# Hypermarket and Low-Priced Competition Event Study Analysis

Daniel Posthumus

December 20, 2024

**Abstract**

In this event study, we examine price heterogeneity across retail stations in the Bay Area. We find that there are considerable price differences between 1) unbranded stations, 2) branded non-hypermarket stations, and 3) hypermarkets–which, defined by the California Energy Commission (CEC), are large retailers that try to compete in retail gas on price, undercutting competition. The hypermarkets appearing in the Bay Area retail data are Costco, Kroger, and Safeway. We then try to understand competition between low-priced competitors (unbranded stations and hypermarkets) and branded stations through investigating the effect of the entry of low-priced competitors on nearby stations through an event study method. Exclusively focusing on the Bay Area because of computational expense, we find that the entry of a low-priced competitor has no clear effect on nearby stations' prices in the year after the entrance, which suggests market segmentation and that consumers' gasoline preferences aren't elastic between different types of gasoline stations. We hope to expand this analysis to all of California, at which point the results may change. We also want to investigate and better understand consumers' elasticities.

## 1 Introduction

**Central question:** How do gas stations compete on retail prices? One way to understand this is to examine the effects of the entry of lower-priced gas stations into geographic markets on nearby stations' prices; do these entries compel nearby stations to lower prices to compete with the new station? Is there a difference if the lower-priced gas stations are branded or unbranded?

Due to computational expense, I was limited to the Bay Area in my analysis. Originally, I analyzed both Los Angeles City and the Bay Area. Because of the relative lack of hypermarkets in Los Angeles, however, I restricted my analysis to the Bay Area. Any analysis I hereafter present was calculated/executed focusing on the Bay Area exclusively.

## 2 Defining Hypermarket Stations

Working off the CEC's definition of a gasoline hypermarket, I classify all stations identifying their store band as one of Costco, Kroger, or Safeway (Walmart, another nationally prominent hypermarket gasoline seller, doesn't appear in the data for the Bay Area) as hypermarkets.

## 3 Cross-Sectional Evidence of Price Differences

Here is a table containing basic information about the distribution of average monthly prices for 2018 in the Bay Area. I restricted the sample to be a fully balanced subpanel, i.e., each gas station included below was present in every month of 2018. This ensures any disparities we see when averaging are because of differences in station characteristics, *not* the composition of the data changing over time. Note that the 'Count' column includes only stations in the balanced subsample, not the whole sample, and that stations which change ownership during 2018 are *not* included in the fully balanced subpanel.

Table 1: Bay Area Station Count and Distribution of Retail Prices in 2018

| Designation | Count | Minimum | 25th Percentile | Median | 75th Percentile | Maximum |
|---|---|---|---|---|---|---|
| Branded | 675 | 2.77 | 3.45 | 3.66 | 3.80 | 4.80 |
| Unbranded | 97 | 2.84 | 3.34 | 3.52 | 3.66 | 4.30 |
| Hypermarket | 21 | 2.77 | 3.17 | 3.35 | 3.46 | 3.72 |

Using the same sample (the fully balanced 2018 subpanel), we can visualize geographic heterogeneity in retail prices with the following heatmap, which has the Bay Area's county boundaries imposed on it.

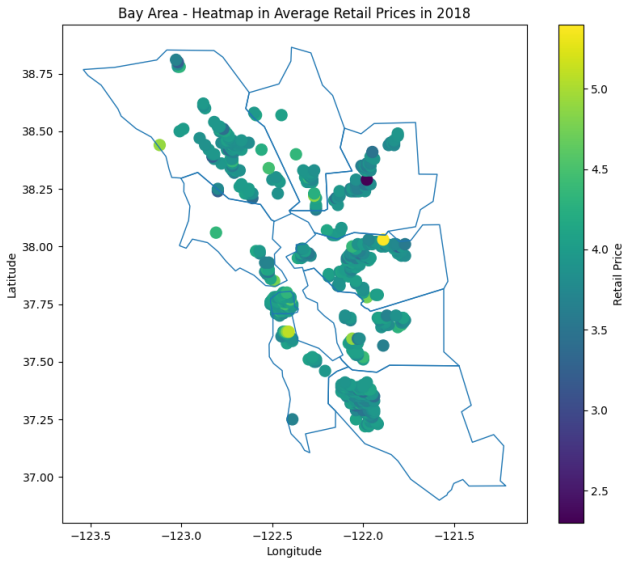



Figure 1: Average 2018 Retail Prices by Gas Station in the Bay Area

In addition to location, we can also analyze the *type* of gas station as a source of variation in retail price. This is demonstrated in the graph below for the Bay Area. These average prices are calculated with the full, unbalanced panel of data:

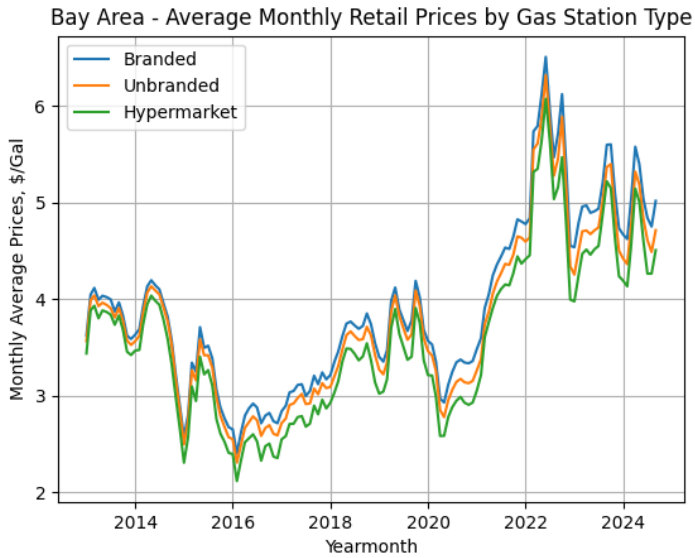



Figure 2: Average Retail Prices by Station Type for the Bay Area

One way we can see if firms are competing on price is to check if there is an association between the distance between a station and low-priced competitors and a station's price. To do so, I estimate the following equation using OLS:

$$\text{price}_{it} = \lambda_t + \beta_1(\text{distance\_from\_nearest\_low\_priced\_competitor})_{it} + \epsilon_{it} \quad (1)$$

for station $i$ at time $t$. $\lambda_t$ is time fixed-effects and distance_from_nearest_low_priced_competitor is the distance a station is from the nearest either hypermarket or unbranded market, depending on the model. Intuitively, we would expect $\beta_1$ to be positive; the further away a station is from a low-priced competitor, the less elastic consumers' demand will be. This would result in stations having higher prices than if they knew consumers could easily choose a nearby low-priced competitor. I report the results from this regression here.

While the results match our expectations, they suffer from a key bias: if low-priced competitor stations tend to be in poorer parts of the Bay Area, where gas prices are lower, then the strong significant positive coefficient on distance from these competitors would just reflect the geographic distribution of low-priced competitors and would not be any kind of signal for competition on price. This motivates us to go deeper into competition through an event study, which follows in the next sections. Because of computational expense, I randomly selected 30% of all observations in the sample and used those observations for this OLS estimation.

Table 2: Results of Regression by Type of Low-Priced Competitor

| Statistic | Unbranded Stations Model | Hypermarkets Model |
|---|---|---|
| Coefficient | 0.379 | 0.361 |
| Standard Error | 0.004 | 0.001 |
| R-Squared | 0.944 | 0.947 |
| Observations | 982494 | 982494 |

# 4 Defining Events

Broadly, there are 3 categories of events we're interested in:

1. **Brand Change**

   A gas station 'remains' in a physical location where a gas station was previously located; now, however, it operates under a different **Store** Brand.

2. **Entry From Nothing**

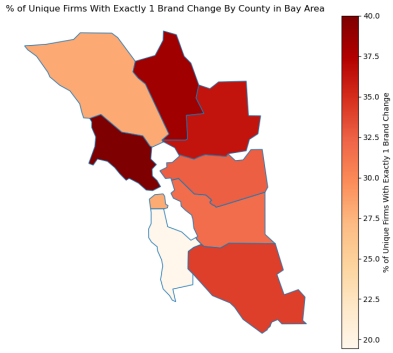
   A gas station opens at a physical location where there was no gas station previously.

3. **Exit to Nothing**

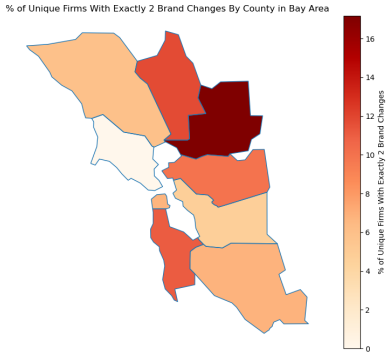
   A gas station closes, and at its physical location no new gas station opens to take its place.

The importance of physical location is a unique feature of the gasoline market; the high fixed cost of opening a gas station (the tank, the pumps, regulatory approvals, etc.) make ownership changes at existing gas station locations much more common than entries from or exits to nothing. We can visualize the relative frequency of these events through what share of gas stations undergo them in a particular geographic area. For example, below I create heatmaps which map the percentage of *all* gas stations that undergo each of these events above, on a county-level in the Bay Area.[1]

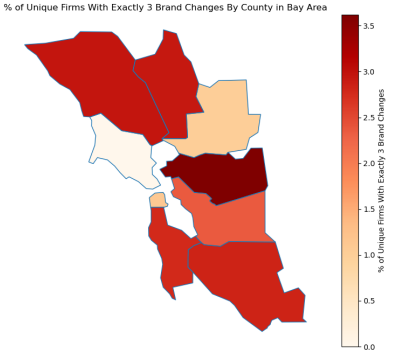
[1]**Note**: Some firms undergo as many as five brand changes in the period of the data; to avoid confusion, I separate out these stations into groups denoted by the *exact* number of brand changes they undergo. Therefore, below, when I write 'Exactly 1 Brand Change', this doesn't include stations that underwent more than one brand change, only those that underwent *exactly* one.



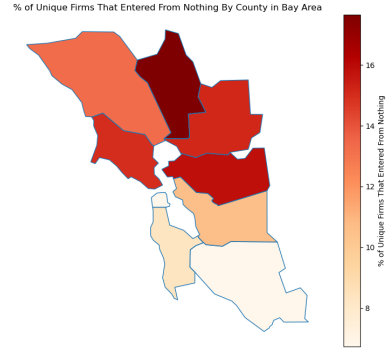

(a) Exactly 1 Brand Change



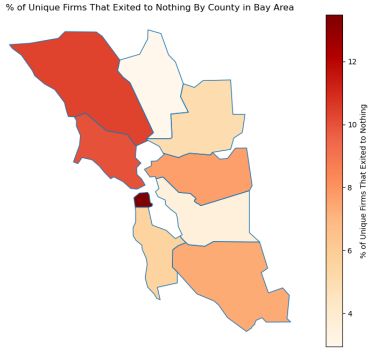

(b) Exactly 2 Brand Change




(c) Exactly 3 Brand Change




(d) Entry From Nothing




(e) Exit to Nothing

Figure 3: County-Level Heatmaps of Event Relative Frequency in Bay Area

Another way to visualize event occurrence, this time capturing intra-county variance, is through a simple scatterplot of geographic coordinates *on the gas station-level*, such as the following which I produce for the Bay Area. These make it clear that events occur throughout the region, although there does appear to be bias in San Jose, where entries from nothing are less common than in the rest of the region (hence the lack of blue points in San Jose).

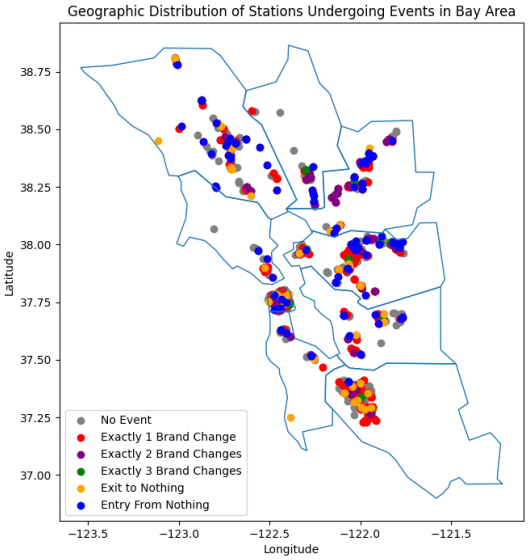



Figure 4: Bay Area Geographic Distribution of Event Occurrence

# 5 Event Study Analysis

## 5.1 Summary Statistics

Below are tables indicating event summary statistics, focusing on low-priced competitors. The table below confirms that hypermarkets sell at a significantly lower price immediately after entering.

Table 3: Bay Area - Summary Statistics for Hypermarkets that Entered From Nothing and Nearby Stations

| Category | Count | Pre-Event Average Price | Post-Event Average Price |
|---|---|---|---|
| Hypermarkets that Entered From Nothing | 5 | N/A | 3.85 |
| Stations Within 3 Mile(s) | 77 | 3.83 | 4.62 |
| Stations Within 3-10 Mile(s) | 242 | 3.79 | 4.54 |
| Stations Within 10-20 Mile(s) | 605 | 3.80 | 4.56 |

Next, I investigate the average prices before and after the brand change when a station transforms from a branded, non-hypermarket station to a hypermarket station. This confirms our intuition, again, as the station went from nearly exactly matching near and far stations in prices to being very depressed in prices after the brand change.

Table 4: Bay Area - Summary Statistics for Stations that Underwent Branded to Hypermarket Changes and Nearby Stations

| Category | Count | Pre-Event Average Price | Post-Event Average Price |
|---|---|---|---|
| Stations that Underwent Branded to Hypermarket Changes | 1 | 3.94 | 3.44 |
| Stations Within 3 Mile(s) | 39 | 3.91 | 3.86 |
| Stations Within 3-10 Mile(s) | 48 | 3.94 | 3.88 |
| Stations Within 10-20 Mile(s) | 42 | 3.97 | 3.99 |

Below, I investigate the entry from nothing of another type of low-priced competitor–this time unbranded gasoline stations. It's clear that unbranded stations average lower prices after entering, while there's also some suggestion that they're located in lower-priced parts of the Bay Area, as nearby stations sell for less than stations 3-10 miles away.

Table 5: Bay Area - Summary Statistics for Unbranded Stations that Entered From Nothing and Nearby Stations

| Category | Count | Pre-Event Average Price | Post-Event Average Price |
|---|---|---|---|
| Unbranded Stations that Entered From Nothing | 33 | N/A | 3.82 |
| Stations Within 3 Mile(s) | 612 | 3.68 | 4.02 |
| Stations Within 3-10 Mile(s) | 1559 | 3.73 | 4.05 |
| Stations Within 10-20 Mile(s) | 2815 | 3.72 | 4.04 |

Below, I investigate another type of event, a brand change when a branded, non-hypermarket station changes brands to become an unbranded station. Prior to the brand change, the station's prices were very similar to stations near and far, while afterwards they were very depressed relative to other stations' prices.

Table 6: Bay Area - Summary Statistics for Stations that Underwent Branded to Unbranded Changes and Nearby Stations

| Category | Count | Pre-Event Average Price | Post-Event Average Price |
|---|---|---|---|
| Stations that Underwent Branded to Unbranded Changes | 47 | 3.55 | 4.23 |
| Stations Within 3 Mile(s) | 1031 | 3.56 | 4.46 |
| Stations Within 3-10 Mile(s) | 2552 | 3.59 | 4.47 |
| Stations Within 10-20 Mile(s) | 4171 | 3.59 | 4.48 |

These analyses are rather limited in that they are static and do not reveal the dynamics of what is happening in the short term (within a year) of event occurrence. For this, I plot below the *unweighted* average prices of stations by their proximity (within 3 miles, within 3-10 miles, and within 10-20 miles) to a station undergoing an event.

It doesn't appear there is a substantial difference in price dynamics for stations near or far from the entering hypermarket. Interestingly, it appears that hypermarkets converge in price with their competitors about 40 weeks after entering. Entering hypermarkets' prices also behave similarly to other stations, just at a lower magnitude.


Figure 5: Hypermarket Entries From Nothing

The below graph, capturing the average prices for the single instance of a gas station changing from branded, non-hypermarket to a hypermarket, suggests that the hypermarket's price followed its competitors (as well as the further away 'control' group) very closely, just at a slightly depressed level. Here, unlike for hypermarket entries from nothing, it does appear there is some correlation between distance from the brand changing station

and price difference, as the further away a station is, the greater its price appears to be–a relationship that holds throughout the first 52 weeks after the brand change occurs.

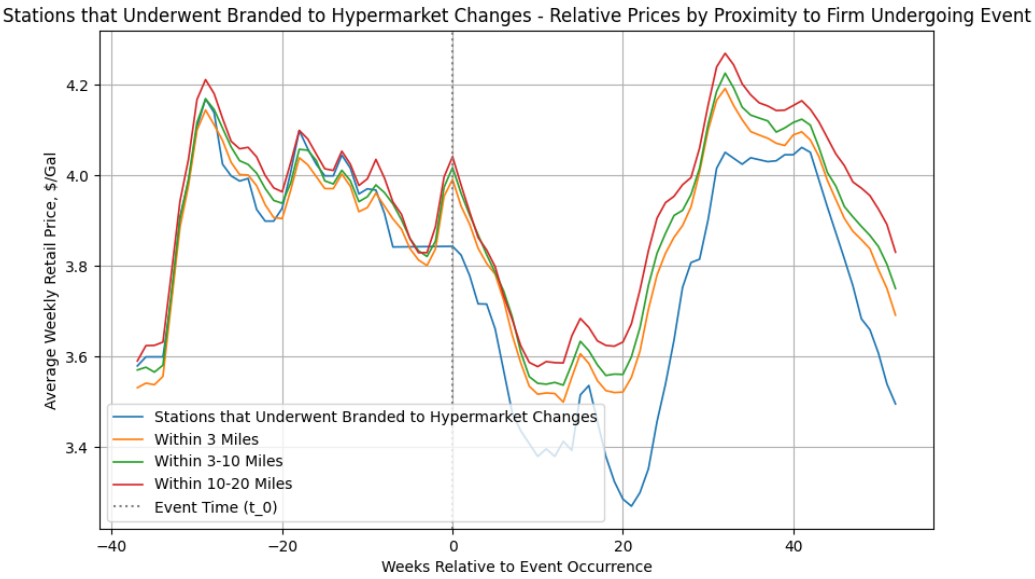



Figure 6: Branded to Hypermarket Brand Changes

The graph below shows that stations which changed from branded, non-hypermarket to unbranded were nearly indistinguishable in price from stations nearby and further away. In the 15-20 weeks before changing to unbranded, though, it appears the branded stations were notably increasing prices.

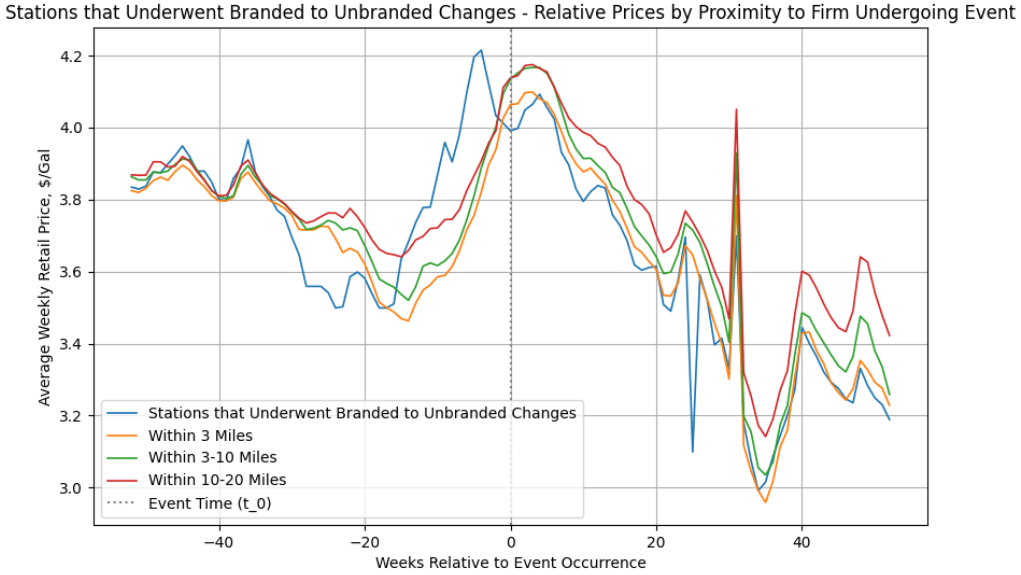



Figure 7: Branded to Unbranded Brand Changes

## 5.2 Difference-in-Difference Approach

Another way to understand price competition is through a Difference-in-Difference (DiD) framework, in which the entry of a low-priced competitor station (either a hypermarket or an unbranded station) is the treatment; then,

we can model the treatment effects over time. Specifically, I estimate the following equation using OLS:

$$\text{price}_{it} = \alpha_i + \lambda_t + \lambda_{\text{relative},t} + \delta \text{within\_3\_mile}_i + \sum_{l=0} \beta_l (\text{relative\_week}_{it} \times \text{within\_3\_mile}_i) + \epsilon_{it} \tag{2}$$

where $\alpha_i$ consists of station location fixed effects for station $i$, $\lambda_t$ is year-week time fixed effects for year-week $t$, $\lambda_{\text{relative},t}$ is relative week fixed-effects, relative_week is the week relative to each event occurrence (relative_week $\in \{-52, 52\}$), and within_3_mile is a dummy equal to one if station $i$ is within three miles of the station undergoing the event being analyzed. I manually fix $\beta = 0$ for the week *immediately* preceding the event occurring to facilitate interpretation of post-event treatment effects. **Note**: stations between 3 and 10 miles away from the station undergoing the event are used as the 'control' group. For the following analyses, I use heteroskedastic-robust standard errors.

In this estimated equation, the interaction term for each relative week is the treatment effect (i.e., the effect of having either a hypermarket or unbranded station enter a geographic market on other stations' prices). To get a sense of the effect of lower-priced competitors' entries, we can compare these effects across time by plotting the estimated treatment effects for each relative week. Below, I have included the 95% confidence bars for each interaction term. I restricted my sample to a fully balanced panel, i.e., restricting to only gas stations that have non-missing observations for each of the 52 weeks before and the 52 weeks after the event occurred.

First, I estimated this equation focusing on hypermarkets entering from nothing. Recall there are 5 instances of this event in the Bay Area, while there is only one instance of a station changing from being branded, non-hypermarket to being a hypermarket. Therefore, I don't include an estimate for the latter type of entry.

The beta coefficients basically suggest that a hypermarket entering had no effect on stations within three miles, with maybe some positive effect around 35-40 weeks after entering. This basically matches our expectations based on Figure 5, where it appears stations at least 10 miles away from the entering hypermarket had lower prices than stations either very close (less than three miles away) or somewhat close (3-10 miles away).

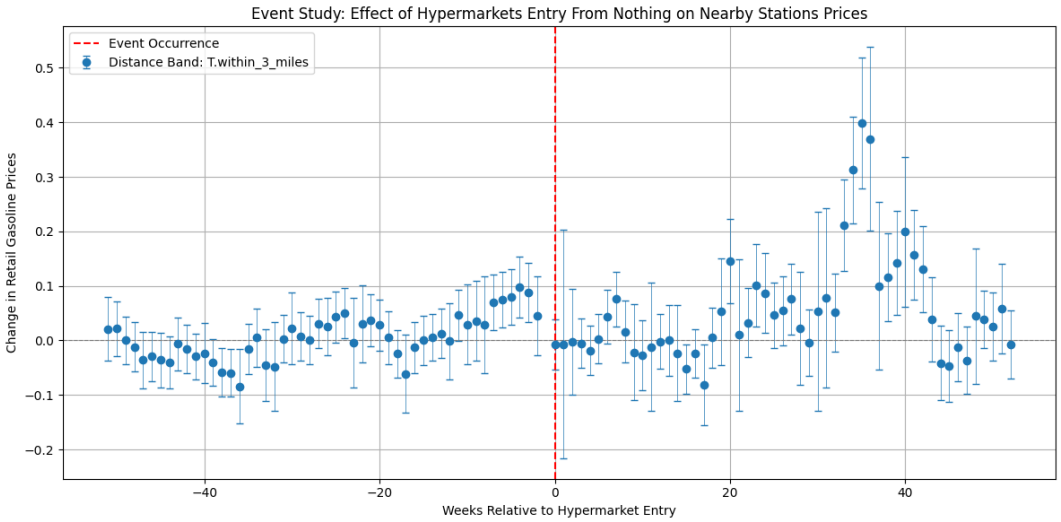



Figure 8: DiD Estimates for Hypermarket Entry From Nothing

Since we're interested in price competition, I was also interested if there was any evidence of nearby stations reacting to unbranded stations entering from nothing. The advantage of this type of event is that the sample size is greater, while it's more difficult to interpret since the whole basis for branded stations charging more is that consumers have some kind of preference for branded gasoline and thus have a higher willingness-to-pay for it. I first focus on unbranded stations entering from nothing. These are noisy, with some evidence of strong negative effects on price between 10 and 20 weeks after entering, although it doesn't appear we can draw any sorts of conclusions about the entrances' effects throughout the first year after it occurs.

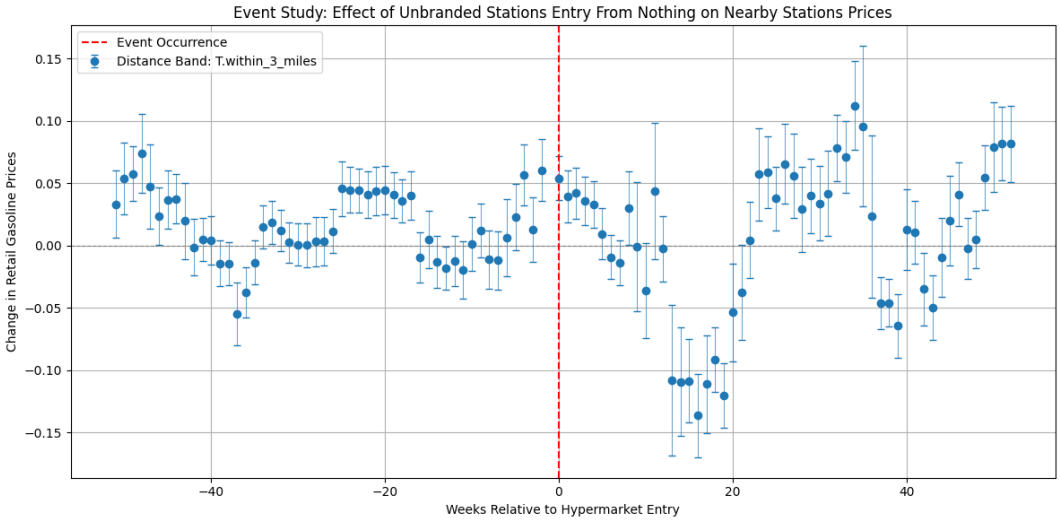



Figure 9: DiD Estimates for Unbranded Entry From Nothing

Next, we can look at branded, non-hypermarket stations changing to being unbranded; we weren't able to estimate the DiD treatment effect for the analogous event for hypermarkets, since there was only one such occurrence. However, there are 47 branded, non-hypermarket to unbranded changes in the data for the Bay Area. It appears that the treatment effects are once again very noisy, but not substantially different from the interaction term estimates from before the event occurring. This leads me to conclude there are no clear effects of branded, non-hypermarket to unbranded changes on nearby stations' prices. Interestingly, the behavior of the treatment effect after the event shadows the behavior in the case of unbranded stations entering from nothing; there's some negative effect immediately before the 20-week mark, before the treatment effect spikes and comes down to near-zero again around week 40.

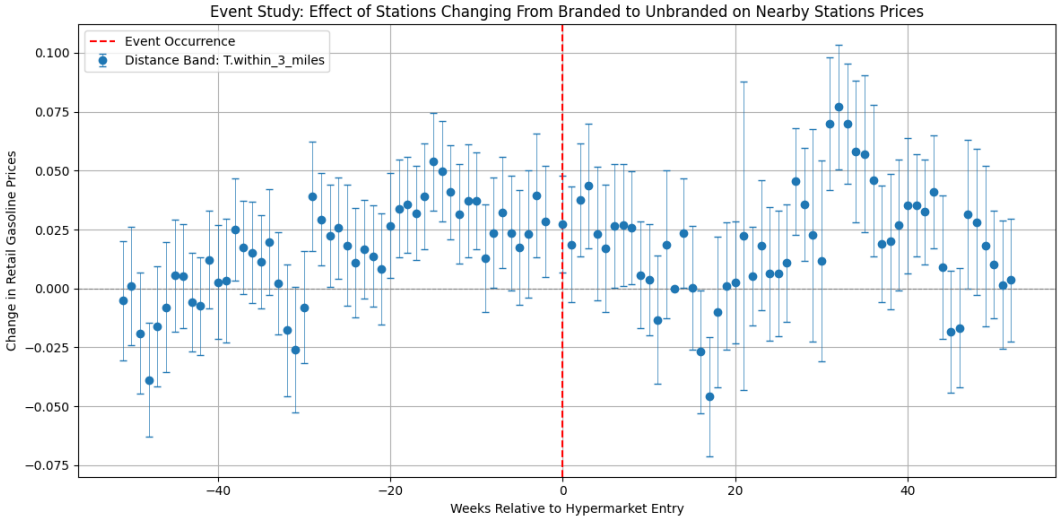



Figure 10: DiD Estimates for Branded to Unbranded Brand Change

# 6 Conclusion

We can make several conclusions from this analysis. First, prices are systematically and significantly different across different types of stations–specifically branded non-hypermarket stations, hypermarkets, and unbranded stations. We demonstrated this was not just a result of geographical heterogeneity in prices, but this significant difference holds even when comparing stations of these types to nearby stations.

Next, to understand competition between these different types of stations, we conducted an event study to understand the effect of a low-priced competitor's entry (either through entering from nothing or through a brand

change) on nearby stations' prices. By running a Difference-in-Difference (DiD) event study design where stations within three miles of the entering station were the treatment group and stations within 3-10 miles were the control group, we found there was no clear evidence of an effect on prices of low-priced competitors entering the market in the Bay Area. This finding matches our descriptive results, in which we collapsed the averages of entering stations and nearby (within three miles), somewhat nearby (within 3-10 miles), and far-away stations (within 10-20 miles) for the year before and year after event occurrence.

We hypothesize these findings are because of market segmentation; the price difference across station types is likely because consumers *do* have strong preferences, associated with a positive willingness-to-pay (WTP) for branded, non-hypermarket gasoline. Thus, even when a low-priced competitor enters a market, branded stations understand that consumers' cross-station type elasticities may be rather weak, so nearby branded stations feel little to no pressure to lower prices to compete with the entering new station. We want to further investigate consumers' price elasticities and cross-station type elasticities. We also hope to expand this analysis to the entire state of California, to see if there is heterogeneity in these findings across different geographic areas. These results may change as we apply this methodology to other areas of the state, particularly as we investigate more rural areas.

# 7 Appendix

## 7.1 Average Prices by Type - Table Representation

Table 7: Bay Area Annual Average Prices by Station Type

| Year | branded | unbranded | hypermarket |
|---|---|---|---|
| 2013 | 3.89 | 3.83 | 3.73 |
| 2014 | 3.78 | 3.72 | 3.58 |
| 2015 | 3.14 | 3.06 | 2.86 |
| 2016 | 2.75 | 2.62 | 2.43 |
| 2017 | 3.07 | 2.96 | 2.76 |
| 2018 | 3.62 | 3.50 | 3.31 |
| 2019 | 3.78 | 3.68 | 3.48 |
| 2020 | 3.29 | 3.13 | 2.91 |
| 2021 | 4.30 | 4.09 | 3.93 |
| 2022 | 5.59 | 5.40 | 5.12 |
| 2023 | 5.02 | 4.79 | 4.58 |
| 2024 | 4.99 | 4.75 | 4.52 |