

# Market Share and Pricing Determination in the Ready-to-Eat Cereal Industry\*

Christin Bivens<sup>†</sup>

May 4, 2018

## Abstract

The ready-to-eat cereal market has been found to be an interesting case study as it has some of the highest profit margins in the grocery industry and the variety of products and advertising employed by major brands gives researchers a plethora of effects to study (Nevo (2001)). Generally, grocery is thought of as retail and therefore imperfect competition by economists. However, RTE cereal displays attributes of an oligopoly structure. By studying price determination and market share in the cereal industry, economists can have insights into the market structure and elasticity of retail trades. This article aims to provide a predictive linear model for market share in the ready-to-eat cereal industry that accounts for market share as a function of pricing, product characteristics of nutritional content, and product size.

## 1 Introduction

The market structure of the ready-to-eat (RTE) breakfast cereal industry provides insights into both imperfect competition and competitive oligopolies. Stemming from grocery retail (which is usually categorized as imperfect competition), RTE cereal companies have managed to create a

---

\*Special thanks to Dr. Myongjin Kim, University of Oklahoma

<sup>†</sup>Department of Economics, University of Oklahoma. E-mail address: christin.e.bivens-1@ou.edu

market where only a few firms are dominant and cost-profit ratios are so high that collusion has been suspected (Nevo (2001)). In addition to making for an interesting case study, the RTE cereal industry also lends itself to economic research through relatively obtainable data on both product characteristics and firm costs (Stanley and Tschirhart 1991).

This paper looks to prior research and data cleaning/visualization methods to create a linear regression capturing market share in the industry. As market share and price have a relationship rooted in reverse causation, a regression to first predict price is necessary. The new predicted variable for price is labeled as  $\widehat{P}_{it}$  for use in the market share model. The response variable  $MS_{it}$  was specified as a function of the relationship between the inside good and outside goods at time  $t$  using a logit model. The explicit equations can be found in the empirical methods section.

The estimated coefficients from the findings were all statistically significant with  $p$ -values below .05 for the market share model. However, only 7 out of 12 coefficients for the pricing regression were significant. Given the time restraints of this research, the ratio of significant results is considered a triumph by the author. Results of the final model for market share found that calorie content, protein content, box size, and sugar content all were positively correlated with market share. Carbohydrate count and price had inverse relationships with market share. Given economic theory on the relationship between pricing and sales, the inverse relationship between price and market share for the RTE cereal industry was to be expected. The strong influence of product characteristics

## 2 Literature Review

"Measuring Market Power in the Ready-to-Eat Cereal Industry" by Aviv Nevo in 2001 is used as a general reference and starting off point for this study. Nevo creates multiple models to capture RTE demand then creates a nested mode with the goal of having a predictive model that is not hindered by unobservables. Time constraints and the scope of this project do not allow for Nevo's thorough approach to be replicated. However, the insight found by Nevo that the high cost-price margins are not a result of collusion in the RTE industry, but rather it is achieved through product differentiation is a benchmark for developing both the pricing and market share models.

Research prior to Nevo showed that despite the large profit margins for cereal, that the demand was not inelastic. "Using a Natural Experiment to Estimate price Elasticity: The 1974 Sugar Shortage and the Ready-to-Eat Cereal Market" by Neslin and Shoemaker 1983 finds that price elasticity for the cereal industry is relatively high. As firms gain market power the demand elasticity for their product decreases. This allows them to increase prices while reducing sales and consequently enjoy higher price-cost margins. Neslin and Shoemaker help to refute the idea that the RTE cereal industry is a price collusive oligopoly by using results from a natural experiment. The sugar shortage of 1974 created an increase in the price of sugar and therefore in an increase in the price of ready-to-eat breakfast cereals. This shock to the market provided data to compare consumer reactions. Consumers bought far less cereal at the heightened prices across multiple brands showing a strong responsiveness to price. Which Neslin and Shoemaker mathematically proved.

For constructing a market share variable the article "Implications of Market Structure for Elasticity Structure" by Russell and Bolton from 1988 was referenced. The authors "develop a parsimonious description of elasticity structure that relates market share and marketing mix variables

(e.g., price) to brand elasticities” (Russell and Bolton (1988)). Their research also concludes that the presence of a market dominating brand affects both pricing and market share for other products in the market. Both their findings with regards to relationships of products in the market and their use of a logit model for market share are included in this research.

The 1991 article ”Hedonic Prices for a Nondurable Good: The Case of Breakfast Cereals” by Stanley and Tschirart was used as reference for determining which variables to include in the pricing and market share regressions. Stanley and Tschirhart exclude calories from their study as they make the observation that factors such as sugar and fiber content greatly affect the caloric content and therefore it is not necessary to include. This research uses that logic to go in the opposite direction. When predictive models become overly laden with variables of interest they lose predictive power. Therefore, this research will include calories in the regression as a means of including the effects of product nutrition characteristics, such as fiber, that can be captured within the caloric content. Stanley and Tschirhart also find that fat content varies very little between cereal products and therefore are not a strong enough contributing factor to include in research. Their insight about fat is likewise used in this paper to further slim down the pricing estimate model.

To decide whether to include certain parent companies or all products in the data for the final model ”The Welfare Economics of Product Variety: An Application to the Ready-to-Eat Cereals Industry” by Scherer in 1979 was considered. Scherer lays the ground work for which brands control the market share in the RTE industry (General Mills, Kellogg Co, PepsiCo, and a few private brands like Quaker that can be umbrella-ed together). The paper asserts that the market share of brands highly affect the success of new products. the author considers whether there is a cannibalistic effect or if firms are purposefully generating high numbers of new products in order

to edge out smaller firms. Because Scherer finds a significant relationship between product variety and market power, this research will focus on the four largest parent companies. This allows for the results to reflect the core behavior of the market as well as which market strategies are effective.

### **3 Data**

Data has been graciously provided by Dr. Kim of the University of Oklahoma from a previous joint research endeavor. The data contains records for 95 weeks of observations from 2001 to 2003 on 18 parent companies within the market. Each record is for a product description (i) for the given week (t). Product description in this case refers to a brand under a parent company that is specific to product size and type (e.g. Captain Crunch Berries 32oz.). The focus will be on products by the dominating companies in the market: General Mills, Kellogg Co., Private Labels, and PepsiCo (Scherer (1979)).

Looking at time series data allows the research to have insights on real pricing and product characteristics without possible issues arising from collecting data from an experiment(Neslin and Shoemaker (1983)). While there are disadvantages to working with observed data like not having an available control group, differences in selling environments, and possible changes in brand quality, the scope of this research is not heavily affected by those shortcomings. The focus of this paper is to model market share based on observable attributes, and therefore there is no need for counterfactual analysis.

Observed variables include time in terms of weeks, months, quarters, and years, product characteristics of nutritional content, box size, and character size on the box, cost inputs for the firm, unit sales per product description, total sales per product description, as well as others that were

not pertinent to this study. For the regressions the data was aggregated by product description per week.

## 4 Empirical Methods

All empirical exercises were done using R and RStudio. In addition to the literature review a process of data cleaning and visualization was used to determine correlation between variables of interest and the response variable. This process included using only one response variable and one variable of interest at a time to create scatter plots, summary statistics, correlation tests, and simple linear regressions. Examples of the plots between price and quantity sold for the four main parent companies can be found in the figures and tables section. This was done on the final focused data set as well as broken down to the four parent companies. The specification by parent company in the visualization stage was to compare effects to ensure that there was some sense of uniformity across the market. That is to say, that there was not an observed effect of a variable on price for one company that did not exist for another therefore throwing off the mean effects for the market.

In line with research and findings through data exploration, the pricing linear regression uses input variables of the observable product characteristics (calorie content, carbohydrate content, protein content, box size, and sugar content) to capture expectations of consumer preferences, cost inputs (corn, wheat, gas, sugar, electricity, and labor costs), as well as dummy variables for both time (weekly) and other products available in the market at weekly time  $t$ .

$$\widehat{P}_{it} = \beta_1 X_{Calories} + \beta_2 X_{Carbs} + \beta_3 X_{Protein} + \beta_4 X_{Size} + \beta_5 X_{Sugar} + \alpha_1 C_{Corn} + \alpha_2 C_{Gas} + \alpha_3 C_{Wheat} +$$

$$\alpha_4 C_{Sugar} + \alpha_5 C_{Electricity} + \alpha_6 C_{Labor} + \alpha_7 D_{Week} + \alpha_8 D_{ProductDescription}$$

(1)

Where  $\hat{P}_{it}$  is a continuous outcome variable for product description  $i$  in week  $t$ , and  $X_{it}$  are characteristics of product  $i$  for which the parameter of interest is  $\beta_j$ ,  $C_{it}$  are input costs for product  $i$  and  $D_{it}$  are dummy variables for time (weeks) and available products in the market for which the parameter of interest is  $\alpha_k$ .

The model for market share ( $MS_{it}$ ) was constructed using the new price variable - as an effort to account for the reverse causality of price and MS, product characteristics (calorie content, carbohydrate content, protein content, box size, and sugar content), and dummy variables for time (weekly) and other products available in the market at weekly time  $t$ .

$$MS_{it} = \log(UnitSales_{it}/Population_t) - \log((Population_t - UnitSales_{it})/Population_t)$$

(2)

Where  $MS_{it}$  is a continuous outcome variable for product description  $i$  in week  $t$  as a function of unit sales for product description  $i$  in week  $t$  in relation to unit sales for the entire market population

at time  $t$ .

$$MS_{it} = \beta_1 X_{Calories} + \beta_2 X_{Carbs} + \beta_3 X_{Protein} + \beta_4 X_{Size} + \beta_5 X_{Sugar} + \alpha_1 \widehat{P}_{it} + \alpha_7 D_{Week} + \alpha_8 D_{ProductDescription}$$

(3)

Where  $MS_{it}$  is a continuous outcome variable for product description  $i$  in week  $t$ , and  $X_{it}$  are characteristics of product  $i$  for which the parameter of interest is  $\beta_j$ , and  $D_{it}$  are dummy variables for time (weeks) and available products in the market for which the parameter of interest is  $\alpha_k$ .

## 5 Research Findings

The results for the pricing prediction regression can be found in Table 1. For  $\widehat{P}_{it}$  the coefficient for caloric content was .144 with a statistically significant  $p$ -value under .01. Indicating that as caloric content increases, firms are more likely to increase prices. For carbohydrates there was an inverse relationship with pricing with a coefficient of -1.847, and a statistically significant  $p$ -value under .01. The reasoning behind this could be as simple as inputs for carbohydrates are cheaper and therefore lower pricing. Indeed when sugar cost is considered there is a statistically significant coefficient of -.031. Sugar inputs would affect carbohydrate content and these are in line with each other. However, sugar content has a coefficient of .061 contradicting this theory - although the  $p$ -value was not significant. Protein and box size both have positive and significant relationships with price (4.824 and .091 respectively). This seems in line with basic economic theory that as inputs



increase so does pricing. Box size not only affects the input of the cereal itself, but the inputs for the box. Protein is often one of the more expensive food costs, therefore, these numbers seem logical. Labor cost (.077), electricity cost (-.042), wheat cost (-.018), and corn cost (.056) all were found to be statistically insignificant. Finally, gas cost was found to have an inverse relationship with price with a significant  $p$ -value of less than .01. This seems counter intuitive as gas will always add to the bottom line for a company. This and the lack of significant results on many of the variables of interest in the regression could be due to the fact that cereal companies enjoy economies of scope - therefore there could be a veiled effect on proper measurements of their product inputs on pricing. The inverse relationship with gas could be due to a decrease in average variable cost for gas as production increases.

For the market share model all coefficients were statistically significant with  $p$ -values below .01, with the exception of sugar content that had a  $p$ -value smaller than .05. Variables of interest that had a positive relationship with market share were caloric content (.577), protein (20.144), box size (.311) and sugar content (.699). The constructed pricing variable had an inverse relationship with market share, captured by a value of -2.873. This is very much in line with economic theory for imperfect or monopolistic competition - that the demand is more elastic than inelastic and consumers are sensitive to price. The only nutritional product characteristic that had an inverse relationship with MS was the amount of carbohydrates at a value of -7.642. these results can be reviewed in Table 2 under the Figures and Tables section.

## 6 Conclusion

The findings in this study reflect different aspects of previous research on the RTE cereal industry and economic theory. While few firms dominate the market and there is a high cost-price margin suggesting an oligopoly structure (Nevo (2001)), the coefficients of the regression reflect that customers are highly responsive to price and product characteristics as is in line with imperfect or monopolistic competition.

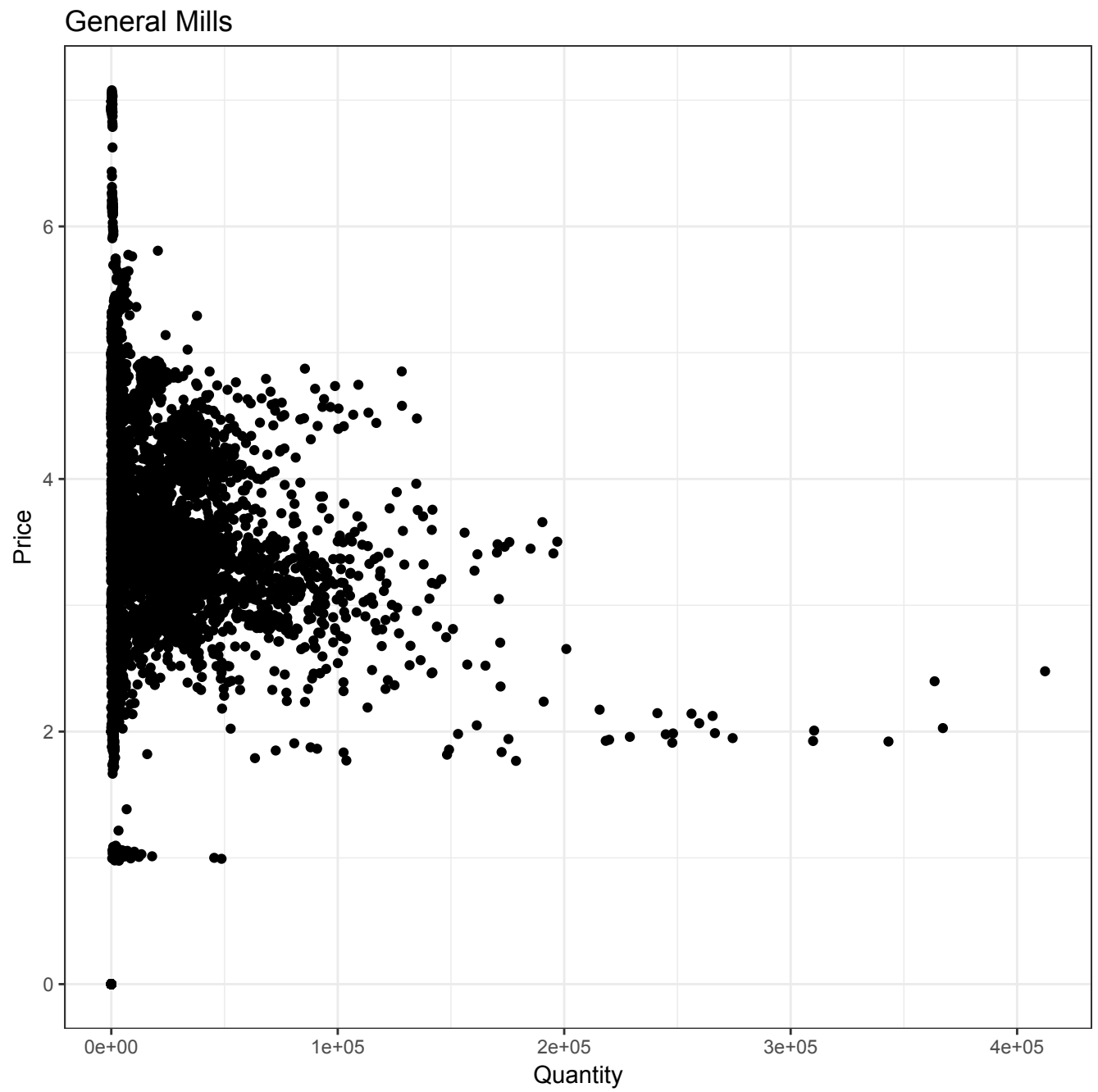
Many firms who enter an imperfectly competitive market, like grocery retail, do so with the hope of increasing their market share in order to escape long term zero profit theories. By studying the effects of pricing and product differentiation in the RTE industry, businesses can gain insight in how to greatly increase their market share and therefore increase profits and longevity. Specifically this linear regression model can be used to help predict market share of a new product by entry firms or an existing firm. This is achieved by using a logit model to predict market share taking into account market share of other products, as well as creating a new price variable that adjusts for the reverse-causal relationship of market share and price.

In addition to commercial use, the findings in this research can help economists to further study the boundaries and characteristics of market structure as industries/firms evolve from one structure to the next. The significant positive correlation between product characteristics and consumer preference show that differentiation is a large part of gaining market power. The less predictable effects of inputs on pricing lend to the theory that RTE cereal companies are able to create a higher price-cost margin by enjoying economies of scope, although this needs further research. Finally, the sensitivity of consumers to price reinforces the theory that demand in grocery is more elastic and therefore reducing price while increasing sales leads to a larger market share.

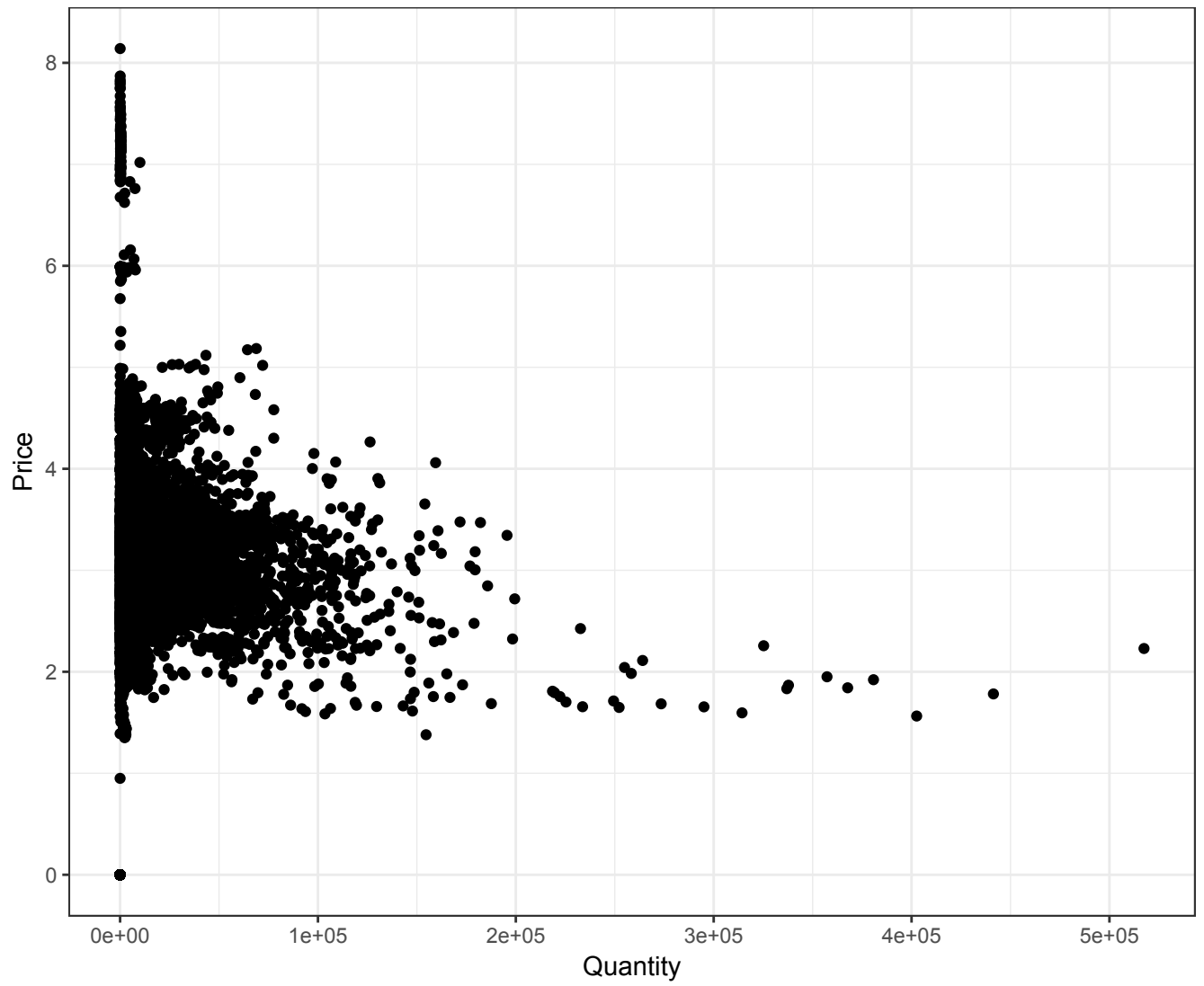
## References

- Neslin, Scott A. and Robert W. Shoemaker. 1983. "Using a Natural Experiment to Estimate Price Elasticity: The 1974 Sugar Shortage and the Ready-to-Eat Cereal Market." *Journal of Marketing* 47 (1):44–57.
- Nevo, Aviv. 2001. "Measuring Market Power in the Ready-to-Eat Cereal Industry." *Econometrica* 69 (2):307–342.
- Russell, Gary J. and Ruth N. Bolton. 1988. "Implications of Market Structure for Elasticity Structure." *Journal of Marketing Research* 25 (3):229–241.
- Scherer, F. M. 1979. "The Welfare Economics of Product Variety: An Application to the Ready-to-Eat Cereals Industry." *The Journal of Industrial Economics* 28 (2):113–134.
- Stanley, Linda R. and John Tschirhart. 1991. "Hedonic Prices for a Nondurable Good: The Case of Breakfast Cereals." *The Review of Economics and Statistics* 73 (3):537–541.

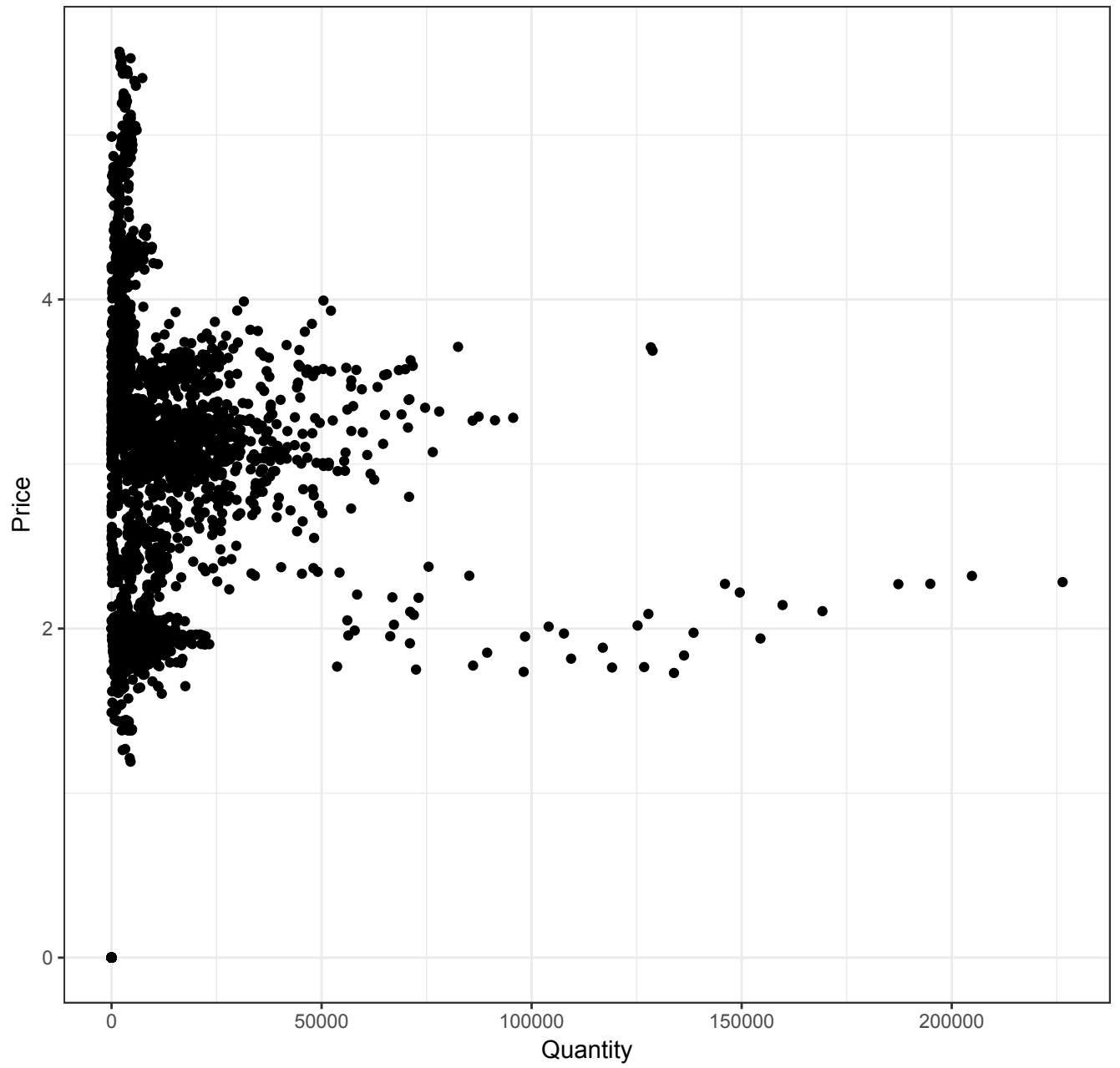
## Figures and Tables



Kellogg



PepsiCo



Private Label

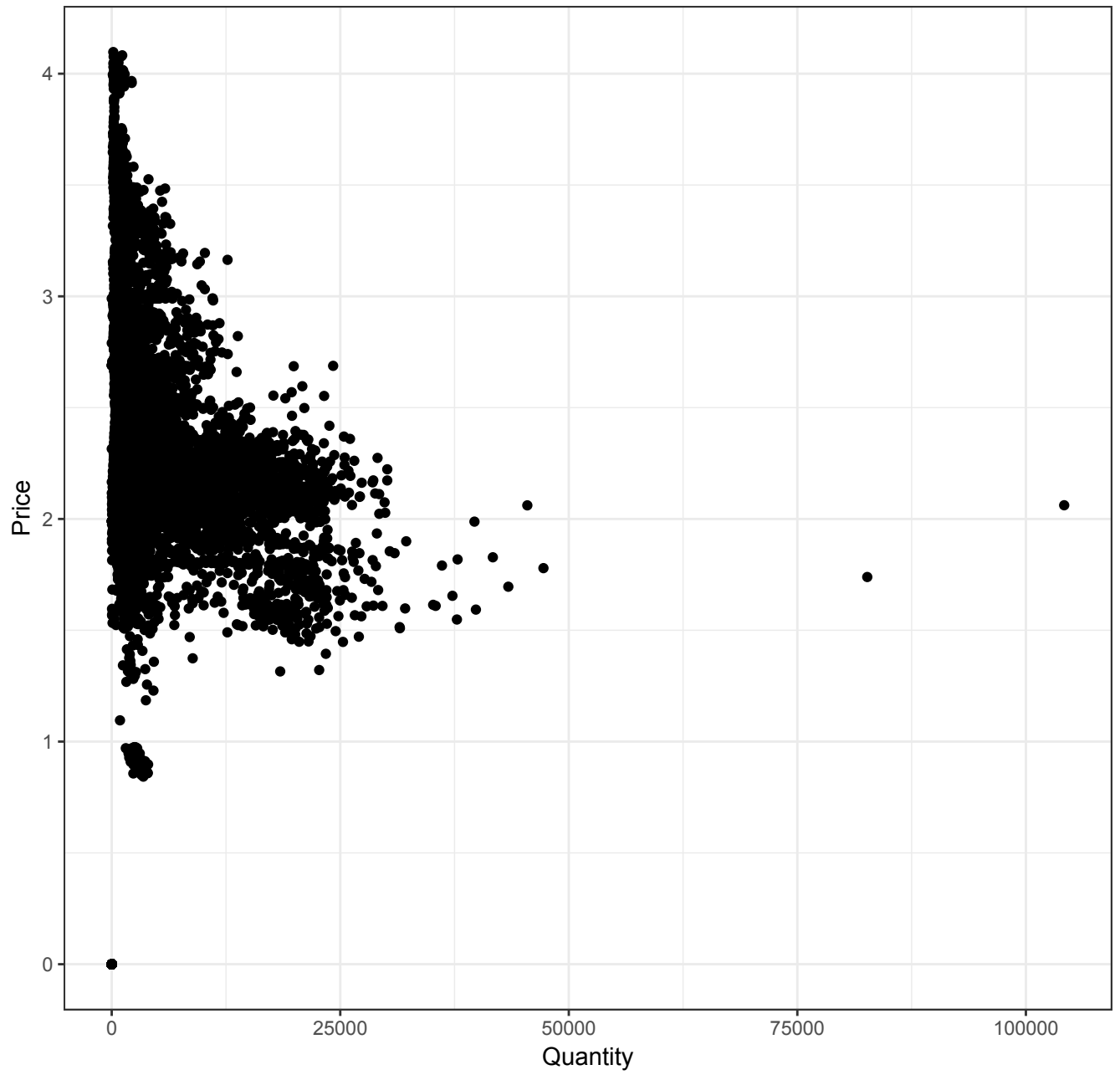


Table 1: Pricing Regression Results

	<i>Dependent variable:</i>
	Pricing
Calories	0.144*** (0.021)
Carbohydrates	−1.847*** (0.353)
Protein	4.824*** (1.131)
Box Size	0.091*** (0.005)
Sugar Content	0.061 (0.100)
Corn Cost	0.056 (0.071)
Gas Cost	−0.173*** (0.046)
Wheat Cost	−0.018 (0.018)
Sugar Cost	−0.031*** (0.004)
Electricity Cost	−0.042 (0.036)
Labor Cost	0.077 (0.172)
Week	0.003* (0.002)
Observations	37,047
R <sup>2</sup>	0.671
Adjusted R <sup>2</sup>	0.666
Residual Std. Error	0.692 (df = 36555)
F Statistic	151.564*** (df = 491; 36555)
<i>Note:</i>	*p<0.1; **p<0.05; ***p<0.01



Table 2: Market Share Regression Results

	<i>Dependent variable:</i>
	Y <sub>it</sub>
Calories	0.577*** (0.069)
Carbohydrates	−7.642*** (1.086)
Protein	20.144*** (3.397)
Box Size	0.311*** (0.024)
Sugar Content	0.699** (0.288)
$\widehat{P}_{it}$	−2.873*** (0.216)
Week	0.002*** (0.001)
Observations	37,047
R <sup>2</sup>	0.465
Adjusted R <sup>2</sup>	0.457
Residual Std. Error	1.980 (df = 36560)
F Statistic	65.280*** (df = 486; 36560)
<i>Note:</i>	*p<0.1; **p<0.05; ***p<0.01