed as the difference between estimated regression coefficients. The Quarters 0-3 and 4-7 rows show the estimated dynamic treatment effects, both point estimates and 95% confidence intervals, transformed to percentage units. The 'Pre-mean' row shows the ratio of means (mean of the dependent variable divided by the mean of the 'as share of' variable in the treatment group) in the quarter before divestiture.

B.5 Acquired firm analysis with alternative control group

Since Safeway operated stores in over 90% of 3-digit zip-codes in California before and after the merger, the estimates in Section 8 are primarily identified using information for the Safeway stores not in California and the stores owned by Delhaize in states with divestitures. We thus also estimate these models using a broader control group, containing all grocery stores not owned by the merging firm or a divestiture buyer located in the states with divestitures, as in Appendix B.2. These results are shown in Table 18; with much smaller standard errors, we find effectively no change in the number of reviews, share of negative reviews, and in the topics mentioned in negative reviews.

Table 18: Summary of difference-in-differences estimates for acquired stores, with all-state control group.

			Topic Share of Negative Reviews		
Time Period	No. Reviews	Share Negative Reviews	Prices	Products	Customer Service
Quarters 0-3	-3%	-8%	4%	2%	0%
	[-9%, 3%]	[-12%, -3%]	[-4%, 13%]	[-1%, 5%]	[-3%, 2%]
Quarters 4-7	1%	-5%	4%	0%	1%
	[-6%, 8%]	[-11%, 1%]	[-3%, 12%]	[-3%, 3%]	[-2%, 4%]
Pre-mean	1.40	0.51	0.34	0.61	0.74

Note: This table shows the results of the difference-in-differences regressions for the acquired store treatment and the all-state control group. Each column represents a different dependent variable; the table presents the estimated effect on either the dependent variable itself, or on the *share* of the dependent variable within a broader category, computed as the difference between estimated regression coefficients. The Quarters 0-3 and 4-7 rows show the estimated dynamic treatment effects, both point estimates and 95% confidence intervals, transformed to percentage units. The 'Pre-mean' row shows either the mean of the dependent variable in the treatment group in the quarter before the treatment, or the ratio of means (mean dependent variable divided by mean of 'as share of' variable).

B.6 Merged stores analysis (combining acquired and acquirer)

In Section 8, we restricted attention to acquired stores, which saw a change in ownership upon the acquisition. However, the merger may have affected incentives for both acquired and acquiring firms' stores. In this section, we examine effects for all merging firms' stores, combining acquired and acquiring firm's stores, in states with divestitures. We find similar null effects in Table 19. As in Section 8, the control group contains all stores in the same state, but different ZIP3, as a treated store.

We now re-estimate Table 19 using a broader control group, containing all grocery stores not owned by the merging firm or a divestiture buyer located in the states with divestitures. We obtain point estimates close to zero and narrower confidence intervals. The results are shown in Table 20.

B.7 Local Competitor Stores

In this section, we show that local competitor stores – chain stores in the same local market (ZIP3) as a divested store, but not involved in the merger – do not see detectable changes in reviews following a divestiture.

Specifically, we examine the effects of divestitures on local competitors: unrelated (non-divested, non-merging) chain stores in the same local market (ZIP3). (As in the main analysis, the control

Table 19: Summary of difference-in-differences estimates for merging (acquirer and acquired) stores.

			Topic Share of Negative Reviews		
Time Period	No. Reviews	Share Negative Reviews	Prices	Products	Customer Service
Quarters 0-3	21%	5%	-11%	-4%	0%
	[3%, 41%]	[-17%, 33%]	[-27%, 9%]	[-14%, 7%]	[-11%, 11%]
Quarters 4-7	5%	16%	-16%	-2%	0%
	[-14%, 30%]	[-8%, 47%]	[-29%, -2%]	[-13%, 11%]	[-11%, 13%]
Pre-mean	1.18	0.47	0.36	0.63	0.74

Note: This table shows the results of the difference-in-differences regressions for the merging store treatment. Each column represents a different dependent variable; the table presents the estimated effect on either the dependent variable itself, or on the *share* of the dependent variable within a broader category, computed as the difference between estimated regression coefficients. The Quarters 0-3 and 4-7 rows show the estimated dynamic treatment effects, both point estimates and 95% confidence intervals, transformed to percentage units. The 'Pre-mean' row shows either the mean of the dependent variable in the treatment group in the quarter before the treatment, or the ratio of means (mean dependent variable divided by mean of 'as share of' variable).

Table 20: Summary of difference-in-differences estimates for all merging stores, with all-state control group.

			Topic Share of Negative Reviews		
Time Period	No. Reviews	Share Negative Reviews	Prices	Products	Customer Service
Quarters 0-3	-1%	-7%	0%	1%	0%
	[-6%, 4%]	[-11%, -3%]	[-6%, 7%]	[-2%, 3%]	[-3%, 2%]
Quarters 4-7	2%	-5%	-1%	-1%	0%
	[-4%, 8%]	[-10%, 0%]	[-7%, 5%]	[-4%, 1%]	[-3%, 2%]
Pre-mean	1.18	0.47	0.36	0.63	0.74

Note: This table shows the results of the difference-in-differences regressions for the merging store treatment and the all-state control group. Each column represents a different dependent variable; the table presents the estimated effect on either the dependent variable itself, or on the *share* of the dependent variable within a broader category, computed as the difference between estimated regression coefficients. The Quarters 0-3 and 4-7 rows show the estimated dynamic treatment effects, both point estimates and 95% confidence intervals, transformed to percentage units. The 'Pre-mean' row shows either the mean of the dependent variable in the treatment group in the quarter before the treatment, or the ratio of means (mean dependent variable divided by mean of 'as share of' variable).

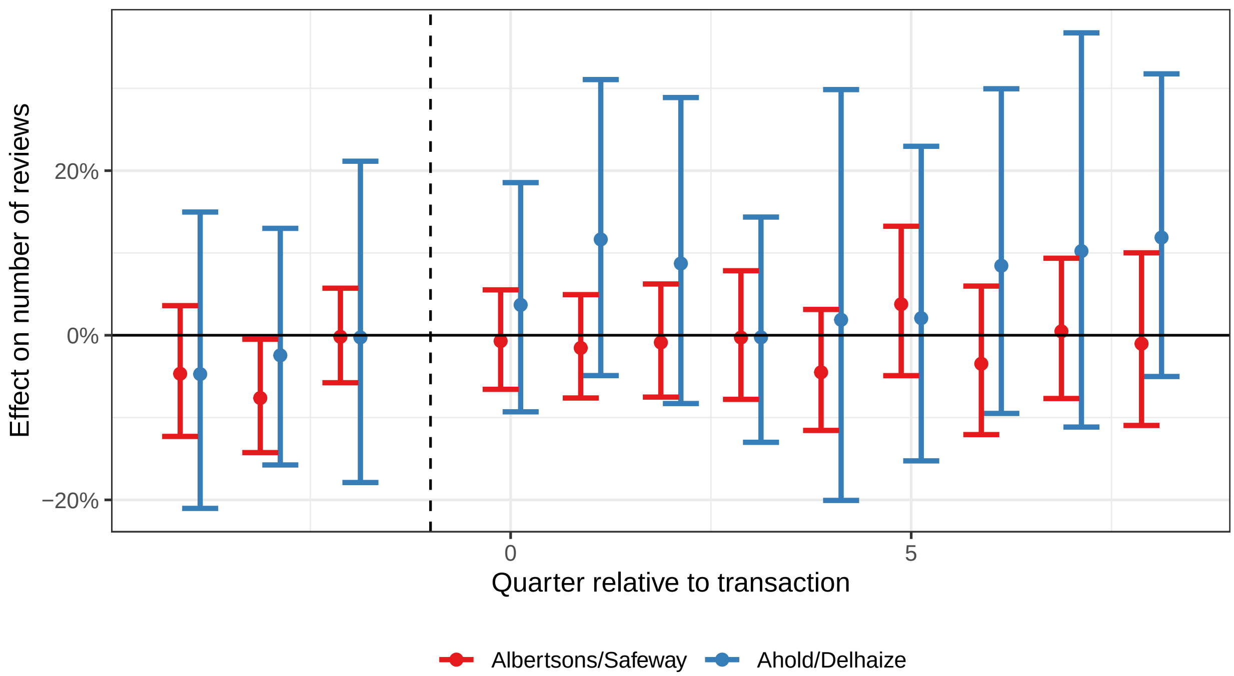

Figure 14: Effect of divestiture on local competitors' number of reviews

group contains non-merging stores in the same state, but different ZIP3, as a treated store.) We do not find any significant effect on the number or topics of reviews. We find a significant but small increase in the share of negative reviews. We report the event study plots in Figure 14 and Figure 15 and the regression coefficients in Table 21.

However, we do not find effects on the share of negative reviews to be robust to the control group used. In Table 22, we show that the positive effect of a local divestiture on the negative share of reviews is no longer significant when the control group is limited to chain stores.

B.8 A&P stores analysis restricted to three large acquirers

The A&P stores analysis in Section 8 looked at former A&P stores that remained in the grocery business, whether as a major supermarket brand or not. In Table 23, we restrict the treatment group to the former A&P stores acquired by Albertsons (owner of Acme), Ahold/Delhaize (owner of Stop & Shop), and ShopRite. (The other A&P store locations are excluded from both treatment and control groups.)

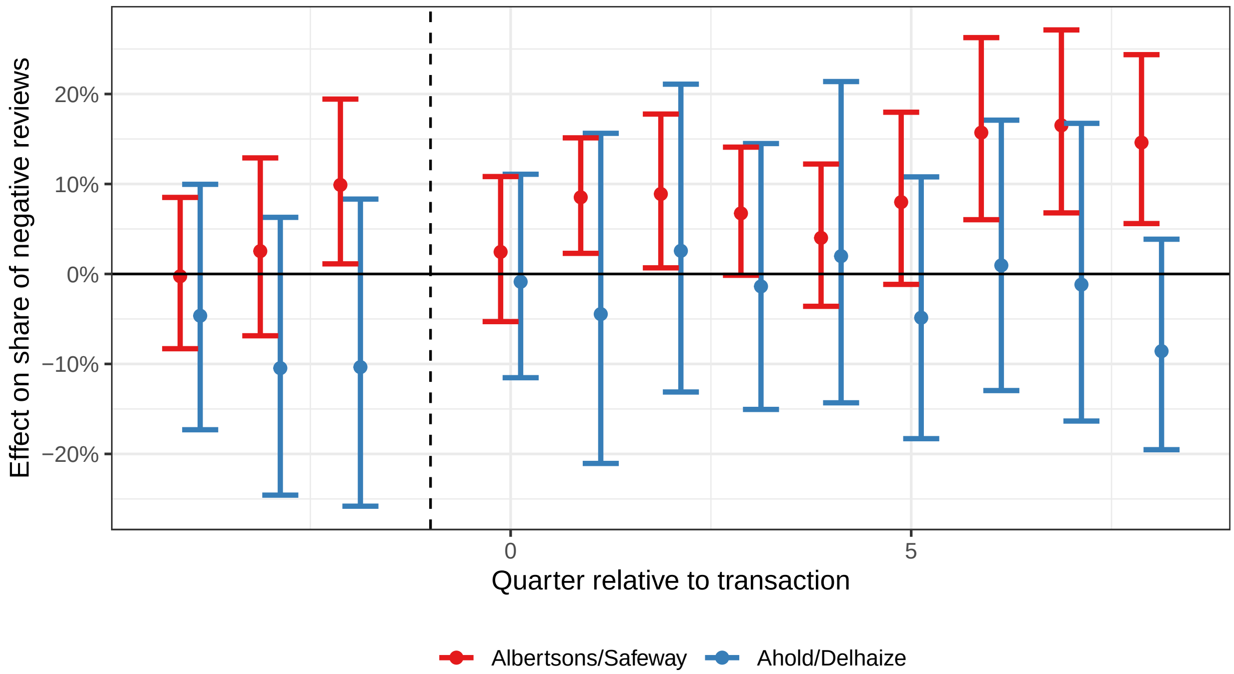

Figure 15: Effect of divestiture on local competitors' share of negative reviews (event study plot)

Table 21: Summary of difference-in-differences estimates for local competitors of divested stores.

			Topic Share of Negative Reviews		
Time Period	No. Reviews	Share Negative Reviews	Prices	Products	Customer Service
Quarters 0-3	0%	6%	-3%	-2%	-1%
	[-5%, 5%]	[1%, 11%]	[-8%, 3%]	[-5%, 1%]	[-4%, 2%]
Quarters 4-7	0%	9%	-3%	-1%	-3%
	[-7%, 7%]	[2%, 16%]	[-8%, 3%]	[-4%, 2%]	[-5%, 0%]
Pre-mean	1.66	0.30	0.39	0.64	0.75

Note: This table shows the results of the difference-in-differences regressions for the local competitor treatment. Each column represents a different dependent variable; the table presents the estimated effect on either the dependent variable itself, or on the *share* of the dependent variable within a broader category, computed as the difference between estimated regression coefficients. The Quarters 0-3 and 4-7 rows show the estimated dynamic treatment effects, both point estimates and 95% confidence intervals, transformed to percentage units. The 'Pre-mean' row shows either the mean of the dependent variable in the treatment group in the quarter before the treatment, or the ratio of means (mean dependent variable divided by mean of 'as share of' variable).

Table 22: Summary of difference-in-differences estimates for local competitors of divested stores.

			Topic Share of Negative Reviews		
Time Period	No. Reviews	Share Negative Reviews	Prices	Products	Customer Service
Quarters 0-3	4%	3%	-3%	-2%	-1%
	[-2%, 10%]	[-2%, 10%]	[-9%, 3%]	[-5%, 1%]	[-4%, 2%]
Quarters 4-7	3%	5%	-2%	-1%	-2%
	[-5%, 11%]	[-2%, 13%]	[-8%, 4%]	[-4%, 3%]	[-5%, 1%]
Pre-mean	1.66	0.30	0.39	0.64	0.75

Note: This table shows the results of the difference-in-differences regressions for the local competitor treatment, when the control group is restricted to chain stores. Each column represents a different dependent variable; the table presents the estimated effect on either the dependent variable itself, or on the *share* of the dependent variable within a broader category, computed as the difference between estimated regression coefficients. The Quarters 0-3 and 4-7 rows show the estimated dynamic treatment effects, both point estimates and 95% confidence intervals, transformed to percentage units. The 'Pre-mean' row shows either the mean of the dependent variable in the treatment group in the quarter before the treatment, or the ratio of means (mean dependent variable divided by mean of 'as share of' variable).

Table 23: Summary of difference-in-differences estimates for post-A&P stores purchased by three large acquirers.

			Topic Share of Positive Reviews		
Time Period	No. Reviews	Share Negative Reviews	Prices	Products	Customer Service
Quarters 0-3	102%	-34%	24%	5%	59%
	[34%, 204%]	[-45%, -20%]	[-51%, 213%]	[-56%, 153%]	[-17%, 202%]
Quarters 4-7	89%	-38%	61%	12%	65%
	[20%, 199%]	[-50%, -23%]	[-50%, 415%]	[-60%, 215%]	[-28%, 281%]
Pre-mean	0.56	0.71	0.19	0.37	0.22

Note: This table shows the results of the difference-in-differences regressions for the A&P stores subsequently acquired by Albertsons, Ahold/Delhaize, or ShopRite. Each column represents a different dependent variable; the table presents the estimated effect on either the dependent variable itself, or on the *share* of the dependent variable within a broader category, computed as the difference between estimated regression coefficients. The Quarters 0-3 and 4-7 rows show the estimated dynamic treatment effects, both point estimates and 95% confidence intervals, transformed to percentage units. The 'Pre-mean' row shows either the mean of the dependent variable in the treatment group in the quarter before treatment, or the ratio of means (mean dependent variable divided by mean of 'as share of' variable).

C Additional Results

C.1 Breakdown of divestiture results by treatment

Table 24 and Table 25 show how our headline difference-in-differences results break down across four treatments: (1) overall, (2) stores divested to Haggen, (3) stores divested to Haggen that did not subsequently close, and (4) stores not divested to Haggen (including all Ahold/Delhaize divestitures).

Table 24: Summary of difference-in-differences estimates, aggregated to four-quarter time periods.

	Effect		
Treatment	Quarters 0-3	Quarters 4-7	Pre-mean
Dep. var.: Review count			
Overall	275% [225%, 333%]	69% [38%, 108%]	0.88
Haggen (all)	288% [232%, 352%]	69% [34%, 113%]	1.23
Haggen (non-closed)	292% [231%, 364%]	87% [48%, 137%]	1.25
non-Haggen	208% [107%, 359%]	66% [5%, 161%]	0.34
Dep. var.: Negative review count, as share of review count			
Overall	44% [23%, 70%]	-27% [-40%, -10%]	0.45
Haggen (all)	37% [17%, 59%]	-40% [-49%, -29%]	0.48
Haggen (non-closed)	39% [18%, 64%]	-38% [-50%, -25%]	0.46
non-Haggen	114% [-3%, 375%]	93% [-13%, 331%]	0.27
Dep. var.: Prices topic (neg.), as share of negative review count			
Overall	104% [68%, 149%]	41% [17%, 70%]	0.39
Haggen (all)	110% [70%, 160%]	39% [14%, 70%]	0.39
Haggen (non-closed)	108% [63%, 166%]	36% [9%, 70%]	0.39
non-Haggen	70% [3%, 180%]	45% [-13%, 142%]	0.35
Dep. var.: Products topic (neg.), as share of negative review count			
Overall	17% [7%, 29%]	11% [0%, 23%]	0.66
Haggen (all)	16% [6%, 27%]	9% [-1%, 21%]	0.68
Haggen (non-closed)	16% [5%, 28%]	9% [-2%, 21%]	0.68
non-Haggen	37% [-15%, 122%]	29% [-21%, 110%]	0.56
Dep. var.: Customer service topic (neg.), as share of negative review count			
Overall	-22% [-28%, -14%]	-10% [-19%, 0%]	0.73
Haggen (all)	-24% [-30%, -18%]	-10% [-18%, -1%]	0.74
Haggen (non-closed)	-26% [-32%, -18%]	-12% [-20%, -2%]	0.76
non-Haggen	5% [-41%, 87%]	2% [-43%, 83%]	0.63

Note: This table shows the results of the difference-in-differences regressions. The dependent variable is presented at the beginning of each section of the table; the table presents the estimated effect on either the dependent variable itself, or on the *share* of the dependent variable within a broader category, computed as the difference between estimated regression coefficients. The Treatment column gives the treatment group: all divested stores (Overall) or selected divested stores. The Quarters 0-3 and 4-7 columns show the estimated dynamic treatment effects, both point estimates and 95% confidence intervals, transformed to percentage units. The 'Pre-mean' column shows either the mean of the dependent variable in the treatment group in the quarter before divestiture, or the ratio of means (mean dependent variable divided by mean of 'as share of' variable).

Table 25: User composition effects of divestitures, difference-in-differences estimates.

Dep. var.	Treatment	Effect Quarters 0-3	Effect Quarters 4-7	Pre-mean
New Grocery Reviewer	Overall	7% [-5%, 20%]	-1% [-11%, 10%]	0.58
	Haggen (all)	10% [-4%, 26%]	0% [-11%, 13%]	0.56
	Haggen (non-closed)	8% [-8%, 26%]	-2% [-15%, 13%]	0.57
	non-Haggen	-6% [-21%, 13%]	-5% [-20%, 12%]	0.70
Never Yelp Elite	Overall	0% [-5%, 5%]	-4% [-10%, 2%]	0.83
	Haggen (all)	0% [-5%, 6%]	-4% [-11%, 3%]	0.83
	Haggen (non-closed)	0% [-7%, 8%]	-4% [-12%, 5%]	0.83
	non-Haggen	-5% [-17%, 8%]	-3% [-16%, 12%]	0.85
Not-Recommended Flag	Overall	-13% [-36%, 19%]	-16% [-40%, 18%]	0.16
	Haggen (all)	-19% [-42%, 14%]	-24% [-47%, 9%]	0.16
	Haggen (non-closed)	-12% [-38%, 24%]	-19% [-44%, 16%]	0.15
	non-Haggen	23% [-49%, 194%]	28% [-49%, 226%]	0.15
Only Yelp Review	Overall	50% [-34%, 240%]	22% [-45%, 168%]	0.03
	Haggen (all)	238% [-18%, 1296%]	111% [-41%, 655%]	0.01
	Haggen (non-closed)	203% [-24%, 1119%]	85% [-48%, 554%]	0.01
	non-Haggen	-38% [-75%, 52%]	-24% [-68%, 84%]	0.12
Only Grocery Review	Overall	11% [-6%, 31%]	7% [-9%, 26%]	0.41
	Haggen (all)	12% [-6%, 35%]	6% [-12%, 27%]	0.40
	Haggen (non-closed)	14% [-5%, 38%]	8% [-10%, 28%]	0.39
	non-Haggen	6% [-25%, 49%]	10% [-24%, 59%]	0.48
Not a Top 10% Reviewer	Overall	1% [-8%, 10%]	-4% [-12%, 5%]	0.72
	Haggen (all)	0% [-9%, 11%]	-5% [-14%, 4%]	0.72
	Haggen (non-closed)	-2% [-12%, 10%]	-7% [-17%, 4%]	0.73
	non-Haggen	5% [-12%, 26%]	4% [-15%, 26%]	0.73
Long Review (tweet)	Overall	9% [1%, 18%]	5% [-3%, 13%]	0.68
	Haggen (all)	8% [-1%, 17%]	3% [-6%, 12%]	0.69
	Haggen (non-closed)	11% [2%, 21%]	6% [-4%, 17%]	0.67
	non-Haggen	16% [-6%, 42%]	17% [-4%, 42%]	0.61
Long Review (median)	Overall	15% [1%, 32%]	9% [-8%, 28%]	0.50
	Haggen (all)	10% [-3%, 25%]	4% [-12%, 22%]	0.52
	Haggen (non-closed)	10% [-6%, 28%]	3% [-15%, 25%]	0.52
	non-Haggen	53% [-6%, 149%]	46% [-12%, 141%]	0.36
Readable Review (median)	Overall	8% [-4%, 21%]	9% [-2%, 22%]	0.50
	Haggen (all)	11% [-2%, 25%]	9% [-3%, 23%]	0.50
	Haggen (non-closed)	7% [-7%, 23%]	6% [-8%, 21%]	0.51
	non-Haggen	-6% [-29%, 25%]	5% [-21%, 40%]	0.55

Note: This table shows the effect of divestiture on the dependent variable ('Dep. var.' column) as a share of the count of reviews, computed as the difference between estimated regression coefficients. The Treatment column gives the treatment group: all divested stores (Overall) or selected divested stores. The Quarters 0-3 and 4-7 columns show the estimated dynamic treatment effects, both point estimates and 95% confidence intervals, transformed to percentage units. The 'Pre-mean' column shows the ratio of the mean of the dependent variable to the mean count of reviews, in the treatment group in the quarter before divestiture.

D Manual Review of Topic Scores

To assess the accuracy of the zero-shot topic scores, we asked three coders to independently, manually assess a sample of 320 reviews. All three coders were undergraduate summer interns at the FTC. We stratified the sample using the interaction of:

1. Relevant transaction (Ahold/Delhaize or Albertsons/Safeway)
2. Divested store (treatment group), same ZIP3 as a divested store, merging store, or control group
3. Before or after transaction
4. Negative/neutral (1-3 stars) or positive (4-5 stars).

For each review in the sample, each coder had to judge the review's relevance to each topic, recording each as yes, no, or unsure.

We find two patterns that suggest that the topic scores perform well. First, when the topic score is higher, the coders mark the review as 'relevant' at a higher rate, with the highest correlations for the "prices" topic. We report rank correlations in Table 26; we find correlations between the number of coders coding a review as relevant and the topic score of 0.79 for prices, 0.66 for customer service, and 0.52 for products. The correlations between the topic score and individual reviewers range from 0.72 to 0.76 for prices, 0.56 to 0.64 for customer service, and 0.47 to 0.50 for products. These correlations are only somewhat lower than between individual reviewers, which range between 0.79 and 0.85 for prices, 0.71 and 0.78 for customer service, and 0.73 and 0.78 for products.

Second, the coders are more likely to mark the review as 'unsure' or disagree among themselves when the zero-shot model is uncertain. We examine the relationship between the zero-shot topic score and the average manual review share over the three coders in Figure 16. The share of reviews for which the manual coders disagree or mark unsure is the highest when the topic score ranges between 0.4 and 0.6.

Table 26: Rank correlations between topic scores and manual review.

Topic	R1 Yes	R2 Yes	R3 Yes	# Yes
Customer Service	0.611	0.560	0.638	0.664
Prices	0.762	0.752	0.724	0.792
Products	0.479	0.496	0.473	0.527

Note: For each topic, we compute the Spearman rank correlation between the zero-shot topic score in $[0, 1]$ and output of manual review. In the first three columns, we compute separate correlations for the three reviewers (0 = Not Relevant, 1 = Relevant). In the fourth column, we compute compute the correlation between the topic score and the sum of the manual reviews (number of reviewers marking Relevant).

Figure 16: Comparison of topic scores and manual review.

Note: Generated using stratified sample of 320 reviews from the following states and time periods: AZ, CA, MT, NV, OR, TX, WA, WY (2013–17); DE, MA, MD, NY, PA, VA, WV (2014–19). Topic scores were calculated for the three topics and range from 0 to 1. Disagree/Unsure indicates that the three reviewers were not unanimous in their assessment, or at least one of the three reviewers marked ‘unsure’.