# Mergers, store locations, and jobs: Evidence from the food retail industry

Debashrita Mohapatra* Rigoberto Lopez† Sandro Steinbach‡

**Abstract**

This article investigates the effects of a merger between food retailers on grocery store locations, employment, and consumer welfare. One major concern with mergers is the potential increase in market power for the merging firms, which could lead to higher prices. However, merging firms and competitors might also reorganize their store locations and workforce to save costs and avoid cannibalization. To assess the welfare implications of such mergers, we employ a structural model of California's food retail sector, where retailers decide employment and store locations based on market dynamics. By analyzing a hypothetical merger between Albertsons and Kroger, our findings shed light on how store locations and employment would have changed if the merger had been approved. Preliminary results indicate the significance of grocery store locations and workforce size in the food retail sector, highlighting a potential reduction in the store network and substantial job cuts resulting from the merger.

*University of Connecticut, debashrita.mohapatra@uconn.edu
†University of Connecticut, rigoberto.lopez@uconn.edu
‡North Dakota State University, email: sandro.steinbach@ndsu.edu

## 1 Introduction

Evaluating the implications of mergers on consumer welfare is a crucial area of focus for antitrust authorities, as mergers can raise concerns about increased market concentration and potential price increases that may negatively impact consumers. While previous research extensively examines the impact of price effects in assessing mergers, there is comparatively less exploration of how changing market structures due to mergers affect firms' investment incentives. Our study contributes to this area by investigating the influence of mergers on store locations and employment levels, which in turn impact consumer surplus and overall welfare. Specifically, we analyze the proposed merger between Albertsons and Kroger in the US food retail industry. Our analysis illuminates how such mergers can prompt firms to adjust their store locations and workforce, ultimately affecting consumer welfare and societal well-being as a whole.

We focus on the $24.6 billion proposed merger between Albertsons and Kroger in the United States. This merger raised questions about what might happen to the closing of stores, in turn leading to job cuts. News article headlines like "Kroger-Albertsons grocery chain merger in the US threatens job cuts as Wall Street eyes paydays" and "Rural towns fear Kroger-Albertsons merger will close stores and raise prices" echo the concerns expressed by various communities and industry observers. Beyond the immediate economic impacts, these news articles furthermore discuss the broader implications for consumers, discussing potential changes in access to quality food and a reduction in food product options within the consumer choice set post-merger. These concerns raise significant questions about antitrust implications, particularly regarding potential consumer welfare loss due to the merger. This paper undertakes a comprehensive evaluation of the effects of the Albertsons-Kroger merger, utilizing a panel dataset from the state of California. Our analysis is two-fold, first assessing the impact on store locations and subsequently exploring changes in the number of employees within the merged entity. Through this research, we aim to provide valuable insights into the consequences of such mergers, contributing to the ongoing discourse in the context of the food retail industry and its broader socioeconomic implications. This paper contributes to the literature on mergers (Fan (2013), Draganska, Mazzeo and Seim (2009), Caradonna, Miller and Sheu (2021), Wollmann (2018), Fan and Yang (2020)) as well as entry/exit in various locational markets (Bresnahan and Reiss (1991), Mazzeo (2002), Seim (2006)).

We employ the National Establishments Times Series (NETS) dataset, which offers annual data on fundamental attributes of food retailers in the United States. Our analysis concentrates on California in 2021, utilizing this dataset to investigate diverse attributes of food retail establishments, such as yearly revenue, location specifics (address, longitude, latitude), and store format. In addition, we integrate this dataset with consumer demographics data sourced from census tract information. This supplementary dataset provides valuable insights into demographic variables such as total population, income levels, household size, gender distribution, age demographics, and the geographical coordinates of a census tract's centroid. We integrate this data with the wage data from the Quarterly Census of Employment and Wages (QCEW) for this analysis.

We operate under the assumption that firms' decisions regarding entry are driven by their revenue-generating motives. When two food retailers merge, they often restructure their store location networks by either expanding or closing stores. If, post-merger, they opt to invest in expanding their networks and hiring more employees, thereby improving consumers' access to stores and staff, it would likely boost consumer welfare. Conversely, if they choose to close stores or cut jobs to reduce costs, it may result in decreased access to stores and staff, thereby diminishing consumer welfare. The ultimate outcome hinges on various factors such as market structure, revenue considerations, and the costs associated with expanding location networks and hiring employees, necessitating empirical investigation. The decision-making process for such restructuring depends on factors like market structure, expected revenue gains, employee hiring costs, fixed entry costs, and exit strategies. By employing a structural model, we uncover these underlying factors and assess consumer welfare by considering both employee restructuring and changes in store locations. This study contributes to existing literature focusing on the impact of mergers on equilibrium prices and welfare effects (Baker and Bresnahan (1985), Werden and Froeb (1994), Banal-Estañol, Macho-Stadler and Seldeslachts (2008), Nevo (2000), Hosken and Ashenfelter (2008), Houde (2012), Savelkoul (2016)). Furthermore, it aligns with recent research examining firms' decisions regarding non-price attributes, such as product portfolio choices, in response to changes in market structure (Draganska, Mazzeo and Seim (2009), Eizenberg (2014), Wollmann (2018), Fan and Yang (2021), Elena Argentesi, Paolo Buccirossi, Roberto Cervone, Tomaso Duso and Alessia Marrazzo (2021), Jorge Alé-Chilet, Cuicui Chen, Jing Li and Mathias Reynaert (2021), Reynaert (2021)).

We utilize a two-stage model that involves a complete information game. In the initial stage, firms decide on their store locations based on shocks affecting the sunk costs associated with market entry and exit. Following this decision, they commit to their chosen locations. In the subsequent stage, firms determine the workforce size in response to shocks in consumer demand. Once the decisions regarding store locations and workforce size are finalized, consumers visit their preferred store to make their grocery purchases.

We employ a consumer expenditure model to gain insights into consumer preferences for various store attributes. This involves connecting store-level revenue data in California with consumer preferences and inferred choices, utilizing a demographic dataset from the 2020 US census at the census tract level in conjunction with the consumer expenditure model. We establish a link between consumer demographics, store attributes, and store revenues by aggregating the inferred choice of individual consumers. We calculate the revenue for each store based on the model parameters and the observed data.

In the following steps, we calculate the range of fixed costs associated with the establishment of a grocery store in a specific location. The widespread presence of grocery stores operated by various retail brands throughout California serves as the primary indicator of entry barriers into local markets. This diversity allows us to estimate the range of fixed costs. To determine the total fixed cost, we adopt a revealed preference approach commonly utilized in empirical entry studies, particularly relying on the fact that the observed store network maximizes revenue in a Nash equilibrium. A food retailer opts to open a grocery store only when the expected revenue exceeds the fixed cost, while the closure of a store occurs if the fixed cost surpasses the expected revenue. Based on these principles, we derive the upper bound of fixed costs for any retail brand store in a local market (such as a census tract in California) if a store is operational within that tract. Conversely, for a brand store absent in the tract, we establish a lower bound. This analysis aligns with existing literature on entry-exit dynamics and fixed cost estimation, employing a moment inequality approach as discussed in references like Pakes et al. (2015), Houde, Newberry and Seim (2023), Berry, Eizenberg and Waldfogel (2016), Wollmann (2018), Mohapatra and Chatterjee (2020).

Using the estimated parameters derived from consumer expenditure, revenue, employment, and fixed cost bounds, we engage in counterfactual simulations to examine the welfare implications of mergers when food retailers reorganize their staff and location networks. This analytical exercise is designed to capture the non-price effects of mergers, emphasizing the post-merger restructuring of employment and location networks. Consequently, the restructuring process is influenced solely by the economic dynamics resulting from the merger, including factors like cannibalization, business stealing, and cost-saving effects. We specifically conduct a counterfactual analysis to assess how the merger between Albertsons and Kroger influences the retailer's decisions regarding the number of employees and the opening or closing of stores post-merger.

## 2 Dataset

This research relies on the National Establishments Times Series (NETS), a dataset that provides yearly data on essential characteristics of food retailers in the U.S. Our focus is on California in 2021, utilizing this dataset to examine various characteristics of food retail stores, including annual revenue, location details (address, longitude, latitude), and store format. We combine this dataset with consumer demographics information sourced from census tract data. This additional dataset provides insights into demographic variables such as total population, income, household size, male/female population, age, and the longitude/latitude of the centroid of a census tract. We integrate this data with the wage data from Quarterly Census of Employment and Wages (QCEW) for this analysis.

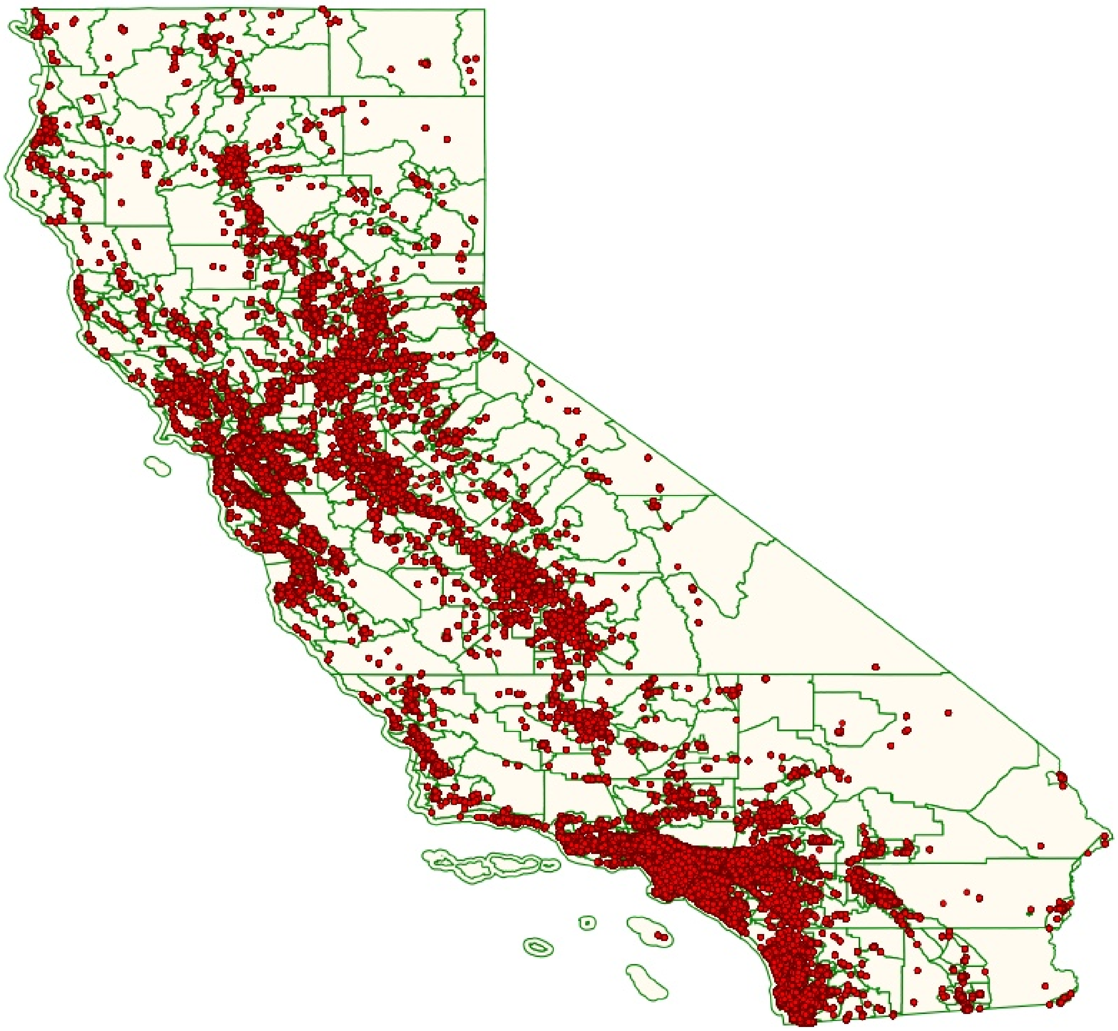
The food retailers can be distinguished in the following formats: Convenience stores, Dollar Stores, Local chain groceries, National chain groceries, Regional chain groceries, Specialty food stores, Supercenter, and Warehouse clubs. Convenience stores are small retail businesses that typically operate extended hours, and offer a limited selection of everyday items such as snacks, beverages, and basic groceries. They are designed to provide convenient access to essential goods for customers in immediate need (e.g. 7-Eleven, Circle K). Dollar stores are retail stores that sell a variety of inexpensive merchandise, often priced at one dollar or less (e.g. Dollar Tree, Dollar General, Family Dollar). They typically offer a wide range of products including household items, party supplies, toys, and sometimes groceries, catering to budget-conscious consumers. Independent grocery stores are retail outlets that operate autonomously, often locally owned and operated by individual entrepreneurs or small business owners. IGRs typically offer a range of grocery items, produce, meats, dairy products, and household goods tailored to the needs and preferences of their local community. Local chain grocery stores are retail outlets that belong to a small chain of supermarkets or grocery stores operating within a specific geographic area (e.g. Wegmans). While they may have multiple locations, they are usually regionally or locally owned and operated, distinguishing them from larger national or international chains. National chain grocery stores are part of large-scale retail chains with multiple locations across a country or even internationally (e.g. Kroger, Safeway, Albertsons). These stores offer a wide selection of grocery items, household goods, and often include additional services such as pharmacies and deli counters. Regional chain grocery stores are similar to national chain grocery stores but operate within a specific region or geographic area, rather than nationwide (e.g. Publix in the Southeastern US, Giant Eagle in the Mid-Atlantic region, Meijer in the Midwest). They may offer a selection of products tailored to local preferences while still benefiting from the economies of scale associated with chain operations. Specialty food stores focus on offering a specific category of food or beverage products, often of higher quality or uniqueness compared to those found in traditional grocery stores. Examples include Whole Foods Market specializing in organic and natural foods, Trader Joe's specializing in unique and gourmet products, Sprouts Farmers Market specializing in fresh and organic produce. Supercenters are large retail stores that combine a traditional grocery store with a wide range of general merchandise offerings, including clothing, electronics, home goods, and more (e.g. Walmart, Target). They operate as one-stop shopping destinations, often with expansive parking lots and extended operating hours. Warehouse clubs are membership-based retail stores that sell a wide variety of products in bulk quantities at discounted prices. These stores often require customers to purchase a membership to access their offerings and typically cater to both individual consumers and small businesses (e.g. Costco, Sam's Club, BJ's Wholesale Club).

A summary of the dataset is presented in Table 1. In 2021, a total of 41,137 stores were in operation. These includes 21,070 specialty food stores, 9,761 independent grocery stores, 4,440 convenience stores, 1,875 dollar stores, and 1,676 national chain groceries. Revenue varies widely across different store formats, with warehouse clubs, supercenters, and national and regional chain groceries generally generating higher revenue per year, as indicated in the table. In our dataset, 7,085 census tracts have at least one grocery store within their locality. The average population of these tracts is 4,703, with an average age of 38 and an average household size of 2.9. On average, consumers need to travel 1.7 miles to reach a grocery store. Figure 1 shows the grocery store network across California.

Table 1: Summary statistics

| Stores | # | Avg Revenue (Million $) | Std dev |
|---|---|---|---|
| Number of stores | 41,137 | 2.24 | 15 |
| Convenience | 4,440 | 1.63 | 34.13 |
| Dollar Store | 1,875 | 1.53 | 2.42 |
| IGR | 9,761 | .62 | 3.77 |
| Local Chain Grocery | 878 | 6.14 | 9.44 |
| National Chain Grocery | 1,676 | 17.09 | 14.50 |
| Regional Chain Grocery | 905 | 10.20 | 11.91 |
| Specialty Food Store | 21,070 | .35 | 3.17 |
| Supercenter | 343 | 29.77 | 19.15 |
| Warehouse club | 189 | 80.82 | 75.53 |
| **Census tract** | **Mean** | **Std. Dev.** | |
| Per cap income ($) | 40516 | 23781 | |
| Population | 4,703 | 1749 | |
| Age | 38.4 | 6.7 | |
| HH size | 2.9 | .73 | |
| Distance (in miles) | 1.7 | 3.3 | |
| # census tracts | 7,085 | | |

Figure 1: Grocery store locations in California State



# 3 Model and Estimation

We employ a two-stage model featuring a complete information game. In the first stage, firms determine the location of their stores by encountering shocks to the sunk cost of entering and exiting the market. Once this decision is made, they commit to that location. Moving on to the second stage, firms determine the number of employees by responding to shocks in consumer demand. Once the store location, and the number of employees are set, consumers visit their preferred store to purchase groceries. In our analysis, we do not observe prices of the products, we focus on the employment decision of the firms in the first stage.

## 3.1 Consumer Expenditure

We use a consumer expenditure model to understand consumer's preferences for different store characteristics (similar approach as in Ellickson, Grieco and Khvastunov (2020)). Our study uses store-level revenue data from all grocery stores in California. We link the store-level revenue information to consumer preferences and implied choices by using census tract-level demographic dataset from 2020 US census, along with model of consumer expenditure. The census tract data is spatially disaggregated and we use the variation of consumer demographics across tracts for our identification. We assume a representative household in a census tract poses the average household characteristics, and traven a distance between tract centroid and a store while buying groceries. A consumer spends a fixed proportion $\alpha$ of his or her income on weekly grocery budget and on the outside good, where $\alpha$ is a parameter to be estimated. Individuals distribute their budget based on a discrete-choice random utility model across a selection of nearby stores, each possessing a location and a set of characteristics $x_s$ (like store format, chain affiliation), alongside an external option. Our model accommodates variations in individuals' preferences for stores, as well as for the outside option, which can vary according to income levels. Hence, we can observe the empirical evidences where the share of income allocated to grocery store expenses (e.g., expenditures on groceries instead of outside goods) decreases as income levels rise. Each consumer engages in a continuous series of purchase decisions to distribute their budget among stores. For each unit of expenditure $i$, a consumer in tract $t$ assesses their utility for spending at store $s$ as follows:

$$
\begin{aligned}
u_{ist} &= u_{st} + \epsilon_{ist} \\
&= \tau_0 d_{st} + \tau_2 d_{st} z_t + \gamma_0 x_s + \gamma_1 (x_s \otimes z_t) + \epsilon_{ist}
\end{aligned} \tag{3.1}
$$

where, $d_{st}$ refers to distance of store $s$ from consumer $i$ in tract $t$, $z_t$ includes tract-level consumer demographics, and $x_s$ includes store characteristics. Consumer's baseline utility from a store is denoted as $u_{st}$. Each purchase decision is subject to an idiosyncratic preference shock, $\epsilon_{ist}$, that follows a Generalized Extreme Value (GEV) distribution. This framework enables the utility of a particular store to be determined by its proximity to consumers, along with store characteristics encompassing factors such as product availability, service quality, and convenience. Individuals are allowed to differ in tastes for distance and other characteristics though heterogeneity in the consumer's (tract-level) demographic variables $z_t$. This allows the utility of different store characteristics (including format type) to vary across observable consumer characteristics such as income.

We also incorporate chain affiliation into $x_s$, capturing unobserved characteristics of specific national chains like product variety and pricing strategies. Since we don't have information on actual prices or product offerings, the store's chain affiliation reflects the overall pricing, quality, and assortment approach of the company, presumed to be determined at the chain level. For instance, some grocery stores may position themselves as budget-friendly, offering limited choices to cater to price-sensitive shoppers, while others may target rich customers who value high-end organic products. Alternatively, conventional supermarket chains often aim for a broader market segment, resulting in less differentiation and increased competition. These diverse strategies appeal to consumers in varying ways. For example, Ellickson and Misra (2008) note that supermarkets adopt different pricing and positioning tactics to target distinct consumer groups based on purchase behavior. Although our model doesn't directly include individual store pricing or quality decisions, we believe these are secondary to the overall chain policies. Considering chain effects also acknowledges that purchasing one dollar's worth of goods at Whole Foods, a high-end grocer, results in a different product assortment compared to spending the same amount at Walmart, which targets lower-income urban populations. Additionally, by combining these chain identifiers with consumer characteristics (such as income), we allow for variations in the utility trade-off between shopping at different chains, like Walmart and Whole Foods, across consumers.

A consumer's utility from the outside good is defined as the representative consumer's (tract-level) demographic characteristics and a set of physical tract characteristics $w_t$, such as population density, that control for the availability of alternative consumption options in the tract's vicinity,

$$u_{0ti} = \lambda_0 w_t + \lambda_1 w_t z_t + \epsilon_{0it}$$

We assume that the household's choice set consists of all stores located within D miles of their resident tract, as well as the outside option, i.e $C_t = \{s : d_{st} \leq D\} \cup 0$. We allow our model to capture the fact that similar formats offer more uniform retail experiences and therefore may compete more intensely within rather than across formats, even after controlling for store characteristics. The nested Generalized Extreme Value (GEV) framework addresses this by incorporating correlation among $\epsilon_{sti}$ terms for stores sharing the same format (i.e., within the same nest). Here, $0 \leq k \leq 1$ represents the parameter governing this correlation, where $\mu_k = 1$ signifies independent shocks within nest $k$ (the case of multinomial logit), and $k = 0$ indicates perfect correlation of $\epsilon_{sti}$ within the nest.

We integrate over the GEV shocks, and derive the share of consumers in tract $t$ budget on groceries and how much they spend on store $s$. We express this as a function of parameter vector i.e. $\theta = (\tau, \gamma, \lambda, \mu)$. Let $C_{t_k}$ represent all the stores within the choice set of tract $t$ belonging to the nest $k$, and $k(s)$ denote the nest to which store $s$ pertains. Subsequently, we define $C_{t,k(s)} = \{q \in C_t : k(s) = k(q)\}$ as the set of stores in the choice set of tract $t$ that are categorized within the same nest as store $s$. Lastly, let $\iota_{ti}$ denote the store where consumer type $t$ allocates expenditure unit $i$. The share of expenditure at store $s$, relative to total spending in tract $t$, can be broken down into the cumulative expenditure on nest $k(s)$ and the outlay at store $s$ as a fraction of all expenditure within $k(s)$ and can be expressed as,

$$p_{st}(\theta) \equiv Pr(\iota_{ti} = s) = Pr(\iota_{ti} \in C_{t,k(s)}) Pr(\iota_{ti} = s | \iota_{ti} \in C_{t,k(s)})$$

Given our distributional assumption, the share of expenditure on stores in $C_{t,k(s)}$ (a grocery store close to tract $t$) is

$$Pr(\iota_{ti} \in C_{t,k(s)}) = \frac{(\sum_{q \in C_{t,k(s)}} e^{u_{qt}/\mu_{k(s)}})^{\mu_{k(s)}}}{\sum_{v=0}^{K} (\sum_{q \in C_{t,v}} e^{u_{qt}/\mu_v})^{\mu_v}}$$

where $u_{st}$ is the baseline utility that consumers in tract t obtain from visiting store s (a function of model parameters defined above). The probability of choosing a particular store s from the set of options included in $C_{t,k(s)}$ is,

$$Pr(\iota_{ti} = s | \iota_{ti} \in C_{t,k(s)}) = \frac{e^{u_{qt}/\mu_{k(s)}}}{\sum_{q \in C_{t,k(s)}} e^{u_{qt}/\mu_{k(s)}}}$$

The unconditional share can be written as

$$p_{st}(\theta) = \frac{e^{u_{qt}/\mu_{k(s)}} (\sum_{q \in C_{t,k(s)}} e^{u_{qt}/\mu_{k(s)}})^{\mu_{k(s)-1}}}{\sum_{v=0}^{K} (\sum_{q \in C_{t,v}} e^{u_{qt}/\mu_v})^{\mu_v}}$$

In theory, we have the option to incorporate further layers of unobserved diversity that rely on store or tract attributes, akin to permitting random coefficients within our utility framework as discussed in Berry, Levinsohn and Pakes (1995). However, by integrating the combined distribution of income and location, we account for a considerable portion of observed heterogeneity. This approach helps in looking into extensive substitution patterns across chains, while maintaining a streamlined analytical framework.

## 3.2 Store Revenues and Employment

We establish a link between consumer demographics, store attributes, and store revenues by aggregating the inferred choice of individual consumers. We calculate the revenue for each store based on the model parameters and the observed data. The revenue generated in store $s$ from expenditures in tract $t$ is essentially the total budget of all consumers in tract $t$ multiplied by the fraction of those expenditures allocated to store $s$.

$$\hat{R_{st}}(\theta, \alpha) = \alpha \, inc_t \, . \, n_t \, . \, p_{st}(\theta)$$

Here, $inc_t$ represents the per capita income in tract $t$, and $n_t$ signifies the total population residing in tract $t$. The model parameter, $\alpha$, reflects the share of income that consumers allocate to overall grocery expenses, encompassing purchases from both the store and outside alternatives. Store $s$ generates revenue from all tracts where it is included in the choice set, which typically includes tracts within a 10-mile radius of its location.

The predicted total revenue of a store from multiple tracts, within D miles can be expressed as

$$\hat{R_s}(\theta, \alpha, \kappa) = \sum_{t \in L_s} \hat{R_{st}}(\theta, \alpha) + \kappa_1 w_{st} + \kappa_2 w_{st}^2 \ - \omega_{st} w_{st}$$

where the stores generate revenue through employees and has an expenditure for employee salaries. Here $\omega_{st}$ is the wage rate of the employees in a census tract $t$, and $w_t$ is the number of employees in a given store. The estimated total revenue for store $s$ is given by:

$$\hat{R_s}(\theta, \alpha, \kappa) = \sum_{t \in L_s} R_{st}(\theta, \alpha)$$

Here, $L_s = \{t : s \in C_t\} = \{t : d_{st} \leq D\}$ represents the set of tracts where store $s$ is included in the choice set. An important aspect of this modeling approach is that it avoids imposing arbitrary geographic market boundaries. Instead, each store is positioned at the center of its own service area. Stores situated in close proximity to one another will have service areas with significant overlap. Consequently, they will exert a stronger competitive influence on each other compared to stores located farther apart, and will compete most intensely for customers in their immediate vicinity.

To estimate the model parameters, we compare the revenue prediction generated by the model with the actual revenues observed in the data, selecting the parameters that minimize the discrepancy between the two. To accommodate potential measurement errors in the revenue data, we presume that the observed revenues for each store undergo a multiplicative shock, which is independent of the exogenous variables and consistent across stores.

$$R_s = e^{\eta_s} \hat{R}_s(\theta_0, \alpha_0, \kappa_0)$$

In this context, $(\theta_0, \alpha_0, \kappa_0)$ represent the true parameters, and $\epsilon_s$ symbolizes the store-level measurement shock. Given these premises, the parameters can be inferred using nonlinear least squares estimation techniques.

$$(\hat{\theta}, \hat{\alpha}, \hat{\kappa}) = \underset{\theta, \alpha}{\text{argmin}} \sum_s (log(\hat{R}_s(\theta, \alpha)) - log(R_s))^2$$

Demonstrating the consistency and asymptotic normality of this estimator is relatively simple. The standard variance-covariance matrix is inferred directly from the nonlinear least squares objective function.

The firms endogenously choose the number of employees in a store. At equilibrium, the store choose optimal number of employees such that it satisfies the equilibrium condition of marginal revenue to marginal cost such that,

$$\frac{\delta R_{st}}{\delta \eta_{slt}} = \kappa_1 + 2\kappa_2 w_{st} - \omega_t$$

where $w_{st}$ is number of employees in store $s$, tract $t$ and $\omega_t$ is wage rate in tract $t$.

### 3.3 Store Location Decision: Stage 1

In the first stage of the model, food retailers make decisions regarding their store locations. In this context, we assume that these retailers are the primary decision-makers for store locations. The information framework is regarded as a game of complete information. Consequently, at the outset of the first stage, all food retailers become aware of the fixed costs linked to a store location. However, during the process of making price and employment decisions, firms are presumed to possess knowledge about the distribution of demand and cost shocks. They observe the realization of these shocks only after stage 2 is realized, after having committed to the store location.

Nash equilibrium implies that given the store location network chosen by the competitors at the equilibrium, any unilateral deviation from a store's locational network decision should not lead to higher expected revenue for the brand, where the expectation is taken over demand and marginal cost shocks. In particular, we consider two types of deviations: the removal of a grocery store in a town and the opening of a store not present in the town.

We consider each census tract in california as a potential location for a grocery store to enter. We first look into the locations where a grocery store of a given retailer is present. We compute the equilibrium revenue conditional on every one-step deviation, removing one store at a time. Inequality in equation 3.2 gives an upper bound of $F_{st}$ for a store-tract $st$ in the data. Intuitively, a food retailer chooses to open a grocery store in a location if the corresponding gain in the revenue dominates the fixed cost for the grocery store. The expression is given by

$$E_\xi R_s(\mathbf{w}, \mathscr{L}_{st}) - F_{st} \geq E_\xi R_s(\mathbf{w}, \mathscr{L}_{st} \setminus s) \tag{3.2}$$

where, $R_s(\mathbf{w}, \mathscr{L}_{st})$ is the equilibrium revenue of a food retailer $s$, $F_{st}$ is the fixed cost of a store associated with a tract $t$, $R_s(\mathbf{w}, \mathscr{L}_{st} \setminus s)$ is the store $s's$ revenue if a store at tract $t$ is removed from its network.

Next, we consider the situation where a grocery store is added. Here, store $s's$ expected revenue should not increase if a store $s\tilde{t}$ is added to a town $\tilde{t}$. The corresponding inequality is:

$$E_\xi R_s(\mathbf{w}, \mathscr{L}_{st}) \geq E_\xi R_b(\mathbf{w}, \mathscr{L}_{st} \cup \tilde{t}) - F_{s\tilde{t}} \tag{3.3}$$

The inequality in (3.3) estimates a lower bound of $F_{s\tilde{t}}$ for any $\tilde{t}$ such that $\tilde{t} \notin \mathscr{L}_{st}$. Note that, we allow the fixed cost to open a grocery store to differ across census tracts. Hence, we allow a retailer's investment in a grocery store to affect the fixed costs but not any other costs. The computed fixed cost of the grocery stores includes the retailer's investment in the opening and maintenance of the store, as well as advertising in the local area.

We estimate the fixed cost bounds of a grocery store in a location by using the inequalities as mentioned in (3.2) and (3.3). We use inequality in equation (3.2) to calculate the upper bound of the fixed cost of opening a store in the tract where it is present. Note that removing a store changes the distance of a food retailer from different tracts. This leads to a change in a store's revenue. We calculate the upper bound as the reduced expected revenue when a store is removed. Similarly, the inequality in equation (3.3) is used to calculate the lower bound of the fixed cost of opening a grocery store in the census tract where it is not present. In this case, we add a store into a tract where it is not present in the data. Suppose grocery stores of other retailers are already present in the town. In that case, we add the store at a location that corresponds to the average over the longitude-latitude of other store' locations. We calculate the distance to a store after it is added and use the distance vector to calculate the expected revenue. The increase in revenue after adding a store into a location serves as the lower bound of the fixed cost of a retailer in a given location.

### 3.4 Identification

After presenting our model and estimation approach, we will now discuss the data variation and necessary assumptions essential for identifying the model parameters. The identification of these parameters stems from observing the geographic diversity in population demographics, store locations, and store revenues. We initially assume that $\epsilon_{its}$ and $\epsilon_{st}$ are independent of stores' decisions regarding residential location and size, as well as consumers' selected locations and observed incomes. Specifically, we assume that consumers perceive store locations as fixed and form opinions about stores' pricing, quality, and product variety at the chain level rather than the individual store level. This allows us to mitigate the endogeneity of these policies using chain fixed effects. While it's plausible that chains may adjust pricing policies store by store based on local demographics (Ellickson and Misra (2008), Hoch et al. (1995)), we consider this concern as less significant for two reasons. Firstly, supermarket firms set prices for numerous products per store, making it impractical for consumers to compute price indices for each outlet. Instead, they likely have a general perception of price differences across chains and use this as a heuristic in selecting their primary store. Secondly, grocery stores typically set prices at the level of broader "pricing zones" rather than individual stores. This strategy allows stores to collectively market their products to a wider area while minimizing costs. Consequently, we assume that it is not efficient for chains to set policies at the individual store level. Although these pricing zones are not typically nationwide, we believe that within-chain variation in pricing and product offerings across a pricing zone is less significant than across-chain variation within the same zone. This latter variation is captured in our framework via chain fixed effects.

Shifting our focus to identifying specific parameters, we first concentrate on $\alpha$, which represents the proportion of total income allocated to grocery expenditures. We identify this parameter by varying the total number of stores in otherwise identical markets and observing the resultant change in total revenue across all stores. Intuitively, adding multiple stores to a market should reduce the share of the outside good towards zero. Eventually, additional stores may not increase total revenue but only redistribute revenue across stores. In this context, $\alpha$ is simply the ratio of total revenue of all stores to the total income of the associated population of consumers. Generally, the change in total revenue in response to altering the number of stores reveals the trade-off between the outside good and the new store while maintaining regional income constant. This allows us to determine the share of the outside good and subsequently identify $\alpha$.

Once we have identified $\alpha$, parameters governing store utility are determined by varying observable characteristics of both stores and consumers and observing the resulting changes in the share of total expenditure of consumers within the catchment area, $L_s$, allocated to each store. For example, consider the impact of distance on store choice. Varying the distance between a tract and a store affects the share of expenditures at that store relative to others in the tract's choice set, which is reflected in the store's revenue relative to others in the same choice set. In conclusion, this framework allows us to compute various statistics revealing the impact of distance and demographics on each firm's revenue. These statistics aid in assessing the model's performance and provide valuable insights for merger analysis, particularly regarding grocery competition's localization, both geographically and by firm and format type. Identification of the fixed cost parameters rely on the variation of locational entry by the food retailers.

# 4 Results

The findings derived from the demand model offer a comprehensive understanding of the intricate dynamics that shape consumer behavior and preferences within the context of grocery stores. Each coefficient in the table signifies the magnitude and direction of influence that specific variables exert on the overall demand for grocery stores. A negative coefficient, such as that observed for Distance (-0.021), suggests that as the distance to a store increases, there is a corresponding decrease in consumer demand, albeit a slight one. Conversely, positive coefficients, exemplified by store types like Local chain (4.53), National chain (15.67), Regional chain (8.608), Supercenter (27.83), and Warehouse (80.07), indicate a notably higher demand for these particular store formats compared to others. These insights are crucial for retailers and policymakers alike, as they illuminate the factors driving consumer choices and enable strategic decision-making regarding store locations, assortment offerings, and marketing strategies. Moreover, the interaction terms such as Income * Distance (-0.082) and Age * Distance (-0.01) shed light on the nuanced interplay between demographic factors like income and age with the influence of distance on consumer demand, highlighting the multifaceted nature of consumer decision-making processes. Overall, these regression results

| | Coeffcient | Std err |
|---|---|---|
| Distance | -0.021 | 0.019 |
| Dollar store | -0.077 | 0.359 |
| IGR | -1.003 | 0.236 |
| Local chain | 4.53 | 0.48 |
| National chain | 15.67 | 0.372 |
| Regional chain | 8.608 | 0.475 |
| Speciality food store | -1.29 | 0.215 |
| Supercenter | 27.83 | 0.732 |
| Warehouse | 80.07 | 0.959 |
| Income | 0.085 | 0.069 |
| Age | 0.008 | 0.01 |
| Household size | 0.0018 | 0.007 |
| Income * Distance | -0.082 | 0.034 |
| Age * Distance | -0.01 | 0.02 |

not only provide valuable insights into consumer preferences but also serve as a foundation for optimizing store operations and enhancing the overall consumer experience within the grocery retail landscape.

## 5 Conclusion

Our study looks into the evaluation of mergers within the food retail sector and their consequential impacts on grocery store locations, employment levels, and overall consumer welfare. The primary focus of our investigation is to unravel the potential repercussions of a merger, particularly concerning the consolidation of market power among the merging firms. The concern over increased market power often raises apprehensions about potential price escalations, which can negatively impact consumer welfare. However, we also acknowledge the strategic reorganization that merging firms and competitors might undertake, aiming to streamline operations, save costs, and avoid cannibalization within the market.

Our analytical approach, centered on a structural model of California's food retail sector, enabled us to scrutinize the implications of a hypothetical merger scenario between Albertsons and Kroger. Through our analysis, we were able to illuminate how such a merger could have influenced store locations and employment patterns had it been approved. The preliminary results gleaned from our investigation underscored the critical role played by grocery store locations and workforce size in shaping the dynamics of the food retail sector. Notably, our findings hinted at a potential reduction in the overall store network and significant job cuts resulting from the merger, indicating profound implications for the industry's landscape and workforce dynamics.

Overall, our study contributes valuable insights into the multifaceted impacts of mergers in the food retail sector, emphasizing the intricate balance between market consolidation, cost-saving strategies, and consumer welfare considerations. The implications drawn from our research underscore the importance of conducting thorough evaluations of mergers, particularly in terms of their non-price effects, such as changes in store locations and employment levels. Such evaluations are paramount for antitrust authorities to make informed decisions, ensuring competition remains robust, consumer interests are protected, and the food retail market thrives in a healthy and competitive environment.

# References

**Alé-Chilet, Jorge, Cuicui Chen, Jing Li, and Mathias Reynaert.** 2021. "Colluding against environmental regulation."

**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.

**Baker, Jonathan B, and Timothy F Bresnahan.** 1985. "The gains from merger or collusion in product-differentiated industries." *The Journal of Industrial Economics*, 427–444.

**Banal-Estañol, Albert, Inés Macho-Stadler, and Jo Seldeslachts.** 2008. "Endogenous mergers and endogenous efficiency gains: The efficiency defence revisited." *International Journal of Industrial Organization*, 26(1): 69–91.

**Berry, Steven, Alon Eizenberg, and Joel Waldfogel.** 2016. "Optimal product variety in radio markets." *The RAND Journal of Economics*, 47(3): 463–497.

**Berry, Steven, James Levinsohn, and Ariel Pakes.** 1995. "Automobile prices in market equilibrium." *Econometrica: Journal of the Econometric Society*, 841–890.

**Bresnahan, Timothy F, and Peter C Reiss.** 1991. "Entry and competition in concentrated markets." *Journal of political economy*, 99(5): 977–1009.

**Caradonna, Peter, Nathan Miller, and Gloria Sheu.** 2021. "Mergers, entry, and consumer welfare." *Georgetown McDonough School of Business Research Paper*, , (3537135).

**Draganska, Michaela, Michael Mazzeo, and Katja Seim.** 2009. "Beyond plain vanilla: Modeling joint product assortment and pricing decisions." *QME*, 7(2): 105–146.

**Eizenberg, Alon.** 2014. "Upstream innovation and product variety in the us home pc market." *Review of Economic Studies*, 81(3): 1003–1045.

**Ellickson, Paul B, and Sanjog Misra.** 2008. "Supermarket pricing strategies." *Marketing science*, 27(5): 811–828.

**Ellickson, Paul B, Paul LE Grieco, and Oleksii Khvastunov.** 2020. "Measuring competition in spatial retail." *The RAND Journal of Economics*, 51(1): 189–232.

**Fan, Ying.** 2013. "Ownership consolidation and product characteristics: A study of the US daily newspaper market." *American Economic Review*, 103(5): 1598–1628.

**Fan, Ying, and Chenyu Yang.** 2020. "Competition, Product Proliferation, and Welfare: A Study of the US Smartphone Market." *American Economic Journal: Microeconomics*, 12(2): 99–134.

**Fan, Ying, and Chenyu Yang.** 2021. "Merger, Product Variety and Firm Entry: the Retail Craft Beer Market in California."

**Hoch, Stephen J, Byung-Do Kim, Alan L Montgomery, and Peter E Rossi.** 1995. "Determinants of store-level price elasticity." *Journal of marketing Research*, 32(1): 17–29.

**Hosken, Daniel, and Orley Ashenfelter.** 2008. *The effect of mergers on consumer prices: evidence from five selected case studies.* Citeseer.

**Houde, Jean-François.** 2012. "Spatial differentiation and vertical mergers in retail markets for gasoline." *American Economic Review*, 102(5): 2147–82.

**Houde, Jean-François, Peter Newberry, and Katja Seim.** 2023. "Nexus Tax Laws and Economies of Density in E-Commerce: A Study of Amazon's Fulfillment Center Network." *Econometrica*, 91(1): 147–190.

**Mazzeo, Michael J.** 2002. "Product choice and oligopoly market structure." *RAND Journal of Economics*, 221–242.

**Mohapatra, D, and Chirantan Chatterjee.** 2020. "Price Control and Access to Drugs: The Case of India's Malaria Market." Working Paper.

**Nevo, Aviv.** 2000. "Mergers with differentiated products: The case of the ready-to-eat cereal industry." *The RAND Journal of Economics*, 395–421.

**Pakes, Ariel, Jack Porter, Kate Ho, and Joy Ishii.** 2015. "Moment inequalities and their application." *Econometrica*, 83(1): 315–334.

**Reynaert, Mathias.** 2021. "Abatement strategies and the cost of environmental regulation: Emission standards on the European car market." *The Review of Economic Studies*, 88(1): 454–488.

**Savelkoul, Ruben.** 2016. "Ex Post Merger Analysis in the Car Market and the Role of Conduct."

**Seim, Katja.** 2006. "An empirical model of firm entry with endogenous product-type choices." *The RAND Journal of Economics*, 37(3): 619–640.

**Werden, Gregory J, and Luke M Froeb.** 1994. "The effects of mergers in differentiated products industries: Logit demand and merger policy." *JL Econ. & Org.*, 10: 407.

**Wollmann, Thomas G.** 2018. "Trucks without bailouts: Equilibrium product characteristics for commercial vehicles." *American Economic Review*, 108(6): 1364–1406.