Empirical Assessment of Partial Antitrust Divestitures

Benjamin S Osenbach
Masters Essay

November 19, 2021

Abstract

In this paper, I assess the strength of of competition regulators' recommendations against "partial" antitrust divestitures. Using a discrete-choice framework, I estimate counterfactual mergers and divestitures in the market for Ready-to-Eat (RTE) cereal. Marginal cost reductions through efficiency gains are estimated by Data Envelopment Analysis (DEA). Partial divestitures are found to significantly increase Consumer Surplus relative to full divestitures when efficiency gains are present. No statistical difference is found in the welfare performance of partial divestitures relative to full divestitures, offering evidence contrary to current regulation

1 Introduction

Over the past thirty years, the analysis of horizontal mergers has undergone an empirical revolution. Tools from the New Empirical Industrial Organization such as random-coefficient demand estimation and merger simulation have displaced older measurements of market concentration and conduct in establishing anticompetitive behavior (Tenn and Yun (2011)). However, the design and implementation of effective merger remedies has not been addressed with the same level of empirical rigor. Although divestitures of products, brands, or assets are a commonly-prescribed structural remedy in horizontal merger cases, their design and implementation have been primarily driven by non-empirical case studies and documentary evidence. Furthermore, as competition authorities have become increasingly cognizant of the role of cost synergies (aka "efficiency gains") to offset the distortions of horizontal mergers (DOJ and FTC (2010)), antitrust divestitures have become more important as a tool to restore competition

In particular, as a result of documentary case studies by the Federal Trade Commission (FTC (1999), FTC (2017)), competition authorities have been reluctant to approve so-called "partial divestitures", i.e. the divestiture of selected assets or brands as opposed to an entire line of business. Findings from these studies had indicated that partial divestitures to new entrants or "carve-out" firms rarely produced viable competition in the industry. Yet although the prevailing wisdom of antitrust regulators has opposed partial divestitures, a number of recent empirical studies in economics have challenged this conclusion. In particular, Jayaratne and Shapiro (2001), Bougette (2010), and Vasconcelos (2010) each provide models in which partial divestitures to an incumbent firm can restore competition and increase consumer welfare

Furthermore, although recent advances in econometrics have enabled feasible non-parametric estimation of efficiency gains, these methods have not yet been incorporated into the divestiture discussion. Beginning with Bogetoft and Wang (2005) and continuing through to Bonnet and Schain (2020), Data Envelopment Analysis (DEA) provides a useful framework for analyzing counterfactual divestitures

In this paper, I estimate the welfare effects of several partial divestitures following a counterfactual merger in the ready-to-eat (RTE) cereal industry. Following Nevo (2001) and Miller and Weinberg (2017), I estimate a nested-logit Demand model of RTE cereal. The supply side of the model is estimated following Miller and Weinberg (2017), and marginal cost reductions arising from efficiency gains are calculated through non-parametric Data Envelopment Analysis (DEA) according to Bonnet and Schain (2020). Finally, counterfactual sets of divestitures are calculated following Friberg and Romahn (2015)

The remainder of this paper is organized as follows. Section 2 reviews the extant empirical literature of antitrust divestitures and efficiency gains. Section 3 describes the methodology and design of the Demand and Supply systems to be estimated. Section 4 presents the results of the estimation, and Section 5 concludes with applications that may be drawn from the model for future economic analysis

2 Background

Antitrust divestitures have long been recognized by competition authorities as a tool to remedy horizontal merger concerns. Elzinga (1969) provides one of the earliest documentary studies of antitrust remedy effectiveness, as well as one of the earliest cautions against partial divestitures. More thorough studies include the FTC's (1999) review of 37 divestiture orders from 1990-94, in which 75% of divestiture orders were deemed to be successful. Moreover, partial divestitures were found to be less successful than divestitures of on-going businesses: 9 out of 15 partial divestitures were deemed viable as opposed to 19 out of 22 full divestitures. A follow-up study (FTC (2017)) of 50 merger remedies from 2006-2016 (plus 39 others), found that 67% of partial divestitures were deemed viable, as opposed to 100% of full divestitures

Despite the influence of these reports, documentary evidence is often insufficient to establish a causal link. On the empirical side of divestiture research, Jayaratne and Shapiro (2001) calibrated a nested logit demand system to argue that partial divestitures may increase efficiency if the divested product is close substitutes with the one(s) being acquired. Although they focus on acquisitions and divestitures of business software, this naturally applies to the market for RTE cereal. Suppose that Firm 1 produces both a kids cereal and an adult cereal. If Firm 1 acquires Firm 2, which also produces a kids cereal, then the merged firm should be required to divest one of the kids cereals instead the adult cereal. Despite their findings, Jayaratne and Shapiro (2001) do not incorporate efficiency gains into their considerations, and their choice of calibration as an estimation strategy also casts doubt on their causal link. Tenn and Yun (2011) study the divestitures from the 2006 Johnson & Johnson acquisition of Pfizer's consumer health division. However, they also fail to consider efficiency gains, and their use of a reduced-form OLS regression of sales on month dummies also muddies the waters for establishing causality.

More recently, more rigorous empirical papers have focused on similar (albeit different) questions. Pham and Prentice (2013) estimate a random coefficients logit demand system for a 1999 merger in the Australian cigarette industry; however their aim is simply to assess the random coefficient logit model's predictions of price changes following a divestiture. They find that although the model successfully predicts the direction of price changes following a counterfactual merger and partial divestiture, it overstates the effects of these price changes. Friberg and Romahn (2015) use data from the 2001 Carlsberg-Pripps merger in the Swedish beer market to estimate the effects of counterfactual divestitures on prices and welfare. After estimating a random-coefficients logit demand model, they randomly select 200,000 candidate packages for divestiture according to a given heuristic. They then use these packages to bootstrap estimates of the price and welfare distributions. Although similar to the question in this paper, Friberg and Romahn's focus on approximating the entire distribution does not address the question of whether partial divestitures can improve consumer welfare relative to full divestitures

More recently, several recent papers in economic theory have also strengthened the case for partial divestitures. Bougette (2010) presents a model of mergers with "cost synergies" (e.g. efficiency gains), showing that mergers with less cost synergies require larger divestitures. Nonetheless, Bougette also demonstrates that the optimal percentage of assets to divest has an upper bound less than 100%; and moreover, this interval shrinks with the number of firms in the market. For instance, an industry with 3 firms will require divesting 27-33% of the merged firms' assets, but an industry with 9 firms only requires divesting 13-17% of the merged firm's assets. With 5 firms as in this paper, the optimal divestiture lies within 20-25% of assets. Bougette also shows that divestitures to a "carve-out" firm (e.g. an entrant) are less efficient than divestitures to an incumbent firm unless technological advantages (such as human capital, "know-how", or distribution networks) are also divested. This justifies a key observation of the FTC's 1999 and 2017 reports, but also strengthens the case for partial divestitures to an incumbent firm. Vasconcelos (2010) also shows that if a merger does not involve every firm in the industry, then the competition authority can only increase surplus if assets are divested to an incumbent firm.

Another consideration furnished by recent papers in economic theory concerns the role of efficiency gains in horizontal mergers. Cosnita and Tropeano (2009) analyze efficiency gains as asymmetric information held by the merging parties. In order to extract the private information from the merging firms, they provide a combination of a high divestiture sales price and a "merger fee" as a direct mechanism. They reason that the more efficient merging parties will opt to pay the merger fee while the less efficient firms will choose the sales price. Likewise, Gayle et al. (2011) show that high upward pricing pressure (UPP) is insufficient to reject a merger in the presence of strong efficiency gains, which provide substantial downward pricing pressure. Consequently, the horizontal merger guidelines of the DOJ and FTC (2010) have established in §10 that "The Agencies will not challenge a merger if cognizable efficiencies are of a character and magnitude such that the merger is not likely to be anticompetitive in any relevant market". Nevertheless, the application of efficiency arguments to divestiture is a recent development in the field. Bonnet and Schain (2020) estimate a standard random coefficients logit demand model of French dairy desserts, but then adjust estimated marginal costs through the non-parametric DEA approach of Bogetoft and Wang (2005). They find that the standard assumption of a 5% reduction in marginal costs following a merger is generally untenable: 44% of their counterfactual mergers produce no efficiency gains, whereas 18% produce gains higher than the 5% rule of thumb, and 38% produce gains smaller than 5%. Bonnet and Schain do not focus on divestitures, but their work indicates that a realistic assessment of the impact of divestitures necessitates at least some estimation of efficiency gains

3 Data

My data set consists primarily of simulated retail scanner data at the product-store-week level for 37 months from January 2009 through January 2012[1]. There are 169,689 observations of 14 RTE cereals from 5 manufacturers at 51 locations across 4 states (Indiana, Kentucky, Ohio, and Texas). Location and month data were used to identify 1887 (=51*37) different markets. Universal product codes, manufacturer, product descriptions, sale prices, units sold, package size, and promotional dummies are all identified in the data set. Two observations are missing sales prices, so I proxy for those with the reported base price. Store chains are not identified in the data. Product descriptions along with the relative frequency of the product in the data set are reported below in Figure 1

Product	Firm	Manufacturer	Category	Name	Relative frequency
1	1	General Mills	All-fam	CHEERIOS	15.189%
2	1	General Mills	All-fam	HONEY NUT CHEERIOS	13.801%
3	2	Kellogg	All-fam	BITE SIZE MINI WHEAT	6.803%
4	2	Kellogg	Kids	FROOT LOOPS	7.870%
5	2	Kellogg	Kids	FROSTED FLAKES	9.711%
6	5	Private Label	All-fam	BT SZ FRSTD SHRD WHT	8.034%
7	5	Private Label	All-fam	HONEY NUT TOASTD OATS	4.497%
8	5	Private Label	Adult	RAISIN BRAN	6.333%
9	3	Post	Adult	FM SZ HNYBNCH OT ALM	4.531%
10	3	Post	Kids	FRUITY PEBBLES	4.989%
11	3	Post	Adult	HNY BN OTS HNY RSTD	5.22%
12	4	Quaker	Kids	CAP N CRUNCH	4.311%
13	4	Quaker	Kids	CAP N CRUNCH BERRIES	5.754%
14	4	Quaker	All-fam	LIFE ORIGINAL	2.956%

Figure 1: Product information and frequency of occurrence in the data set

Product availability across markets is presented below in Figure 2. Products 1-6 and 8 are available in all markets, and product 7 is present in every market except market 1863 (College Station, TX, 2012 Month 1). Entry is observed early in the data set for products 9-11, and exit is observed late for products 12-14. In total, products 9-14 are present in approximately 88.4% of the markets on average.

Information on product characteristics is matched from Carnegie Mellon's Data and Story Library[2]. In particular, cereal size, the number of calories per serving, grams of sugars, grams of carbohydrates, grams of protein, grams of fat, milligrams of sodium, grams of fiber,

[1]Source: Dunhumby *Breakfast at the Frat* `https://www.dunnhumby.com/source-files/`

[2]Source: `https://dasl.datadescription.com/datafile/cereals/`

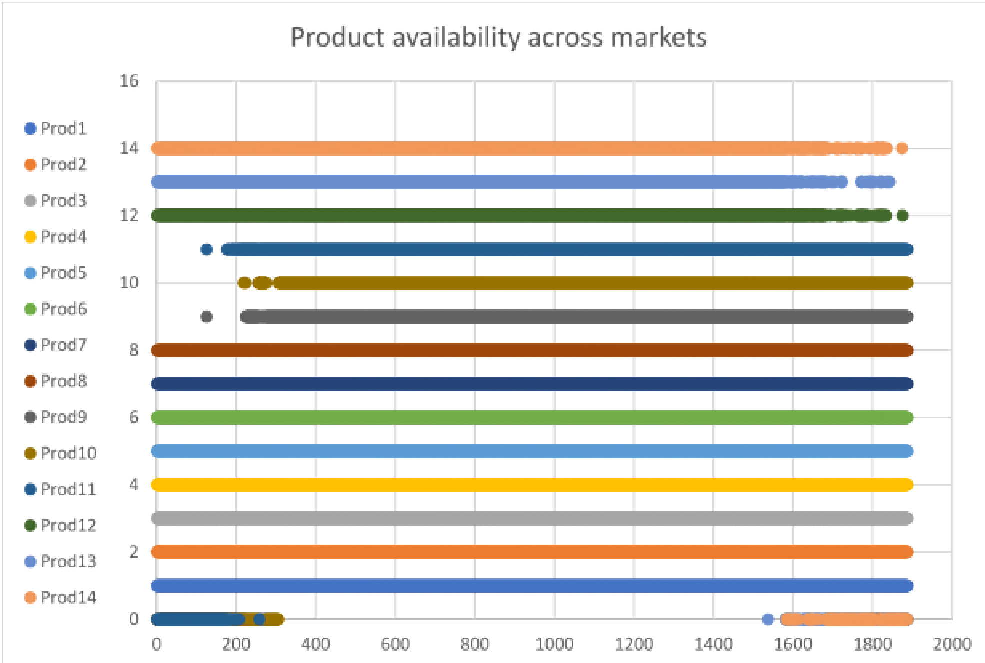

Figure 2: Product availability across markets. Products 9-11 enter late and products 12-14 exit early

and milligrams of potassium are all identified for each product. These characteristics are not observed to change across markets

Following Nevo (2001), demographic variables such as average family income, age, and number of children ages 3-5, 6-13, and 14-17 were estimated for each month and Metropolitan Statistical Area through draws from the Basic Monthly Current Population Survey[3]. MSA codes were matched to counties in the original dataset through the US Cities Database[4]. City population for each month and city was estimated from the Population and Housing Unit Estimates Datasets[5]. The relevant market size for each location-week is estimated as city population multiplied by seven bowls of RTE cereal per week

On the Supply side, the prices of sugar, paperboard (index), wages, corn, rice, wheat, and oats (index) are taken from the Federal Reserve Economic Database[6]

[3]Source: https://www.census.gov/data/datasets/time-series/demo/cps/cps-basic.html

[4]Source: https://simplemaps.com/data/us-cities

[5]Source: https://www.census.gov/programs-surveys/popest/data/data-sets.html

[6]Source: https://fred.stlouisfed.org/ series PSUGAUSAUSDM, WPU09141105, CES3000000003, PMAIZMTUSDM, PRICENPQUSDM, PWHEAMTUSDM, and WPU012203

4 Methodology

To evaluate alternative divestitures, I employ a nested logit model of Ready-to-Eat (RTE) cereal following Nevo (2001) and Miller and Weinberg (2017). The Supply side of the model is estimated following Miller and Weinberg (2017). A counterfactual merger is simulated between General Mills and Kellogg, followed by divestitures of candidate packages to the remaining incumbent firms as well as a "carve-out" entrant firm. Efficiency gains are calculated as marginal cost reductions according to the Data Envelopment Analysis approach of Bonnet and Schain (2020). Finally, welfare is estimated and compared for each outcome.

4.1 Demand

In standard random coefficients nested logit models of demand (e.g. Nevo (2001), Miller and Weinberg (2017)), consumer i's indirect utility from consuming one serving of brand $j = 1, ..., 14$ RTE cereal in market location $r = 1, ..., 51$ at time $t = 1, ..., 37$ is given by

$$u_{ijrt} = x_j'\beta_i - \alpha_i p_{jrt} + \tau_j + \tau_t + (\xi_{jrt} + \bar{\epsilon}_{ijrt}) \tag{1}$$

where x_j is a vector of observable characteristics. The absence of rt subscripts reflects that there is no observable change in these characteristics across time or markets. The retail store price of the cereal is p_{jrt}, time- and product- fixed effects are given by τ_j and τ_t, and ξ_{jrt} consists of quality characteristics unobserved to the econometrician. The random coefficients α_i, β_i are often parameterized by

$$\begin{bmatrix} \alpha_i \\ \beta_i \end{bmatrix} = \begin{bmatrix} \alpha \\ \beta \end{bmatrix} + \Pi D_i + \Sigma v_i$$

where D_i includes demographic variables, $v_i \sim N(0, I)$ represents unobserved demographic characteristics, and Σ is a scaling matrix.

Different assumptions on the distribution of $\bar{\epsilon}_{ijrt}$ result in different formulas for the predicted market shares. Following the standard nested logit model à la Miller and Weinberg (2017), I assume that $\bar{\epsilon}_{ijrt}$ is iid Type 1 Extreme Value (T1EV). Let two groups $g = 0, 1$ be defined such that group 0 has the outside good and group 1 has the inside goods. Then

$$\bar{\epsilon}_{ijrt} = \zeta_{igrt} + (1 - \sigma)\epsilon_{ijrt}$$

where ϵ_{ijrt} is T1EV, ζ_{igrt} has the unique distribution that makes $\bar{\epsilon}_{ijrt}$ T1EV, and σ is the canonical nesting parameter. Values of σ closer to 1 indicate stronger substitution within groups, whereas values closer to 0 indicate stronger substitution across groups.

In general, the predicted share of good j in market rt is given by

$$\widehat{s}_{jrt} = \int_{A_{jrt}} dP^*(v_i) dF(\bar{\epsilon})$$

where A_{jrt} is the set of consumer attributes that would result in buying product j in market rt; and P^* and F are population distribution functions for v_i and ϵ, respectively. General elasticities (and by extension, share derivatives) are given by

$$\frac{p_{krt}}{s_{jrt}}\frac{\partial s_{jrt}}{\partial p_{krt}} = \begin{cases} -\frac{p_{krt}}{s_{jrt}} \int \alpha_i s_{ijrt} * (1 - s_{ijrt}) dP^*(v_i) dF(\epsilon) & j = k \\ \frac{p_{krt}}{s_{jrt}} \int \alpha_i s_{ijrt} * s_{ikrt} dP^*(v_i) dF(\epsilon) & j \neq k \end{cases}$$

The primary advantage of the full random coefficients nested logit model is its allowance for very flexible substitution patterns. Standard logit models of Demand exhibit the restrictive Independence of Irrelevant Alternatives property, by which a rise in the price of a kids' cereal (like Cap n' Crunch) would result in equal substitution to both another kids' cereal and to an adult cereal (like Raisin Bran). The full RCNL model is substantially more realistic; however it is also substantially more computationally expensive in practice. To reduce the computational burden while still maintaining reasonable substitution patterns, I choose to drop the random coefficients (i.e. $\Pi, \Sigma = 0$) while retaining the nesting structure.

Let the mean utility be defined as $\delta_{jrt} = x_j'\beta - \alpha p_{jrt} + \tau_j + \tau_t + \xi_{jrt}$. Normalizing the mean utility of the outside good to zero, we have $u_{i0rt} = \epsilon_{i0rt}$. Berry (1994) provides a simple analytic share inversion for the nested logit model given by $\delta_{jrt} = \ln(S_{jrt}) - \ln(S_{0rt}) - \sigma \ln(\overline{S}_{jrt|g})$; where S_{jrt} is the observed market share of product j (i.e. $S_{jrt} = q_{jrt}/M_{rt}$), S_{0rt} is the observed share of the outside good, and $\overline{S}_{jrt|g}$ is the share of product j as a fraction of the group share, i.e. $\overline{S}_{jrt|g} = s_{jrt}/\sum_{j\in g} s_{jrt}$. This leads to a simple regression of

$$\ln(S_{jrt}) - \ln(S_{0rt}) = x_j'\beta - \alpha p_{jrt} + \tau_j + \tau_t + \sigma \ln(\overline{S}_{jrt|g}) + \xi_{jrt} \tag{2}$$

For the nested logit model, the predicted market share of good j in market rt is given by Berry (1994) as

$$\widehat{s}_{jrt}(\delta_{jrt}, \sigma) = \frac{e^{\delta_{jrt}/(1-\sigma)}}{D_{grt}^{\sigma}\left[\sum_g D_{grt}^{(1-\sigma)}\right]} \tag{3}$$

where $D_{grt} = \sum_{j\in g} e^{\delta_{jrt}/(1-\sigma)}$. Share derivatives are given by Ackerberg and Crawford (2009) as

$$\frac{\partial S_{jrt}}{\partial p_{krt}} = \begin{cases} -\alpha S_{jrt} * \left(\frac{1}{1-\sigma} - \frac{\sigma}{1-\sigma} S_{jrt|g} - S_{jrt}\right) & j = k \\ \begin{cases} \alpha S_{krt}\left(\frac{\sigma}{1-\sigma} S_{jrt|g} + S_{jrt}\right) & j,k \text{ in same group (i.e. inside goods)} \\ \alpha S_{jrt} * S_{krt} & \text{otherwise (i.e. } j,k = 0) \end{cases} \end{cases} \tag{4}$$

4.2 Demand Estimation

Estimation of the Demand system proceeds by 2SLS using the optimal weighting matrix. The two sources of endogeneity in the model include prices and the within-group share. Prices are likely to be endogenous since the mean valuation of quality ξ_{jrt} is known to firms when making their pricing decisions, but is not identified separately from the error term. The within-group share $\overline{S}_{jrt|g}$ is a function of the LHS of the regression, and is therefore endogenous by construction. One advantage of the nested logit model is that even if characteristics x_j are endogenous, this bias will not be transmitted in estimation to the price coefficient α or the share coefficient σ; allowing for accurate estimation of share derivatives (Ackerberg and Crawford (2009)). As estimation of the Demand system serves primarily to identify the share derivatives for the Supply side, including at least one instrument for price and one for the within-group share are sufficient to accurately estimate the model.

Following guidance from BLP (1995), Nevo (2001), Villas-Boas (2007), and Friberg and Romahn (2015), price instruments include brand dummies, quarterly dummies (to control for unobserved costs over time), and state/regional dummies (to account for unobserved transportation costs). I also include three observed cost-shifter dummies: whether the product was part of an in-store promotional display, whether the product was part of the in-store circular, and whether the product had a temporary shelf-tag price reduction. Nevo (2001) and Friberg and Romahn (2015) both caution at length against using product fixed effects as regressors. Full inclusion of all product fixed effects would result in a singular instrument matrix, and inclusion of most product fixed effects substantially weakens the other price instruments. To minimize this issue, I follow Friberg and Romahn (2015) by only including fixed effects for products which appear in every market (i.e. products 1-8).

In most applications of Demand estimation, product characteristics are assumed to be exogenously determined by firms in advance. This allows for the construction of "BLP instruments" (i.e. functions of the other firms' characteristics), which are likely to be strongly correlated with intensely rival products. Although exogeneity of characteristics is more difficult to justify when entry and exit of products are observed (as they are here), I follow Friberg and Romahn (2015) by computing average characteristics for each market. Results from this exercise are presented in Appendix A. Since there is minimal variation in average characteristics across markets, I argue that product characteristics are plausibly exogenous in this data set. Two sets of BLP instruments are constructed using the sums and averages of other firms' characteristics.

In the market for RTE cereal, advertising expenses are a substantial piece of firm's expenses and marginal costs (see Nevo (2001)). Although advertising costs and expenses are not identified in my data set, ignoring them would likely lead to another source of price endogeneity through an omitted variables bias. Following Bonnet and Schain (2020), estimating the reduced-form equation of prices on input costs W_{jrt} given by

$$p_{jrt} = W'_{jrt}\gamma + \tau_j + \tau_m + \eta_{jrt}$$

(where τ_m is a set of manufacturer dummies) produces a residual estimate of $\widehat{\eta}_{jrt}$ that non-parametrically captures the effect of omitted variables (such as advertising) in ξ_{jrt}. This term is included as an instrument for price in the regression.

A particular concern of choosing the nested logit model is that estimation results may be driven by the choice of nests. To mitigate against this, I re-estimate the model with $g = 0, 1, 2, 3$ groups corresponding to the outside good and All-family, Kids, and Adult categories identified in the data. Like Miller and Weinberg (2017), I use the number of products within each group for each market as an instrument for the nested logit parameter. This instrument is valid as long as the structural error term is uncorrelated with the number of products in each market.

4.3 Results of Demand Estimation

The results from the Demand estimation are presented below. The first specification uses two groups corresponding to the inside goods and the outside good. The second specification uses 3 groups for the inside goods as described above.

Variable	Parameter	One-nest	Category-nest
Price	$-\alpha$	-0.778	-0.085
		(0.010)	(0.0115)
Nesting parameter	σ	0.975	0.019
		(0.010)	(0.0103)
Intercept	β_0	-21.347	-11.798
		(0.073)	(0.0722)
Calories	β_1	0.243	0.001
		(0.001)	(0.001)
Sugars	β_2	-0.676	0.039
		(0.004)	(0.005)
Carbo	β_3	-0.504	0.009
		(0.002)	(0.002)
Protein	β_4	-0.353	0.096
		(0.009)	(0.010)
Fat	β_5	-1.713	-0.311
		(0.014)	(0.015)
Sodium	β_6	-0.012	0.005
		(0.0001)	(0.0001)
Fiber	β_7	-2.474	0.575
		(0.0186)	(0.020)
Potassium	β_8	0.048	-0.013
		(0.0004)	(0.0004)
Adjusted R-squared		0.09167	0.09592

Figure 3: Results from the Demand Estimation. Standard errors are given in parentheses

The category-nest structure does not appear to offer strong improvements over the inside-good nest structure. Not only are the parameter estimates substantially attenuated in the category-nest structure, but the estimates of σ in both specifications are consistent with robust substitution across groups. Smaller estimates of α, σ also introduce computational issues in calculation of the share derivatives. Therefore, I choose to focus on the inside-good nest estimates for the remainder of the analysis. Each of the characteristics are significant at the 1% level, and the magnitude and signs are consistent with expectations. Consumers like calories, and dislike fat and fiber. Estimated average cross-price elasticities of Demand across all markets are presented below in Figure 4

4.4 Supply

Following Nevo (2001) and Miller and Weinberg (2017), each manufacturing firm f produces a set of goods $\mathcal{F}_f$ to maximize their profits given by

$$\Pi_f(p) = \sum_{j \in \mathcal{F}_f} (p_{jrt} - mc_{jrt}) M s_{jrt}(p) + \kappa \sum_{j \notin \mathcal{F}_f} (p_{jrt} - mc_{jrt}) M s_{jrt}(p)$$

where $\kappa \in \{0, 1\}$ is a conduct parameter that captures the internalization of coalition pricing externalities as defined by Miller and Weinberg (2017). This leads to the familiar Bertrand first order conditions:

$$s_{jrt}(p) + \sum_{k \in \mathcal{F}_f} (p_{krt} - mc_{krt}) \frac{\partial s_{jrt}(p)}{\partial p_{krt}} + \kappa \sum_{k \notin \mathcal{F}_f} (p_{krt} - mc_{krt}) \frac{\partial s_{jrt}(p)}{\partial p_{jrt}} = 0$$

which can be vectorized across products for each market as

$$s_{rt}(p_{rt}) + \left[\Omega(\kappa) \circ \left(\frac{\partial s_{rt}}{\partial p_{rt}} \right)' \right] (p_{rt} - mc_{rt}) = 0$$

$$\Rightarrow p_{rt} = mc_{rt} - \left[\Omega(\kappa) \circ \left(\frac{\partial s_{rt}}{\partial p_{rt}} \right)' \right]^{-1} s_{rt}(p_{rt}) \quad (5)$$

where $\circ$ denotes the Hadamard (element-wise) product and $\Omega(\kappa)$ is the ownership matrix (i.e. $\Omega_{ij} = 1$ if products i and j are produced by the same firm). Estimation of alternative mergers (and divestitures) is simply a matter of changing the elements in $\Omega(\kappa)$. Letting $\Omega_0(\kappa)$ and $\Omega_1(\kappa)$ denote the pre- and post-merger ownership matrices, pre-merger marginal costs can be estimated as

$$\widehat{mc}_{rt} = p_{rt} + \left[\Omega_0(\kappa) \circ \left(\frac{\partial s_{rt}}{\partial p_{rt}} \right)' \right]^{-1} s_{rt}(p_{rt})$$

Product	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)
(1)	-112.24	2.94	3	0	0	2.69	0.76	0	0	0	0	0	0	0.9
(2)	2.87	-87	3.34	0	0	3	0.84	0	0	0	0	0	0	1.01
(3)	2.31	2.63	-100.7	0	0	2.41	0.68	0	0	0	0	0	0	0.81
(4)	0	0	0	-93.47	3.17	0	0	0	0	1.46	0	1.56	1.72	0
(5)	0	0	0	2.17	-91.11	0	0	0	0	1.45	0	1.56	1.72	0
(6)	1.86	2.13	2.17	0	0	-67.74	0.55	0	0	0	0	0	0	0.65
(7)	2.17	2.48	2.53	0	0	2.28	-56.79	0	0	0	0	0	0	0.76
(8)	0	0	0	0	0	0	0	-55.68	1.64	0	2.07	0	0	0
(9)	0	0	0	0	0	0	0	2.56	-102.93	0	3.22	0	0	0
(10)	0	0	0	1.77	2.57	0	0	0	0	-82.81	0	1.27	1.4	0
(11)	0	0	0	0	0	0	0	2	1.98	0	-101.99	0	0	0
(12)	0	0	0	1.52	2.22	0	0	0	0	1.02	0	-79.43	1.21	0
(13)	0	0	0	1.86	2.72	0	0	0	0	1.25	0	1.34	-79.29	0
(14)	2.29	2.62	2.67	0	0	2.4	0.67	0	0	0	0	0	0	-88.21

Figure 4: Estimated average Cross-Price Demand elasticities across all markets. Group 1 (“All-family”) is highlighted in yellow and group 2 (“Kids”) in cyan. The cell at row j and column k corresponds to the percentage change in the share of product j for a 1% increase in the price of product k

4.5 Divestitures

I estimate a counterfactual acquisition of General Mills by Kellogg, the two largest observed producers in my data set. In order to ensure availability of all products for accurate computation of share derivatives, I choose August 2010 (2010 month 8) as the merger date, which corresponds to markets 970 through 1020 in the data set. Drawing insights from Jayaratne and Shapiro (2001) as discussed in section 2, I only consider divestitures of products in the Kids category for the merged firm. This leads to a total of 5 candidate packages for partial divestiture: {1}, {2}, {3}, {1,3}, and {2,3}. The package {1,2} corresponds to a full divestiture of the Cheerios© line from the merged firm, so it is used as a baseline for comparison. Counterfactual prices are re-estimated as each candidate package is assigned to firms 3-5 (Private Label, Post, and Quaker), as well as "firm 0", which is a carve-out firm that only holds the divestiture package

4.6 Efficiency Gains

Following Bonnet and Schain (2020), I estimate reductions in merging firms' marginal costs with Data Envelopment Analysis. DEA is a non-parametric approach which estimates a production frontier from a linear combination of a firm's inputs and outputs. Pooling the inputs and outputs of the merged firm (and subtracting divested assets), DEA then projects the merged firm onto the frontier. In the radial input-oriented model, this projection estimates potential reductions in the merged firms' combined inputs that could obtain the same level of output (Álvarez, Barbero, and Zofío, (2020)), which is analogous to cost savings (Bonnet and Schain, (2020)). The shape of the frontier depends on the assumed underlying production technology. Since the RTE cereal industry is fairly well-established and shows no entry or exit around the merger date, I assume that this underlying technology exhibits Constant Returns to Scale.

Formally, let s_{fc} be the market share of firm f in producing category $c \in \{1, 2, 3\}$ (All-family, Kids, and Adult) of RTE cereal. Let $y_{fc} = Ms_{fc}$ denote the total output of firm f in category c, where M is the total market size at the merger date. To estimate firm's inputs, marginal costs are regressed on input prices W_{kct} and product and manufacturer fixed effects

$$\widehat{mc}_{jrt} = \tau_j + \tau_m + W_{kct}\rho + \varepsilon_{jrt}$$

The estimated amount of input k for firm f to produce category c is given by $x_{fck} = y_{fc} * \widehat{\rho}_{fck}$. Let U denote the set of merging firms. The radial input-oriented "potential overall gains" from merging (E^U) are given by the linear program

$$E^U = \min_{E,\lambda} E \quad \text{st}$$

$$\begin{aligned} & E * \left[\sum_{f\in U} x^f\right] \geq \sum_{f\in F} \lambda_f x^f \\ & \sum_{f\in U} y^f \leq \sum_{f\in F} \lambda_f y^f \\ & \lambda \geq 0 \end{aligned} \tag{6}$$

To illustrate, suppose that firms A and B produce a scalar output y with the use of two inputs x_1, x_2. Furthermore, let $F_f(x_1^f, x_2^f)$ denote the production function of firm f using inputs x_1, x_2. Suppose that $F_A(3,9) = 15$ and $F_B(9,3) = 15$, and the efficient frontier has $(x_1, x_2) = (4,4)$. Calculation of E^U pools together the merging firm's inputs, so $(x_1^U, x_2^U) = (3+9, 9+3) = (12,12)$. The merged firm can therefore reduce both inputs by 2/3 to achieve the efficient frontier, hence $E^U = 1/3$ of the merged firms' inputs are needed to achieve the frontier

Accounting for efficiency gains, adjusted marginal costs are given by

$$\widehat{mc}^*_{jrt} = \begin{cases} \widehat{mc}_{jrt} - \left[(1-E^U) * W_{kct}\widehat{\rho}\right] & \text{for merging firms} \\ \widehat{mc}_{jrt} & \text{all other firms} \end{cases}$$

and the simulated post-merger prices p^*_{rt} can be obtained as the solution to

$$p^*_{rt} = \widehat{mc}^*_{rt} - \left[\Omega_1(\kappa) \circ \left(\frac{\partial s_{rt}}{\partial p_{rt}}\right)'\right]^{-1} s_{rt}(p^*_{rt}) \tag{7}$$

5 Estimation Results and Inference

The estimation routine[7] loops over 25 candidate divestiture packages (including the standard merger simulation). On each iteration, the ownership matrix, product shares, share derivatives, and DEA marginal cost reductions are numerically estimated. Counterfactual prices are estimated using MATLAB's `fsolve` command. As average prices of each product range between \$1-\$4 across markets (see Figure 5 below), I use a uniformly-generated random vector with a mean of 2.5 as a starting point. With a maximum number of iterations at 1500 and a tolerance of 1e-12, each of the estimates converged with a function value less than 1e-13 at the optimum.

[7] Code for the estimation can be found at `https://github.com/torculus/Antitrust-Divestitures`

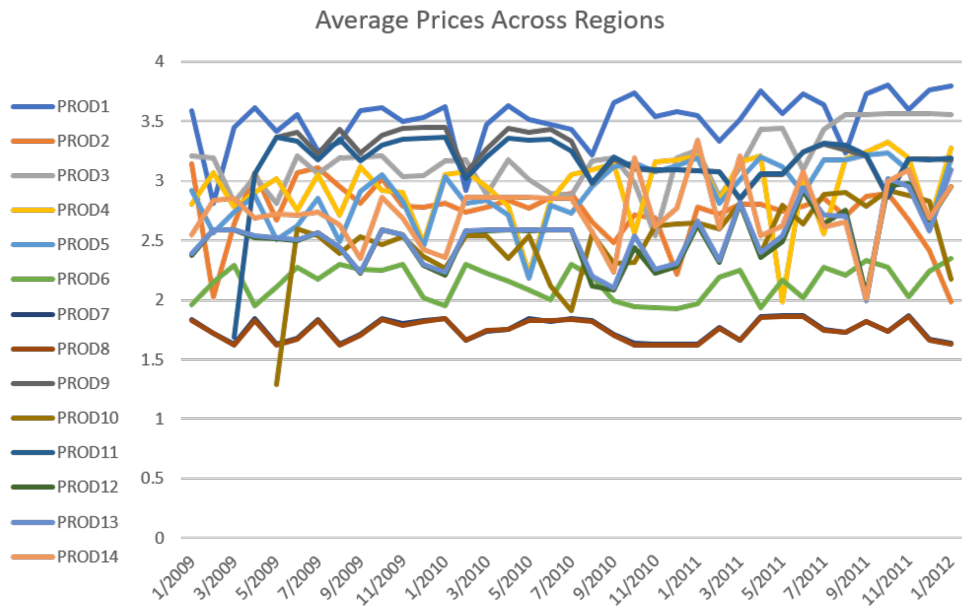

Figure 5: Average prices of products 1-14 across all 51 locations for each month

In practice, the estimated "potential overall gains" (E^U) for each loop are equal to 1, implying that there are no efficiency gains in the merging parties. To validate this, I computed input and output slacks on each iteration, which were zero to machine precision. To counteract potential limitations in my data set or choice of merger date, in what follows I also present results for an assumed 5% reduction in marginal costs, which is standard in other merger analyses (Bonnet and Schain, (2020))

Average percentage changes in counterfactual prices are presented below in Figure 6. Specifications 2-6 correspond to the six candidate packages discussed above, while the package {1,2} corresponds to the pre-merger outcome and is used as a baseline. Specification 1 corresponds to the pure merger simulation in which no products are divested. Without divestitures, prices are estimated to be 2.965% higher without efficiency gains (and 0.478% higher assuming 5% efficiency gains), with a standard deviation of 4.486% (4.119%, respectively). This reflects the unilateral effects of the merger

Partial divestitures to incumbents tend to result in lower price increases than full divestitures, and this effect is more pronounced with efficiency gains. I interpret this as an artifact of my data set. Since products 1 and 2 (the Cheerios© line) are the most-represented products in my data set, it may be that a full divestiture transfers too much market power to incumbent firms, resulting in larger price increases than partial divestitures. Furthermore, full divestitures to the "carve-out" firm generally result in lower price increases relative to divestitures to incumbents. This finding validates Vasconcelos' (2010) argument that the carve-out firm cannot readily incorporate the market power relative to incumbent firms

Variable	Baseline {1,2}	Spec 1 {}	Spec 2 {1}	Spec 3 {2}	Spec 4 {3}	Spec 5 {1,3}	Spec 6 {2,3}
Divestiture package: No efficiency gains							
Divesting to Incumbents:							
Average % Price change	3.837		0.451	3.663	0.497	0.474	2.975
SD of % price change	(4.897)		(0.88)	(5.332)	(0.895)	(0.794)	(4.287)
Divesting to "carve-out" firm 0							
Average % Price change	0.533		-0.057	0.585	-0.02	-0.038	0.557
SD of % price change	(0.073)		(0)	(0)	(0)	(0.027)	(0.04)
All products: No efficiency gains							
Divesting to Incumbents:							
Average % Price change	2.965	1.799	1.797	2.47	1.797	1.794	2.636
SD of % price change	(4.486)	(2.31)	(2.263)	(3.553)	(2.263)	(2.273)	(3.605)
Divesting to "carve-out" firm 0							
Average % Price change	1.76		1.76	1.721	1.76	1.721	1.76
SD of % price change	(2.338)		(2.338)	(2.365)	(2.338)	(2.365)	(2.338)
Divestiture package: assuming 5% efficiency gains							
Divesting to Incumbents:							
Average % Price change	1.135		-0.018	0.591	0.02	0.001	0.462
SD of % price change	(5.501)		(0.067)	(0.701)	(0.069)	(0.064)	(0.606)
Divesting to "carve-out" firm 0							
Average % Price change	0.457		-0.057	0.186	-0.02	-0.038	0.194
SD of % price change	(0.063)		(0)	(0)	(0)	(0.027)	(0.006)
All products: assuming 5% efficiency gains							
Divesting to Incumbents:							
Average % Price change	0.478	-0.214	0.029	-0.492	0.033	0.276	0.189
SD of % price change	(4.119)	(3.45)	(3.243)	(4.55)	(3.241)	(3.094)	(3.847)
Divesting to "carve-out" firm 0							
Average % Price change	0.241		0.026	-0.568	0.03	0.27	0.109
SD of % price change	(4.095)		(3.325)	(4.66)	(3.323)	(3.173)	(3.936)

Figure 6: Counterfactual Price Results. Each cell is in percentage terms. The first two row specifications ignore efficiency gains, while the last two assume a 5% reduction in the merging parties' marginal costs. The first and third row specifications measure only the prices of products in the divestiture package, while the second and fourth row measure all products

Finally, estimates of Consumer Surplus both with and without efficiency gains are calculated using $CS = \frac{1}{|\alpha|}\exp(\sum_j \delta_j)$. Results from this exercise are presented below in Figure 7. Without efficiency gains, 18 candidate divestitures create greater surplus than the pure merger; whereas with 5% efficiency gains, 4 packages (the partial divestitures of product 2) beat the merger. Divestitures to the private label (firm 5) generally result in the lowest surplus. The change in Consumer Surplus between partial and full divestitures ($\Delta CS_{partial} - \Delta CS_{full}$) was examined with a simple heteroskedastic difference-in-means t-test, and the result was statistically significant only when efficiency gains were included (P-val = 0.066). This validates Bougette's (2010) argument that higher efficiency gains increase the efficacy of partial divestitures. The difference in means was not statistically significant without efficiency gains, suggesting that partial divestitures perform just as well as full divestitures in this data set, contrary to current regulation guidelines

6 Conclusion

Although antitrust divestitures are a commonly-prescribed structural merger remedy, there is scant empirical justification for the preference of full divestitures over partial divestitures. This paper attempts to connect several independent literatures in economics by examining the welfare effects of partial asset divestitures relative to full divestitures when efficiency gains are present. By simulating a merger of two RTE cereal manufacturers, partial divestitures were not found to result in statistically different changes in Consumer Surplus relative to full divestitures without efficiency gains. However, partial divestitures were found to significantly increase surplus relative to full divestitures when efficiency gains were included. This result provides empirical justification for one of the main results in Bougette (2010). Furthermore, divestitures to incumbent firms were found to more adversely affect welfare relative to divestitures to the carve-out firm, justifying a main result in Vasconcelos (2010).

References

[1] Ackerberg, D. and Crawford, G. (2009) *Estimating Price Elasticities in Differentiated Product Demand Models with Endogenous Characteristics.* Working paper. `http://www.econ.ucla.edu/ackerber/acnew5.pdf`

[2] lvarez, Inmaculada C.; Barbero, Javier and Zofo, Jos L, (2020) *A Data Envelopment Analysis Toolbox for MATLAB.* Journal of Statistical Software. Volume 95, Issue 3, pp 1-49. `http://dx.doi.org/10.18637/jss.v095.i03`

[3] Berry, Steven T. (1994). *Estimating Discrete-Choice Models of Product Differentiation.* The RAND Journal of Economics, vol. 25, no. 2, 1994, pp. 242262. JSTOR, `www.jstor.org/stable/2555829`

[4] Berry, Levinsohn, and Pakes. (1995). *Automobile Prices in Market Equilibrium.* Econometrica, Vol. 63, No. 4. (Jul., 1995), pp. 841-890.

Iteration	Consumer Surplus (No efficiencies)	Consumer Surplus (5% efficiencies)	Best to worst (No efficiencies)	Best to worst (5% efficiencies)
1	3.768	3.943	133	20
13	3.772	3.916	233	23
14	3.772	3.916	123	24
15	3.76	3.915	134	25
23	3.775	3.997	20	1
24	3.774	3.996	130	233
25	3.633	3.979	23	234
33	3.771	3.918	234	230
34	3.771	3.918	124	30
35	3.761	3.917	24	34
123	3.775	3.913	230	33
124	3.774	3.911	10	35
125	3.525	3.851	13	10
133	3.775	3.891	14	14
134	3.775	3.891	120	13
135	3.754	3.89	30	15
233	3.775	3.925	34	123
234	3.774	3.924	33	124
235	3.589	3.901	1	120
10	3.772	3.916	35	235
20	3.775	3.997	15	134
30	3.771	3.918	135	130
120	3.771	3.909	25	133
130	3.775	3.891	235	135
230	3.772	3.923	125	125

Figure 7: Consumer Surplus estimates. Iteration 1 corresponds to the pure merger simulation without divestiture, while iteration *abc* indicates products *a* and *b* are divested to firm *c*. For instance, iteration "230" indicates that package {2,3} is divested to firm 0 (the carve-out). Partial divestitures are highlighted in yellow

[5] Bogetoft, P., and Wang, D. (2005). *Estimating the Potential Gains from Mergers.* Journal of Production Analysis vol. 23, 145171 (2005). `https://doi.org/10.1007/s11123-005-1326-7`

[6] Bonnet and Schain (2020). *An Empirical Analysis of Mergers: Efficiency Gains and Impact on Consumer Prices* , Journal of Competition Law & Economics, Volume 16, Issue 1, March 2020, Pages 135, `https://doi.org/10.1093/joclec/nhaa001`

[7] Bougette, P. (2010). *Preventing merger unilateral effects: A NashCournot approach to asset divestitures.* Research in Economics, Volume 64, Issue 3, 2010, pp 162-174,`https:`

//doi.org/10.1016/j.rie.2010.03.001.

[8] Cosnita, A. and Tropeano, J.P. (2009). *Negotiating remedies: Revealing the merger efficiency gains*, International Journal of Industrial Organization, Volume 27, Issue 2, 2009, pp 188-196,https://doi.org/10.1016/j.ijindorg.2008.06.003

[9] Department of Justice and Federal Trade Commission (2010). *Horizontal Merger Guidelines*. Washington, D.C.: U.S. Dept. of Justice, 2010

[10] Elzinga, Kenneth G. (1969). *The Antimerger Law: Pyrrhic Victories?* The Journal of Law & Economics, vol. 12, no. 1, 1969, pp. 4378. JSTOR, www.jstor.org/stable/724979

[11] Federal Trade Commission (1999). *A study of the Commission's Divestiture Process*. Staff report, Federal Trade Commission

[12] Federal Trade Commission (2017). *The FTC's Merger Remedies 2006-2012*. Staff report, Federal Trade Commission

[13] Friberg, R. and Romahn, A. (2015). *Divestiture requirements as a tool for competition policy: A case from the Swedish beer market*. International Journal of Industrial Organization, vol 42 2015

[14] Gayle, W., Marshall, R.C., Marx, L.M. et al. (2011). *Coordinated Effects in the 2010 Horizontal Merger Guidelines*. Review of Industrial Organization vol 39, 3956 (2011). https://doi.org/10.1007/s11151-011-9303-y

[15] Jayaratne, J. and Shapiro, C. (2001). *Simulating Partial Asset Divestitures to 'Fix' Mergers*, International Journal of the Economics of Business, 7:2, 179-200, DOI: 10.1080/13571510050084523

[16] Knittel, Christopher R., and Konstantinos Metaxoglou. (2011). *Challenges in Merger Simulation Analysis*. American Economic Review, 101 (3): 56-59. https://www.aeaweb.org/articles?id=10.1257/aer.101.3.56

[17] Miller, N. and Weinberg, M. (2017). *Understanding the Price Effects of the MillerCoors Joint Venture*. Econometrica, Vol. 85, No. 6, (November 2017), 17631791

[18] Prentice, D. and Pham, V. (2013). *A Random Coefficients Logit Analysis of the Counterfactual: A Merger and Divestiture in the Australian Cigarette Industry*, SSRN Electronic Journal. 10.2139/ssrn.2287327.

[19] Nevo, A. (2000) *A Practitioner's Guide to Estimation of Random Coefficients Logit Models of Demand*, Journal of Economics & Management Strategy, Wiley Blackwell, vol. 9(4), pages 513-548, December

[20] Nevo, A. (2001). *Measuring Market Power in the ReadytoEat Cereal Industry*. Econometrica, 69: 307-342. https://doi.org/10.1111/1468-0262.00194

[21] Tenn, S. and Yun, JM. (2011). *The Success of Divestitures in Merger Enforcement: Evidence from the J&J - Pfizer Transaction.* International Journal of Industrial Organization 2011;29; 273-282.

[22] Sofia Berto Villas-Boas (2007). *Vertical Relationships between Manufacturers and Retailers: Inference with Limited Data*. The Review of Economic Studies, vol. 74, no. 2, 2007, pp. 625652. JSTOR, www.jstor.org/stable/4626153

[23] Vasconcelos, Helder. (2010). *Efficiency Gains And Structural Remedies In Merger Control*. The Journal of Industrial Economics, vol. 58, no. 4, 2010, pp. 742766. JSTOR, www.jstor.org/stable/40985719

[24] Weinberg, M. (2011). *More Evidence on the Performance of Merger Simulations*. The American Economic Review, Vol. 101, No. 3, (MAY 2011), pp. 51-55

A Average Characteristics across Markets

For each market, average characteristics were calculated as the average of characteristics in the available products. Average characteristics exhibit remarkable stability across markets: the range of average calories across markets is just 10.28% of the total range of calories across products. Choice sets become decreasingly fat- and sodium-rich later in the data set due to the entry of more Adult cereals

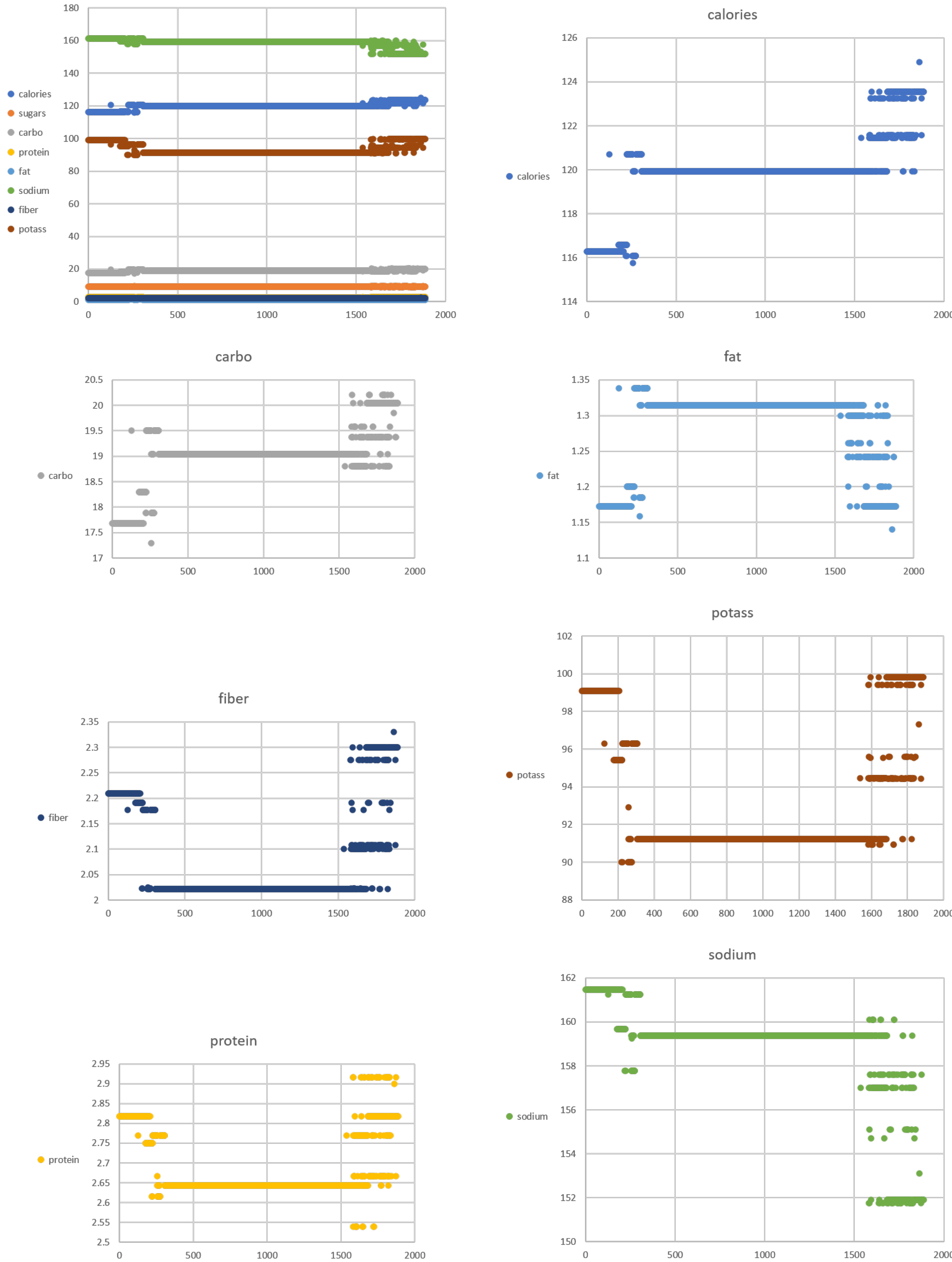

Figure 8: Average Characteristics across markets 1-1887. Average characteristics exhibit a range of 9.15 calories (10.28% of the range across products), 0.67 grams of sugars(5.56%), 2.92 grams of carbohydrates(8.58%), 0.378 grams protein(7.56%), 0.198 grams of fat(6.62%), 9.705 milligrams of sodium(3.35%), 0.309 grams of fiber(5.32%), and 9.818 milligrams of potassium(4.57%)