Demand Estimation of Oatmeal

Edward Gilligan

January 2026

Abstract

This paper estimates consumer demand for oatmeal using Dominick's supermarket scanner data and applies a structural discrete-choice framework to evaluate market behavior and counterfactual strategies. Leveraging a Berry (1994) logit model using rich product and store–week fixed effects, the analysis recovers price elasticities, examines substitution patterns, and quantifies revenue and consumer-welfare implications around a range of hypothetical interventions. My results across oatmeal products, notwithstanding other outcomes, indicate that price sensitivity exhibits particularly notable elasticities with instant varieties. My counterfactual simulations reveal that analogous, modest price increases generate predictable revenue gains with small welfare losses, while targeted or aggressive discounts—especially for high-elasticity high-share UPCs—produce large consumer-welfare improvements at the cost of significant revenue declines. These findings highlight the central tradeoff between profitability and consumer benefit in pricing decisions and underscore the importance of accounting for product heterogeneity and substitution dynamics when designing category-level pricing strategies.

1 Introduction

This note studies consumer demand for oatmeal and evaluates a set of market counterfactuals to estimate price elasticities, assess consumer welfare, and characterize how consumers respond to changes in prices and product attributes. Oatmeal provides a useful setting for this analysis: it is a differentiated packaged good with substantial variation across brands and formats, and demand likely shifts seasonally and across consumer types. In such markets, simple correlations between prices and quantities can be misleading due to endogenous pricing, unobserved quality differences, and promotional activity. To address these concerns, I adopt a structural demand framework that recovers underlying preference parameters and enables welfare analysis under simulated pricing, packaging, and promotional scenarios.

The analysis uses scanner data from Dominick's Finer Foods provided by the University of Chicago Booth School of Business Kilts Center.[1] The data cover more than seven years of weekly observations from approximately 100 stores in the Chicago metropolitan area and include over 3,500 Universal Product Codes (UPCs), each corresponding to a distinct brand–size–variety combination. For each UPC, the dataset records quantities sold, prices, profits, and product characteristics. Store-level demographic variables derived from the 1990 U.S. Census are also available, along with weekly customer counts, which I use to define market size at the store–week level. After merging these components and applying sample restrictions, the resulting panel contains 752,069 store–week–UPC observations.

I construct market shares as the quantity sold of each UPC divided by the total number of customers visiting a given store in a given week, then estimate a logit demand model in which the dependent variable is the log of a product's market share relative to an outside option—defined as non-purchase of oatmeal in that store–week. I consider two specifications: a parsimonious baseline model and a richer specification incorporating extensive product and store–week fixed effects. Standard errors are clustered at the store level to account for within-store correlation over time. Price endogeneity is mitigated through two channels: the use of revenue-weighted prices reduces measurement error from low-volume transactions, while store–week fixed effects absorb common demand shocks, allowing for more credible estimation of price sensitivities and welfare measures.

The results reveal a strong and statistically significant negative relationship between price and market share, indicating substantial consumer responsiveness to price changes. This relationship remains stable across specifications that control flexibly for product-level heterogeneity and local market conditions, suggesting that the estimated price sensitivity reflects underlying consumer preferences rather than unobserved product quality or store-specific demand shocks. The implied own-price elasticity for an average product is approximately 1.75, indicating that a one percent increase in price reduces quantity demanded by roughly 1.75 percent. Notably, instant oatmeal exhibits considerably higher price sensitivity than regular oatmeal, a pattern consistent with stronger competition from ready-to-eat breakfast alternatives in the convenience segment. These findings motivate the counterfactual analysis that follows, which uses the estimated demand system to quantify revenue and welfare implications of alternative pricing strategies.

[1] I would like to thank the Kilt's Center at the University of Chicago Booth School of Business for granting public access to this dataset.

This paper makes three contributions to the empirical analysis of differentiated product markets. First, it provides detailed evidence on price sensitivity and substitution patterns within the oatmeal category using high-frequency scanner data, documenting systematic differences in demand elasticities between instant and regular varieties that have implications for category management. Second, it demonstrates that economically plausible demand elasticities can be recovered using a Berry (1994) logit framework with rich fixed effects, even without explicit cost instruments or complete promotional data. The controls-as-instruments strategy proves effective in this setting, offering a practical approach for researchers working with retail scanner data. Third, it quantifies the revenue and consumer-welfare consequences of five alternative pricing and packaging strategies through counterfactual simulations, highlighting the tradeoffs between profitability and consumer surplus when elasticities vary across product types. Taken together, these contributions shed light on pricing behavior in packaged food markets and provide guidance for retailers considering uniform price changes, targeted discounts, or package-size adjustments.

The remainder of the paper is organized as follows. Section 2 describes my methodology, including the data, its limitations, the construction of market shares, and the demand estimation framework. Section 3 presents the main results, including price elasticities, and evaluates counterfactual pricing and packaging scenarios. Section 4 concludes with a summary of the findings and directions for future research.

2 Methodology

2.1 Data

The data are drawn from Dominick's Finer Foods scanner data and span more than seven years of weekly observations from the late 1980s through the mid-1990s (dom, 2024). The unit of observation is a store–week–UPC, where a UPC corresponds to a distinct brand–size–variety combination. Dominick's operated roughly 100 stores in the Chicago metropolitan area, making the scanner data geographically concentrated in this region. The analysis focuses on oatmeal products as defined by Dominick's category classifications.

I restrict the estimation sample to economically relevant transactions and products through several filters. Zero-quantity observations are dropped because the Berry inversion requires strictly positive shares; this restriction is standard in scanner-data applications of logit demand. I exclude extreme prices by trimming the 1st and 99th percentiles of the revenue-weighted price distribution. Revenue-weighted prices mitigate the influence of measurement error from low-volume transactions and ensure that the effective price reflects the price faced by the marginal unit sold rather than idiosyncratic small purchases. To reduce sparsity and ensure reliable identification of substitution patterns, I retain only products sold in at least half of store–weeks. After applying these restrictions, the final sample consists of 752,069 store–week–UPC observations.

Observed product characteristics include package size (measured in ounces), an indicator for whether the product is labeled as instant, and a brand identifier, among others. Figure 1 displays the distribution of prices across all observations, which range from \$1.04 to \$4.19 with a mean of \$2.76. This variation across brands, package sizes, and over time provides the identifying variation used in demand estimation. The substantial price dispersion—spanning more than a four-to-one range—reflects both differences in product attributes and temporal fluctuations driven by promotional cycles and competitive adjustments.

Figure 1: Histogram of Prices

figures/hist_price.png

In addition to product-level scanner data, the dataset includes store-level demographic information constructed from the 1990 U.S. Census for the Chicago metropolitan area, capturing income, age composition, education, household size, and ethnicity. Because these characteristics are time-invariant, they are absorbed by store fixed effects in the baseline specification and do not enter directly. In alternative specifications, I interact these demographics with time-varying covariates to explore heterogeneity in price sensitivity across different consumer populations, though these results are not the focus of the present analysis.

Market size is measured using Dominick's store–week customer counts, representing the total number of shoppers visiting each store in a given week. Product-level market shares are computed as units sold of each UPC divided by the corresponding store–week customer count. This approach treats all store visitors as potential oatmeal buyers, implicitly defining the outside option as choosing not to purchase oatmeal during that shopping trip. Observations missing prices, quantities, or market shares are excluded. The dependent variable in the logit demand estimation is the log of each product's market share relative to this outside option.

2.2 Limitations

Several limitations qualify the interpretation of the demand estimates. First, the logit specification imposes the independence of irrelevant alternatives (IIA) assumption, which implies proportional substitution patterns across products. Under this restriction, a price increase for one product reallocates demand toward other products in proportion to their baseline market shares, regardless of product similarity. Substitution toward close competitors—such as between instant and regular oatmeal within the same brand—may therefore be understated, while substitution toward dissimilar products may be overstated. More flexible models, such as nested logit or random-coefficients specifications, would relax this assumption by allowing richer correlation structures across products. However, nested logit requires credible nesting definitions that may be arbitrary in a category with multiple overlapping dimensions of similarity, while random-coefficients models demand either cost-side instruments or sufficient price variation within narrowly defined product groups—neither of which is readily available in the Dominick's data.

Second, while the controls-as-instruments strategy substantially mitigates price endogeneity, residual correlation between prices and unobserved demand shocks may persist. Identification relies on high-dimensional product and store–week fixed effects to absorb common sources of price–demand correlation in the absence of explicit supply-side cost shifters such as input prices or wholesale contracts. Remaining variation comes from within-product price changes over time, which may reflect short-run promotional adjustments, inventory management, or responses to idiosyncratic sales fluctuations. Promotional activity is measured imperfectly due to incomplete coding in the scanner data, raising the possibility that the estimated price coefficient partially captures endogenous promotional behavior rather than exogenous cost-driven variation. This concern is mitigated by the use of revenue-weighted prices and the absorption of store–week-level promotional intensity through fixed effects, but it cannot be fully eliminated without access to retailer cost data.

Third, the model is static and abstracts from intertemporal considerations. Consumers may stockpile oatmeal during promotions, shift purchases across weeks in response to temporary price changes, or develop brand preferences that evolve through repeated purchases. These dynamics—including habit formation and state dependence—are not captured in a single-period logit framework, which treats each shopping occasion as an independent choice problem. Dynamic demand models would address these patterns but require panel data tracking individual consumers over time, which are unavailable in the aggregated scanner data. In addition, the analysis examines a single product category within a single geographic market during a specific historical period. Dominick's operated exclusively in the Chicago metropolitan area in the late 1980s and early 1990s, and demand elasticities may differ in other regions with different demographics, competitive conditions, or consumption patterns. Contemporary retail environments also feature online shopping, larger assortments, and different promotional strategies, all of which could alter the observed substitution patterns.

Finally, market size is proxied by store–week customer counts, which treat all store visitors as potential oatmeal buyers regardless of whether they considered purchasing the category. This approach potentially overstates the size of the outside option and understates inside good market shares, biasing elasticity estimates toward zero. An alternative approach would define potential buyers more narrowly—for instance, as households purchasing breakfast foods—but such refinements are infeasible without individual-level basket data. Despite these limitations, the qualitative patterns of substitution and the relative elasticities across product types remain informative for understanding demand behavior within the oatmeal category. The estimated elasticities are economically plausible and consistent with findings from similar packaged goods categories, and the counterfactual exercises that follow provide directionally useful guidance even if the precise magnitudes should be interpreted with appropriate caution.

2.3 Estimation Method and Model

To estimate consumer demand for oatmeal products, I adopt the discrete-choice framework of Berry (1994). Consumers choose among a finite set of differentiated products, including an outside option, and observed market shares arise from the aggregation of individual utility-maximizing choices. The key object of interest is the product-level mean utility, which captures the average valuation of each product net of price and observed characteristics.

Market shares s_{jst} are defined as the units sold of product j in store s during week t divided by the store-week customer count M_{st}. The outside option share is

$$s_{0st} = 1 - \sum_{j \in \mathcal{J}_{st}} s_{jst},$$

where $\mathcal{J}_{st}$ denotes the set of products available in store s during week t. Utility for consumer i from product j is

$$u_{ijst} = \delta_{jst} + \varepsilon_{ijst},$$

where ε_{ijst} is independently and identically distributed, yielding standard multinomial logit choice probabilities. Mean utility decomposes as

$$\delta_{jst} = \mathbf{X}_{jst}\beta - \alpha p_{jst} + \xi_{jst},$$

where p_{jst} is the revenue-weighted price, $\mathbf{X}_{jst}$ are observed product characteristics such as package size and the instant indicator, and ξ_{jst} captures unobserved demand shocks including unmeasured quality attributes and promotional effects.

The Berry inversion recovers mean utility from observed shares:

$$\delta_{jst} = \log(s_{jst}) - \log(s_{0st}).$$

Substituting this expression into the mean utility equation yields the estimating specification

$$\log(s_{jst}) - \log(s_{0st}) = \mathbf{X}_{jst}\beta - \alpha p_{jst} + \xi_{jst}. \tag{1}$$

This resembles a conventional demand equation but retains price endogeneity because prices may correlate with ξ_{jst} through retailer responses to unobserved demand shocks or promotional strategies.

Identification of α is achieved through high-dimensional fixed effects. Product fixed effects absorb time-invariant quality differences across UPCs, while store–week fixed effects control for common shocks affecting all products within a store-week, including seasonal demand patterns, local competition, and store-level promotional intensity. After conditioning on these controls, remaining variation comes from within-product price changes over time. This variation plausibly reflects short-run cost-driven adjustments or idiosyncratic promotional decisions rather than systematic responses to product-specific demand shocks, providing the identifying assumption underlying the controls-as-instruments strategy. Section 2.4 discusses the plausibility of this assumption and remaining sources of endogeneity in detail.

Own-price elasticities follow directly from the logit structure:

$$\varepsilon_{jst}^{\text{own}} = -\alpha \, p_{jst} \, (1 - s_{jst}).$$

This expression incorporates substitution both to the outside option and to competing products, with the $(1 - s_{jst})$ term capturing the probability that a consumer currently purchasing product j switches to any alternative. Consumer welfare changes under counterfactual scenarios are computed using the standard log-sum formula from discrete-choice theory, which aggregates utility changes across all available products and the outside option. The welfare measure, derived in Section 3.2, allows for transparent comparison of consumer surplus effects across different pricing interventions.

Estimation is performed using the REGHDFE estimator in Stata, which efficiently absorbs high-dimensional fixed effects through iterative demeaning. Standard errors are clustered at the store level to account for within-store correlation in demand shocks over time.

2.4 Remaining Endogeneity

The estimation strategy relies on high-dimensional fixed effects to purge the primary sources of correlation between prices and unobserved demand shocks. To clarify the remaining sources of identifying variation, decompose the unobserved demand component as

$$\xi_{jst} = \lambda_{st} + \mu_j + \eta_{jst},$$

where λ_{st} captures market-level shocks common to all products in a store–week, μ_j represents time-invariant product attributes, and η_{jst} denotes residual idiosyncratic shocks.

Store–week fixed effects absorb λ_{st}, eliminating demand shifts that affect all products within a store in a given week. These include seasonal consumption patterns, foot traffic fluctuations driven by weather or local events, and broad promotional intensity such as store-wide sales or category features. UPC fixed effects absorb μ_j, removing persistent product heterogeneity such as brand reputation, typical shelf placement, and stable perceived quality differences. After conditioning on these controls, identifying variation arises from within-product price changes over time. These price movements plausibly reflect cost-driven adjustments—such as changes in wholesale prices or distribution costs—or idiosyncratic short-run promotional decisions that vary independently across products within the same store-week.

Under this decomposition, the controls-as-instruments approach succeeds if remaining price variation is orthogonal to η_{jst}, the only unobserved component not absorbed by fixed effects. The exclusion restriction requires that within-product price changes, after conditioning on store–week and UPC fixed effects, do not systematically respond to product-specific demand shocks. This assumption is plausible in the scanner data setting because most short-run price variation reflects pre-planned promotional schedules, wholesale cost pass-through, or inventory clearance rather than real-time demand monitoring. Retailers typically set prices at the category or chain level with limited ability to respond to high-frequency, product-specific sales fluctuations within individual stores.

Nevertheless, three potential sources of residual endogeneity merit consideration. First, promotions remain imperfectly measured in the scanner data, and some promotional events may coincide with unobserved demand shocks that enter η_{jst}. If retailers systematically promote products when category-level demand is high, the estimated price coefficient could partially reflect demand-driven promotional timing. Second, measurement error in recorded prices—arising from unobserved manufacturer coupons, loyalty card discounts, or unreported markdowns—may induce spurious correlation between prices and quantities if such discounts are applied non-randomly. Third, multi-product pricing decisions may generate coordinated adjustments across UPCs in response to localized demand conditions, violating the assumption that price changes are independent across products conditional on fixed effects. For instance, retailers may simultaneously discount multiple UPCs within a brand when sensing category-wide demand weakness, introducing correlation between prices and η_{jst}.

These considerations suggest that while the fixed-effects strategy substantially reduces endogeneity, the estimated price coefficient may still reflect traces of demand-driven promotional behavior or coordinated pricing responses. The direction of any remaining bias is ambiguous: endogenous promotions during high-demand periods would bias the price coefficient toward zero, while inventory-driven price cuts during weak demand would bias it away from zero. Empirically, the stability of the price coefficient across specifications with varying levels of fixed effects—shown in Section 3.1—suggests that such contamination is limited. Moreover, the estimated elasticities fall within the range reported for similar packaged goods categories in studies with access to explicit cost instruments, providing external validation of the identification strategy.

2.5 Model Fit

The performance of the estimated demand model is assessed by examining the alignment between predicted mean utilities and observed market shares, as well as the plausibility of the implied behavioral responses. Figure 2 plots observed Berry logits, $\ln(s_{jt}) - \ln(s_{0t})$, against the model's fitted values. The fitted relationship passes through the center of the observed distribution without exhibiting systematic bias across the support, indicating that the model captures the underlying relationship between product characteristics, prices, and realized market shares. The scatter displays modest dispersion around the fitted line, reflecting idiosyncratic variation in choices that the deterministic component of utility cannot fully explain. This pattern is expected in discrete-choice models and does not indicate misspecification.

Figure 3 presents the distribution of own-price elasticities implied by the estimated parameters. Most mass lies between approximately -1.5 and -2.5, with a concentration near -2. This range is consistent with empirical evidence for differentiated consumer packaged goods and falls squarely within estimates reported for breakfast cereals, yogurt, and other grocery categories in prior scanner-data studies. The distribution avoids implausible outcomes such as near-zero or positive elasticities, which would suggest insensitivity to price or perverse demand responses. The thin left tail indicates that few products exhibit extreme price sensitivity, suggesting that the estimated elasticities are not driven by outlier observations or products with anomalous substitution patterns. Taken together, these diagnostics indicate that the model rationalizes observed purchasing behavior and generates economically coherent price responses across the full range of products and market conditions.

Figure 2: Actual vs. Fitted Berry Logit Values

figures/fitted_vs_actual_large.png

Figure 3: Distribution of Own–Price Elasticities

figures/hist_own_elast.png

3 Discussion

3.1 Initial Estimation Results

Table 2 presents estimates from the Berry logit regression of product market shares relative to the outside option on revenue-weighted prices at the store-week-UPC level. The coefficient on price is negative and highly statistically significant in both specifications, indicating that higher prices reduce the likelihood that a consumer purchases a given oatmeal product. In the naive model, which includes only store–week and UPC fixed effects, the estimated coefficient is -0.8645. The richer specification, which additionally controls for interactions between product characteristics and time-varying market conditions, yields a coefficient of -0.6608. The attenuation in the latter specification reflects the absorption of additional sources of correlation between prices and demand, suggesting that the naive estimate may partially capture systematic variation in unobserved product attributes or promotional intensity across store-weeks.

Using the estimated coefficient from the richer specification and the average price and market share of the top 50 UPCs—\$2.76 and 0.043, respectively—the implied own-price elasticity of demand is approximately -1.75. This indicates that a one percent increase in price reduces quantity demanded by roughly 1.75 percent, placing oatmeal in the elastic portion of the demand curve. This elasticity is economically plausible and consistent with estimates for other differentiated packaged goods: studies of breakfast cereals report elasticities ranging from -1.5 to -2.5, while yogurt and refrigerated juice categories exhibit similar magnitudes. The stability of the price coefficient across specifications—moving from -0.86 to -0.66 as controls are added—suggests that the estimated relationship is not an artifact of omitted variables but rather captures genuine consumer price sensitivity.

Importantly, the elasticity varies meaningfully across product types. As shown in Figure 9 and discussed further in Section 4, instant oatmeal exhibits substantially higher price sensitivity than regular oatmeal, with median elasticities near -2.3 compared to -1.5 for regular varieties. This heterogeneity reflects differences in the competitive environment: instant products face closer substitutes from ready-to-eat cereals and breakfast bars, whereas regular oatmeal competes primarily within a narrower set of hot breakfast options. These patterns motivate the counterfactual analysis that follows, which exploits variation in elasticities to evaluate targeted pricing strategies.

3.2 Notes on Counterfactuals

I consider five counterfactual pricing and packaging scenarios designed to capture realistic retailer decision problems. Scenario 1 applies a uniform 10% price increase to all products, representing a category-wide margin improvement strategy. Scenario 2 raises prices by 15% for the top 20 products by market share, reflecting a premium-targeting approach that exploits high-share UPCs' perceived inelasticity. Scenario 3 increases package size by 15% for smaller products without altering prices, effectively reducing the price per ounce and testing consumer responsiveness to value-per-unit changes. Scenario 4 reduces prices by 25% for products with the highest estimated price elasticities, simulating an aggressive promotional strategy aimed at elastic segments. Scenario 5 applies a 25% price reduction to products labeled "INST" in their descriptions, targeting the instant oatmeal subcategory specifically. In all cases, product characteristics other than the targeted price or size adjustment are held fixed, and the analysis abstracts from potential supply-side constraints or competitor responses.

To simulate demand responses, I exploit the invertibility of the logit model to trace how changes in prices or product attributes propagate through the choice set and reallocate demand across products and to the outside option. The baseline for each counterfactual is defined at the store–week–UPC level using observed prices, quantities, and revenues from the estimation sample. For each scenario, I construct a counterfactual mean utility $\delta_{jst}^{\mathrm{CF}}$ by modifying the relevant component of δ_{jst} according to the scenario-specific policy rule. For price changes, this involves updating the price term $-\alpha p_{jst}$ to $-\alpha p_{jst}^{\mathrm{CF}}$, while for size changes, the modification adjusts the relevant element of $\mathbf{X}_{jst}$ if package size enters mean utility directly.

Counterfactual market shares are computed by applying the logit formula to the modified utilities:

$$s_{jst}^{\mathrm{CF}} = \frac{\exp(\delta_{jst}^{\mathrm{CF}})}{\sum_{k \in \mathcal{J}_{st}} \exp(\delta_{kst}^{\mathrm{CF}})},$$

where $\mathcal{J}_{st}$ denotes the set of products available in market (s, t). Counterfactual quantities follow directly from the implied choice probabilities:

$$Q_{jst}^{\mathrm{CF}} = s_{jst}^{\mathrm{CF}} \cdot M_{st}.$$

Revenue under each scenario is computed as $R_{jst}^{\mathrm{CF}} = P_{jst}^{\mathrm{CF}} \times Q_{jst}^{\mathrm{CF}}$, where P_{jst}^{CF} denotes the counterfactual price. Changes in quantities and revenues are measured relative to their baseline counterparts, allowing for assessment of both the direction and magnitude of demand responses.

Consumer welfare is evaluated using the standard log-sum formula from discrete-choice theory, which aggregates utility changes across all available products. The change in expected consumer surplus, expressed in dollar-equivalent units, is given by

$$\Delta\mathrm{CS}_{st} = \frac{1}{-\bar{\alpha}} \left[\ln\left(\sum_{k \in \mathcal{J}_{st}} \exp(\delta_{kst}^{\mathrm{CF}}) \right) - \ln\left(\sum_{k \in \mathcal{J}_{st}} \exp(\delta_{kst}) \right) \right],$$

where $\bar{\alpha}$ denotes the estimated price coefficient. This expression measures the compensating variation required to make consumers indifferent between the baseline and counterfactual choice sets. Welfare changes are aggregated across store–weeks to obtain total consumer surplus effects for each scenario, enabling comparison of the distributional consequences of alternative pricing strategies.

For each counterfactual, I report product-level changes in quantities and revenues as well as market-level aggregates for total sales, total revenue, and consumer welfare. All simulations preserve the internal consistency of the logit model by ensuring that counterfactual market shares sum to one within each store–week and that the outside option absorbs any residual probability mass. This framework allows for transparent comparison of demand responses, revenue outcomes, and welfare tradeoffs across the five policy interventions, clarifying how different pricing strategies balance firm profitability against consumer surplus.

3.3 Counterfactual Revenue Effects

Baseline total revenue varies substantially across stores and weeks. As shown in Figure 4, the distribution is strongly right-skewed: most store–weeks generate oatmeal revenue below \$2,000, while a small fraction exceed \$5,000. This heterogeneity reflects variation in store size, assortment breadth, and customer traffic, with larger suburban locations and weeks featuring promotional activity producing disproportionately high revenues. The baseline distribution serves as the benchmark against which all counterfactual outcomes are evaluated.

Figure 5 summarizes changes in total revenue relative to baseline across the five scenarios. By construction, these changes reflect only the modeled interventions—holding market size, assortment composition, and all other factors fixed—and therefore isolate the demand-side effects of each pricing or packaging strategy. The uniform 10% price increase generates a moderate revenue gain of approximately \$589,519, indicating that aggregate demand is sufficiently inelastic for broad price increases to raise revenue despite reduced quantities sold. This outcome aligns with the estimated category-level elasticity of -1.75: the percentage decline in quantity is smaller than the percentage increase in price, yielding a net positive revenue effect.

Figure 4: Baseline Total Revenue

figures/baseline_total_revenue.png

Figure 5: Revenue Comparison across Scenarios

figures/revenue_comparison_bar.png

In contrast, increasing prices by 15% for the top 20 products results in a revenue decline of roughly \$1,070,041. Concentrating price increases on high–market-share products triggers more elastic demand responses than anticipated by their average elasticities, as consumers substitute not only toward competing oatmeal products but also toward the outside option. The quantity reductions magnify and outweigh the higher per-unit prices, demonstrating the risks of aggressive targeted increases on popular UPCs. A similar revenue loss—approximately \$1,070,055—arises under the package-size expansion scenario. Increasing size without adjusting prices effectively reduces the price per ounce, inducing substitution toward larger packages and away from smaller, higher-margin products. This reallocation shifts the product mix toward lower per-unit profitability without generating sufficient volume to offset the margin compression.

The largest revenue losses occur under the targeted discount scenarios. Reducing prices by 25% for high-elasticity products and for instant oatmeal leads to revenue declines of \$2,581,382 and \$2,670,303, respectively. Although these discounts substantially increase sales volumes—with quantity gains exceeding 30% in some store-weeks—the associated reductions in per-unit revenue dominate. The elasticity of demand, while high, is not sufficiently high to generate proportional quantity increases that would preserve revenue. Moreover, substitution patterns exacerbate the losses: consumers drawn to discounted products disproportionately come from higher-priced UPCs within the category, cannibalizing margin on products that would have sold at baseline prices. These outcomes illustrate the risks of aggressive price cuts in segments characterized by high price sensitivity, particularly when substitution occurs predominantly within the category rather than from the outside option.

Taken together, the counterfactual results reveal that broad, moderate price increases are more effective at raising revenue than narrowly targeted or aggressive discounts. Uniform adjustments distribute demand responses across the full product set, preventing concentration of quantity losses on any single high-share UPC and yielding more predictable revenue effects. In contrast, targeted strategies amplify substitution patterns and require careful calibration to avoid unintended losses, particularly when applied to products with high baseline market shares or pronounced price sensitivity. These findings underscore the importance of accounting for both own-price elasticities and cross-price substitution patterns when designing pricing and packaging interventions.

3.4 Distribution of Welfare Changes and Revenue–Welfare Tradeoff

Figure 6 displays the distribution of consumer welfare changes across the five counterfactual scenarios, measured in dollars per consumer per week. Several clear patterns emerge. Scenarios involving price reductions—discounts applied to highly elastic products and to instant oatmeal—generate positive welfare gains for most consumers, with the distributions exhibiting substantial right skewness. The instant oatmeal discount produces the largest effects, with a median gain of approximately \$0.20 per consumer per week and a long right tail extending beyond \$0.70. This reflects both the prevalence of instant oatmeal in consumer baskets and the concentration of elastic demand in this subcategory. The elasticity-targeted discount yields a median welfare gain of roughly \$0.10, with the interquartile range spanning \$0.06 to \$0.14. In both cases, the right-skewed distributions indicate that consumers with strong preferences for the discounted products experience disproportionately large benefits, while those with weaker attachment to the affected categories see minimal welfare changes.

Figure 6: Welfare Comparison across Scenarios

figures/welfare_comparison_boxplot.png

By contrast, scenarios involving price increases or package-size expansions exhibit muted welfare effects clustered tightly around zero. Increasing package size by 15% produces negligible welfare changes because the effective price reduction is modest and spread across products with relatively low baseline shares. Consumers benefit slightly from improved value per ounce, but the effect is dampened by the limited salience of package-size differences and the tendency for size increases to affect smaller, less popular products. Similarly, raising prices for the top 20 products by 15% affects a limited subset of consumers—those with strong preferences for high-share UPCs—and results in a near-zero median welfare change, though the distribution exhibits modest negative skew as heavy users of premium products experience larger losses. The uniform 10% price increase leads to small but systematic welfare losses, with a median decline of approximately \$0.25 per consumer per week. Although modest in magnitude, these losses are broadly distributed across consumers, reflecting the pervasive impact of a category-wide price increase. Importantly, welfare responses vary systematically with consumer characteristics: individuals with high price sensitivity experience larger swings under price changes, while consumers with strong brand attachment or habitual purchasing patterns exhibit smaller responses. This heterogeneity underscores the distributional consequences of pricing policies and suggests that aggregate welfare measures may mask significant variation in individual-level impacts.

To examine how consumer welfare gains relate to firm outcomes, Figures 7 and 8 plot changes in consumer welfare against changes in total revenue for the elasticity-based discount scenario. Each point corresponds to a store–week observation. In the full sample (Figure 7), the relationship is predominantly negative: store–weeks with larger welfare gains tend to experience greater revenue losses, with the cloud of points forming a clear downward-sloping pattern. While a small number of observations in the upper-right quadrant generate positive revenue alongside welfare improvements—likely reflecting store-weeks where discounts attract consumers from the outside option rather than cannibalizing existing purchases—these cases are rare and do not alter the overall pattern. The clustering of observations in the lower-right quadrant indicates that substantial welfare gains, often exceeding \$0.10 per consumer per week, are typically associated with revenue losses of \$2 to \$5 per store–week.

Figure 7: Revenue–Welfare Tradeoff for the Elastic Scenario

`figures/welfare_revenue_tradeoff_all.png`

After trimming extreme outliers—defined as store-weeks in the top and bottom 2.5% of the welfare and revenue change distributions—the negative relationship becomes more pronounced (Figure 8). Store–weeks with welfare gains exceeding \$0.15 per consumer per week almost uniformly correspond to revenue declines greater than \$3.50. The fitted regression line, with a slope of approximately -20, highlights the steep inverse association between consumer surplus gains and firm revenue: each additional dollar of consumer welfare improvement corresponds to roughly \$20 in lost revenue at the store-week level. This stark tradeoff reflects the mechanics of discounting elastic products: price reductions generate large quantity increases, but these increases are insufficient to offset the margin compression, particularly when substitution draws volume from higher-priced UPCs within the category. The dispersion of points around the fitted line reveals meaningful heterogeneity across markets. Some store–weeks experience moderate welfare gains with limited revenue loss, likely corresponding to markets with substantial untapped demand from the outside option, while others realize large consumer benefits at substantial cost to the firm, reflecting markets dominated by within-category substitution. This heterogeneity likely stems from differences in the local composition of elastic and inelastic consumers, competitive intensity, and baseline promotional activity.

Figure 8: Revenue–Welfare Tradeoff for the Elastic Scenario (Excluding Outliers)

figures/welfare_revenue_tradeoff_excl_top5.png

Taken together, the welfare results indicate that sharp price reductions generate the largest consumer benefits, particularly for products with high demand elasticity or widespread consumption, such as instant oatmeal. Median welfare gains range from \$0.10 to \$0.20 per consumer per week, translating to annual gains of \$5 to \$10 per consumer if sustained, but the distributions are wide, reflecting substantial heterogeneity in price sensitivity and product preferences across the consumer population. At the same time, these gains are accompanied by substantial revenue losses, often exceeding \$3 per store–week or roughly \$150 annually per store. In contrast, uniform price increases deliver predictable revenue gains at the cost of modest but broadly distributed welfare losses, while targeted price increases or package-size adjustments generate little movement in either dimension. The analysis reveals a fundamental tradeoff: policies that substantially increase consumer surplus tend to erode firm revenue through margin compression and unfavorable substitution patterns, whereas revenue-enhancing strategies impose diffuse welfare costs that, while individually small, accumulate across a broad consumer base. Evaluating pricing and product interventions therefore requires careful consideration of both the magnitude and distribution of these effects, as well as explicit judgments about the relative weights attached to firm profitability and consumer welfare.

4 Conclusion

The counterfactual exercises reveal a clear tension between firm revenue and consumer welfare. Modest, uniform price increases raise category revenue while imposing small, diffuse welfare losses (Figure 6), whereas deep, targeted discounts on highly elastic products generate substantial welfare gains but produce large revenue losses (Figure 5). After trimming extreme observations, the negative relationship between welfare and revenue becomes steep and unambiguous, with each dollar of consumer surplus improvement corresponding to roughly \$20 in lost store-week revenue (Figure 8). This stark tradeoff reflects the underlying mechanics of demand substitution and margin compression in differentiated product markets.

Figure 9: Own-price Elasticities by Product Type

`figures/elasticity_by_type.png`

Differences in demand sensitivity across product types drive these patterns. Instant oatmeal exhibits considerably higher price elasticity than regular oatmeal (Figure 9), with median elasticities near -2.3 compared to -1.5 for regular varieties. This difference likely reflects the competitive environment: instant products face stronger competition from ready-to-eat breakfast alternatives—cereals, breakfast bars, and portable options—where convenience is the primary value proposition. Instant consumers, already paying a premium for time savings, are more willing to switch when prices rise or when competing products become relatively cheaper. In contrast, regular oatmeal buyers exhibit greater commitment, driven by health considerations, taste preferences, or frugality, and face a narrower set of close substitutes within the hot breakfast category. Price reductions on instant products therefore trigger outsized quantity responses, generating strong welfare gains but reducing revenue: unit increases fail to offset margin drops, especially when substitution draws volume from higher-margin regular oatmeal UPCs rather than from the outside option.

Within the elastic discount scenario, quantity gains concentrate in the lowest price quartile while higher-priced quartiles lose volume (Figure 10), illustrating price-led down-trading within the category. Total units increase by approximately 15%, but the product mix shifts toward less profitable items, compressing average margins and reducing aggregate revenue despite higher sales volumes. This substitution dynamic appears at the individual UPC level as well: Figure 11 shows volume gains clustering among lower-priced products—often smaller package sizes or value brands—while losses concentrate among higher-priced premium UPCs. The negative welfare-revenue relationship thus reflects not only margin compression from

Figure 10: Net Quantity Change by Price Quartile in Elastic Scenario

figures/net_quantity_by_price.png

lower prices but also unfavorable composition effects as consumers trade down within the assortment.

Figure 11: UPC-level Winners and Losers in Elastic Scenario

figures/winners_losers_elast.png

Category structure further constrains revenue outcomes. A small number of high-share UPCs account for most category volume (Figure 12), with the top 10 products representing roughly 40% of total sales. Price increases on these products create substantial downside risk: even modest elasticities, when applied to large market shares, generate significant quantity losses that can outweigh per-unit margin gains. This dynamic explains the revenue decline in the premium-targeted scenario (Figure 5), where concentrating price increases on the top 20 UPCs backfired despite their seemingly strong market positions. In contrast, broad, modest adjustments distribute price effects across the full assortment, preventing concentration of quantity losses on any single high-volume product and generating aggregate revenue gains with minimal welfare losses.

These insights suggest a practical pricing strategy for the oatmeal category. Retailers should implement modest, uniform price increases—in the range of 5–10%—while excluding the most elastic instant UPCs to preserve revenue improvement and limit welfare losses (Figure 9). If discounting is necessary to drive store traffic or maintain loyalty program engagement, it should be shallow (10–15%), time-limited (one to two weeks), and targeted at lower-priced, high-elasticity UPCs where incremental volume gains are strongest (Figure 10). Retailers must monitor higher-priced UPCs closely during promotional periods to detect and respond to persistent volume declines that signal cannibalization (Figure 11). Price adjustments on high-share UPCs—particularly the top 10 products—should remain conservative, with increases capped at 5% and promotional depth limited to avoid triggering disproportionate demand responses (Figure 12).

Several actionable recommendations emerge from the analysis. First, favor modest, broadly applied price increases over targeted adjustments. Uniform changes reliably raise revenue with minimal welfare losses by distributing demand responses across the product set and preventing concentration effects. Second, avoid deep discounts on highly elastic items, particularly instant oatmeal, where sharp quantity responses erode margins and induce down-trading toward lower-margin UPCs. The revenue costs of such promotions—often exceeding \$2,500 per scenario—substantially outweigh the benefits of increased volume. Third, maintain tight control over instant oatmeal pricing. Given median elasticities exceeding -2, even small price deviations induce substantial substitution toward ready-to-eat breakfast alternatives, eroding category share over time. Fourth, avoid increasing package sizes without corresponding price adjustments. Size expansions effectively lower unit prices, encourage down-trading, and depress category revenue through composition effects. Finally, if promotional activity is required for competitive or strategic reasons, structure discounts to be shallow, brief, and concentrated on lower-priced, high-elasticity UPCs where incremental volume gains are strongest and downside risk to higher-margin products is minimized.

The analysis reveals consistent pricing dynamics within the oatmeal category that generalize to other differentiated packaged goods markets. Consumers exhibit substantial price sensitivity, with elasticities varying systematically across product types based on competitive positioning and the availability of close substitutes. Aggressive or prolonged discounts prove costly in terms of both revenue and product mix, as quantity gains concentrate in low-margin items while higher-margin products lose volume. The most effective pricing strategy combines restrained, category-wide price adjustments—leveraging the inelasticity of aggregate demand—with disciplined avoidance of deep promotions on elastic items and tight control over high-elasticity subcategories like instant oatmeal. When discounts are strategically necessary, they should be structured to minimize cannibalization of higher-margin products while balancing revenue objectives against manageable consumer welfare impacts. These findings underscore the importance of accounting for product heterogeneity, substitution dynamics, and category structure when designing pricing interventions in differentiated product markets.

5 Appendix: Summary Statistics and Regression Results

Table 1: Summary Statistics for Key Variables

Variable	Mean	SD	Min	Max
Berry logit: $\ln(s_j) - \ln(s_0)$	-7.92	0.97	-10.82	0.17
Price	2.76	0.62	1.04	4.19
price_wmean	2.76	0.62	1.04	4.19
size_oz	18.51	8.55	5.00	42.00
instant	0.34	0.47	0.00	1.00
income	10.62	0.29	9.87	11.24
age9	0.14	0.02	0.05	0.19
age60	0.17	0.06	0.06	0.31
educ	0.23	0.11	0.05	0.53
ethnic	0.16	0.19	0.02	1.00
quantity	10.62	21.53	1.00	10816.00
rev	807.52	447.08	2.75	29199.62
share	.001	0.00	0.00	0.41
Observations	752,069			

Notes: Summary statistics are calculated at the store-week-UPC level.

Table 2: Regression Results: Berry Logit

	Naive Model (1)	**Large Model** (2)
Dependent variable:	Berry logit: $\ln(s_j) - \ln(s_0)$	
Mean price per store-week-UPC	-0.8645***	-0.6608***
	(0.0073)	(0.0367)
Observations	752,067	727,432

Notes: Standard errors are in parentheses. All regressions include store $\times$ week and UPC fixed effects. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Figure 12: Price vs. Market Share

figures/price_vs_share_labeled.png

Figure 13: Market Share vs. Price for top 50 UPCs

figures/market_share_top50.png

Figure 14: Histogram of Sizes

figures/hist_size_oz.png

References

(2024). Dominick's dataset. Kilts Center for Marketing, University of Chicago Booth School of Business.

Berry, S. T. (1994). Estimating discrete-choice models of product differentiation. *The RAND Journal of Economics*, pages 242–262.