Title
Identification of demand in differentiated products markets

Author
Megerdichian, Aren

Publication Date
2010

UNIVERSITY OF CALIFORNIA, SAN DIEGO

Identification of Demand in Differentiated Products Markets

A dissertation submitted in partial satisfaction of the
requirements for the degree
Doctor of Philosophy

in

Economics

by

Aren Megerdichian

Committee in charge:

Professor Halbert White, Chair
Professor Richard Carson
Professor Karsten Hansen
Professor Dominique Lauga
Professor Michael Noel

2010

The dissertation of Aren Megerdichian is approved, and it is acceptable in quality and form for publication on microfilm and electronically:

____

____

____

____

____
Chair

University of California, San Diego

2010

TABLE OF CONTENTS

LIST OF FIGURES

ACKNOWLEDGEMENTS

First and foremost, I would like to thank Professor Halbert White, from whom it was an honor and privilege to learn econometrics. I am forever indebted to him for generously providing me with his time and guidance during the past three years. Without his invaluable insights and input, this dissertation would not be possible. I hope that I have done justice to his important work on identification and causal inference in structural economic systems. I am also indebted to Professors Richard Carson, Michael Noel, Karsten Hansen, and Dominique Lauga, who provided helpful comments and suggestions that benefitted this dissertation in countless ways. Finally, I am extremely grateful to Amjad Malik for providing me with the scanner data that I used in my dissertation.

I would like to acknowledge Sam Dastrup, Hiroaki Kaido, Daniel Lima, Xun Lu, and Danielken Molina, all brilliant young scholars who entered the Ph.D. program with me in 2005. They provided me with helpful feedback at various stages of my dissertation. I would like to especially acknowledge Xun Lu's exceptional work on our joint paper on identification in nonseparable logit models, which can be found in the third chapter of my dissertation.

Last but not least, I owe special thanks to my wife, Bree, my parents, Mac and Sylvia, and my sister, Teena. Although I still don't think they quite know what I've been doing at UCSD for the past five years, they never ceased to support me and my endeavors.

This dissertation is dedicated to Nora and Ella.

VITA

2000	Bachelor of Science - Economics, Quantitative California State Polytechnic University, Pomona
2005	Master of Arts - Economics University of California, Santa Barbara
2010	Doctor of Philosophy - Economics University of California, San Diego

ABSTRACT OF THE DISSERTATION

Identification of Demand in Differentiated Products Markets

by

Aren Megerdichian

Doctor of Philosophy in Economics

University of California, San Diego, 2010

Professor Halbert White, Chair

This dissertation contains four essays at the intersection of econometrics and industrial organization. In all my chapters, I rely on a detailed set of supermarket scanner data on ready-to-eat cereals. In Chapter 1, I examine identification of price effects for differentiated product markets by relying on a conditional form of exogeneity that is an alternative framework to standard instrumental variables. I simulate price changes in the cereal industry arising from potential mergers between firms, one of which took place in 2008.

In Chapter 2, I continue to employ conditional exogeneity to identify the effect of market price on demand for differentiated products. The analysis here departs from past studies of demand in several ways, including relaxing the preva-

lent assumption that observed product characteristics are exogenous. Estimates of implied price-cost margins based on the conditional exogeneity framework are far more reasonable and stable compared to estimates based on standard instrumental variables procedures.

In Chapter 3, we (coauthored with Xun Lu) relax the omnipresent assumption that indirect utility takes a linear-separable parametric form in standard logit models of demand. We rely on conditional independence to structurally identify and nonparametrically estimate the average marginal effect of market price on consumer demand. We find that the effect of price on demand is monotonically increasing in price, resulting in high-priced goods having less elastic own price elasticities, and thus higher implied price-cost margins, which addresses a well-known concern in empirical industrial organization

In Chapter 4, I examine a firm's decision to raise price overtly (by increasing the dollar amount of a good) versus a hidden price change (by decreasing the contents in a good's package). I conduct a comprehensive set of empirical analyses in order to assess the impact of hidden price increases on expenditure share and profitability. During July 2007, General Mills decreased the cereal content for 20 out of 23 of their products in my sample of scanner data. A key finding is that consumers tend to notice hidden price changes on smaller-sized boxes of cereal, leading them to substitute to larger-sized boxes of cereal.

Chapter 1

Identification of Price Effects in Demand Systems with an Application to Merger Simulation

1.1 Introduction

Researchers often find the need to estimate demand systems. For example, demand systems are of central importance for analyzing product market delineation,[1] as well as determining competitive effects from mergers and new product introductions.[2] Endogeneity of the right-hand-side variables in the estimating demand equations, typically prices, is often a concern for practitioners. When price exogeneity fails, the researcher is not able to identify the causal or structural effect of price on quantity in the demand equation, but rather is only able to estimate predictive effects. Failure of demand curve identification has a rich history in economics, and is often addressed with instrumental variable (IV) estimators using cost shifters, or proxies for cost shifters, as exogenous instruments. This chapter, in contrast, relies on a conditional form of exogeneity for identification that does not require the instruments to be exogenous.

[1]See, for example, Rubinfeld (2000), Werden (1998), and Scheffman and Spiller (1996).

[2]See Nevo (2000b), footnote 1, for a list of numerous important studies in industrial organization.

Price endogeneity may arise for several reasons, and understanding the underlying sources is important for remedying the problem. The price-setting behavior in differentiated product markets is different than the supply-demand framework that researchers may expect to encounter in homogeneous product markets under perfect competition. The latter is the textbook example of why demand identification may fail to hold, while the former is the focus of this chapter. Research in retail pricing suggests that prices set for many products sold in supermarkets are not simply the result of the intersection between supply and demand, but rather are strategically set for a variety of competitive reasons. To the extent that manufacturers and retailers are setting prices based on factors that also affect demand, these confounding variables are potentially important sources of price endogeneity in the demand system. For example, competition among supermarkets within a geographic market for the same consumers is one potential source of price endogeneity; stockpiling behavior by consumers is another.

This chapter examines several issues of interest in econometrics, industrial organization (IO), management strategy, and quantitative marketing. First, I determine that simultaneity as a source of price endogeneity is relatively unfounded when estimating demand for differentiated products found in supermarkets, especially when the scanner data is disaggregated and high frequency. Second, instruments that have come to be widely known among IO practitioners as "Hausman instruments"[3] appear to suffer from instrument weakness when the data is high-frequency. Finally, I propose a system of structural equations describing consumer demand and retailer price-setting behavior based on previous research in IO and marketing, which sheds light on new sources of price endogeneity that can be conveniently addressed using Chalak and White's (2010) theoretical work on extended instrumental variables.

The empirical application undertaken here is estimating product-level demand for ready-to-eat cereals with the almost ideal demand system.[4] Demand

[3]Hausman (1997) proposes using prices in other cities to instrument for price in a particular city. This idea is evaluated in detail throughout this chapter.

[4]For an application of the conditional exogeneity framework to discrete choice models of demand, see Chapter 2.

estimation for cereal products has been the subject of econometric identification concerns in past research in IO; most notably, Hausman (1997) and Nevo (2001). A new, highly detailed set of supermarket scanner data that differs from the two aforementioned papers is employed here. It contains supermarket chain-level detail within each city at a weekly frequency, as well as stock-keeping unit detail for each of the cereal products in the sample. Information on prices, quantities, and promotional variables are available. I apply the conditional exogeneity approach of Chalak and White (2010) and the standard IV technique with Hausman instruments separately to estimate demand, and then to estimate price changes in the cereal industry resulting from simulated mergers, one of which took place in 2008. The conditional exogeneity assumption and conditioning instruments yield more reasonable results compared to standard instrumental variables estimation.

The organization of this chapter is as follows. Section 2 contains a discussion of price endogeneity as it relates to demand estimation for a general differentiated product. Section 2 also provides a structural system of equations that describes price-setting by firms/retailers and the associated demand response by consumers, which sheds light on the conditioning instruments that are needed to identify the effect of price on demand. Section 3 contains a discussion of standard identification assumptions, such as regressor and instrument exogeneity, as well as a discussion of the conditional exogeneity identification strategy. Section 4 presents demand estimation results for the ready-to-eat cereal industry using a detailed set of scanner data. Section 5 contains merger simulations employing demand estimates based on the different identification assumptions. Section 6 concludes.

1.2 Demand and Price Endogeneity

Consider estimating demand for a differentiated product found in supermarkets. Here, I focus on a product segment that is considered shelf-stable or nonperishable, such as cereal and soft drinks.[5] Typically, the effect of price on demand is of utmost interest to the researcher, while other variables, which may or may not be exogenous, can be captured by an unobservable term. A general demand function that results is given by

$$Q_{r,g,t} = d\left(P_{r,g,t}, U^{Q}_{r,g,t}\right), \tag{1.1}$$

where Q is the quantity purchased by consumers, P is a $1 \times J$ vector of observable prices for $j = 1, 2, ..., J$ goods in a relevant product segment, and U captures unobservable variables that drive demand. Throughout this chapter, the subscript r denotes the supermarket retail chain, g denotes geographic market (i.e., metro area), and t denotes the time period.[6] Price endogeneity reflects the dependence between U and P, represented by the notation $P \not\perp U$, where $\perp$ and $\not\perp$ denote independence and dependence, respectively.

1.2.1 Typical Sources of Price Endogeneity

A typical concern is that the prices and quantities observed in market data are jointly set by supply-demand equilibria, known as simultaneity. Simultaneity in the context of demand occurs if the variation in price and quantity observed in the data is due to variation in both supply and demand, in which case it becomes problematic to regress quantity against price to identify the demand curve.[7] While traditional supply-demand analysis for homogenous goods in a perfectly competitive market gives rise to identification problems due to simultaneity, it is not a reasonable concern in the estimation of demand for differentiated products in supermarkets.

Hausman (1997) points out that prices set in a given week may be considered predetermined under the assumption that supermarkets do not alter their

[5] A different set of considerations may need to be incorporated into the analysis for perishable goods, such as produce, dairy, and meat. Such goods are typically not prone to stockpiling by consumers.

[6] E.g., Ralphs-San Diego-Week of 5/1/2006.

[7] This is a classic example in econometrics textbooks as it relates to perfectly competitive markets. See, for example, chapter 10 of Kennedy (1992), chapter 9 of Hamilton (1994), and chapter 3 of Hayashi (2000).

prices to equilibrate supply and demand in a given week. Bresnahan's subsequent comment calls this assumption into question. The literature on supermarket pricing, most notably Lal and Matutes (1994), Pesendorfer (2002), and Hosken and Reiffen (2004a), find that supermarkets exercise a considerable amount of discretion in determining temporary promotional price cuts, known as sales, which according to Hosken and Reiffen (2004b) accounts for a considerable amount of retail price variation in supermarkets.[8] This suggests that supermarket prices are likely not being set in a traditional supply-demand equilibrium characteristic of perfectly competitive markets, which is the textbook source of simultaneity.

Moreover, one would have to believe that an unobservable shock to demand in time t occurs in such a way that the price setter observes the shock with enough time to appropriately change the price in time t. For example, consider the effects of national advertising (unobserved) on demand and price. For national advertising to impact the estimated coefficient associated with price in the demand system, national advertising and price would need to be correlated. There is little evidence in the literature to suggest that supermarkets are cognizant of national advertising campaigns undertaken by manufacturers when setting prices, yet the possibility of it appears to be a concern in some IO studies. Bresnahan's (1997) comment on Hausman's (1997) demand study of RTEC point to national advertising as a phenomenon that affects both demand and price, thus rendering Hausman's instruments invalid. Nevo (2001) presents this caveat as well.

For national advertising and prices in supermarkets to be correlated, it would require that the price setters at supermarkets to not only have precise, credible information on the time periods in which the national advertising campaigns are being run by manufacturers, but also information on the efficacy of the advertising with enough time to change prices. As noted by Hausman (1997) and Rubinfeld (2000), supermarkets typically set prices well in advance of effective dates. It is thus unlikely that the prices set by retailers are immediately impacted by the national advertising campaigns of manufacturers, which suggests that si-

[8]They find that about 45 percent of the variation of RTEC prices relative to modal prices is due to temporary price promotions. Similar downward price deviations are noted in Hosken and Reiffen (2001), Pesendorfer (2002), Besanko et al. (2005), and Eichenbaum et al. (2008).

multaneity may not be a problem in differentiated product markets that undergo frequent price changes.[9]

1.2.2 Sources of Price Endogeneity for Differentiated Products

The prices of one retailer may be correlated with prices of another retailer within a particular geographic market for a variety of reasons: (i) Different supermarkets compete for the same pool of consumers in a geographic market; (ii) Different retailers within the same geographic location may have common underlying marginal costs; (iii) Manufacturers may have complex trade promotion schemes among the different supermarket retail chains they deal with; and (iv) Price collusion may exist among supermarkets within a geographic market.[10]

Consider the first case in which two or more different supermarket retail chains compete for the same consumers within a geography. Pesendorfer (2002) analyzes the pricing behavior of supermarkets and finds that competition between retailers for accumulated shoppers is a consideration of retailers' decision to set ketchup on sale. Lal and Matutes (1994) develop a model in which different supermarket chains within a geographic market compete for consumers through price cuts and accompanying local advertising. Hoch et al. (1995) find that the own price elasticity of demand for cereal estimated from scanner data on various Dominick's supermarket locations becomes more elastic the closer competing stores are and the bigger the competing stores are. Shankar and Bolton (2004) also find that competitors affect pricing decisions. Such evidence suggests that prices set by one retail chain may depend on the pricing strategy of other nearby retailers, and that consumers are willing to purchase from different stores depending on prices and promotions.

---

[9] One potential counter-point is that manufacturers determine the national advertising to commit to and may simultaneously determine the level of trade promotion activity or wholesale cost to commit to. This could be a plausible case in which price and national advertising are correlated.

[10] It is possible that firms engage in a form of tacit collusion because they fall into an uncommunicated pattern of price promotions based on observing historical price patterns.

Consider again the demand function for a particular good given by (1.1). Let $-r$ denote supermarket retail chains not r. The existence of price correlation among the retailers within the same geographic market, g, is implied by the condition $P_{r,g,t} \not\perp P_{-r,g,t}$. If the retailers are competing for the same consumers, and the marginal group of consumers are willing to switch between retailers as discussed above,[11] then it's plausible that the demand for a good in one retail chain depends on the pricing behavior of another retail chain within the same geographic market. Such competition for consumers is unobserved in the estimating equation, and would lead to the endogeneity condition that $P_{r,g,t} \not\perp U^Q_{r,g,t}$, which follows from $P_{r,g,t} \not\perp P_{-r,g,t}$ and $P_{-r,g,t} \not\perp U^Q_{r,g,t}$. This potential source of price endogeneity has gone largely unexamined in the IO demand estimation literature, and is examined further in later sections.

Another potential source of price endogeneity stems from consumers' stockpiling behavior for storable goods. Stockpiling can be described as consumers' behavior of purchasing excessive quantities of a good during low-price time periods for consumption in later time periods. In the context of cereal, for example, a consumer may purchase two boxes of a particular cereal when it is on sale, consuming one box during the current time period, while storing or stockpiling the second box for consumption during a later time period.

Hendel and Nevo (2003) and Hendel and Nevo (2006) find stockpiling behavior to be present for many nonperishable products found in supermarkets. The reason this may lead to price endogeneity in demand estimation is as follows. Suppose in time $t-1$ a particular good is temporarily set on sale and consumers stockpile the good, and further suppose that the stockpiling in time $t-1$ leads consumers to inherently purchase less in the next time period t, regardless of the price of the good in t. If the price set in week t depends in some way on the price set in the previous week $t-1$, then the endogeneity condition $P_{r,g,t} \not\perp U^Q_{r,g,t}$ arises because $P_{r,g,t-1} \not\perp P_{r,g,t}$ and $P_{r,g,t-1} \not\perp U^Q_{r,g,t}$. There is some evidence suggesting that price in week t does in fact depend on prices in previous time periods. Pesendorfer

[11] Hosken and Reiffen (2004a) discuss a similar idea called "cream skimming," which is a term they use to describe the act of consumers mixing between different supermarket chains to buy the products on sale at each retailer.

(2002) finds that the probability of a temporary price promotion for ketchup is an increasing function of the time elapsed since the last promotion. Figure 1.4, discussed in Section 4, presents price equations showing that the price of cereal in time t depends extensively on lagged prices as well as the time that has elapsed since the last promotion.

Competition between retailers within the same geographic market and stockpiling behavior are two plausible sources of price endogeneity for storable differentiated products found in supermarkets. These two issues are explored further in the next section with a structural system of equations describing price and demand.

1.2.3 A Structural System Describing Demand and Price

This section provides a structural system of equations that describe the data generating processes (DGP) of price-setting and associated demand response by consumers in a hypothetical product market sold in supermarkets. The system sheds light on the factors that drive both price and demand, which are potential sources of price endogeneity. An explicit set of equations and variables becomes particularly important in later sections for identifying the effect of price on demand with the conditional exogeneity assumption $P \perp U^Q | W$, where W is a set of conditioning instruments that contains drivers of or responses to common underlying determinants of both price and demand.

Demand is given by the structural equation

$$Q_{r,g,t} = d(P_{r,g,t}, U^Q_{r,g,t}), \tag{1.2}$$

where Q and P denote quantity and price(s), respectively, and U^Q denotes factors of demand that are unobservable to the researcher, such as preferences and stockpiling behavior.

The point of interest is the interaction between $P_{r,g,t}$ and $U^Q_{r,g,t}$, for this interaction drives the price endogeneity in the demand function, and, perhaps more importantly, sheds light on the assumption needed to identify the coefficients associated with $P_{a,g,t}$. The variables that comprise $U^Q_{r,g,t}$ will be specific to the analysis at hand as well as the available data. In the context of demand estimation for a typical differentiated product market found in supermarkets, $U^Q_{r,g,t}$ is likely to include consumers' aggregate preferences, denoted $PREF_{g,t}$, the proportion of shoppers in a geographic market that are not store-loyal, denoted $SHPR_{r,g,t}$, and the stockpiling behavior of consumers at a retailer in a geographic market, denoted $STKPL_{r,g,t}$. I.e., the unobservable variables are the set of variables $U^Q_{r,g,t} = \{PREF_{g,t}, SHPR_{r,g,t}, STKPL_{r,g,t}\}$. Other unobservable variables, such as national advertising, are certainly expected to affect demand, but do so through consumers' preferences, and are thus incorporated later in the structural equation for preferences.

The next primary structural equation of interest is for price, given by

$$P_{r,g,t} = f\left[E(D_{r,g,t}), U^P_{r,g,t}\right]. \tag{1.3}$$

The price that supermarket retail chain r located in g sets during t is determined by expected demand, denoted by $E(D_{r,g,t})$, and unobservable variables that drive price, captured by $U^P_{r,g,t}$, most likely related to marginal cost. Although an increase in marginal cost, embodied in $U^P_{r,g,t}$, will lead to an increase in price, it is not clear that an increase in expected demand will necessarily lead to an increase in price. Chevalier et al. (2003) find that prices actually decrease for many goods found in supermarkets during periods of expected seasonal demand peaks, such as summer months and holidays.[12] They conclude such evidence supports a Lal and Matutes (1994) loss-leader strategy by supermarkets to compete for consumers through (locally) advertised prices. Hosken and Reiffen (2004a) have similar findings.

Consider next the structural equation for expected demand given by

$$E(D_{r,g,t}) = g(\widetilde{U}^Q_{r,g,t}), \tag{1.4}$$

where $\widetilde{U}^Q_{r,g,t}$ denotes the variables contained in $U^Q_{r,g,t} = \{PREF,\ SHPR,\ STKPL\}$ that also determine prices. If all of the unobserved variables comprising $U^Q_{r,g,t}$ play a role in the price-setting process, then $\widetilde{U}^Q_{r,g,t} = U^Q_{r,g,t}$. The importance of equations

[12] For example, they find that price of beer drops during Memorial Day, Fourth of July, and Labor Day.

(1.2) - (1.4) taken together are that they directly define the problem of price endogeneity in the demand system. That is, $P_{r,g,t} \not\perp U^Q_{r,g,t}$ since $E(D_{r,g,t}) \not\perp U^Q_{r,g,t}$ by equation (1.4), and $P_{r,g,t} \not\perp E(D_{r,g,t})$ by equation (1.3). Further, by combining equations (1.4), (1.3), and (1.2), the following nested function is obtained

$$Q_{r,g,t} = d\{f[g(\widetilde{U}^Q_{r,g,t}), U^P_{r,g,t}], U^Q_{r,g,t}\}.$$

It is clear that $P_{r,g,t} \not\perp U^Q_{r,g,t}$ since price is a function of $\widetilde{U}^Q_{r,g,t}$ and $\widetilde{U}^Q_{r,g,t} \not\perp U^Q_{r,g,t}$ because of common elements between $\widetilde{U}^Q_{r,g,t}$ and $U^Q_{r,g,t}$. The econometric implication of price endogeneity is that the coefficients associated with $P_{r,g,t}$ in equation (1.2) cannot be identified; meaning they do not have a causal or structural interpretation due to the existing correlation between price and unobserved variables in the demand system. Even if (1.2) is linearly separable, for example $Q_{r,g,t} = P'_{r,g,t}\beta_o + U^Q_{r,g,t}$, the OLS estimate of β_o would at best provide an estimate useful only for predictive purposes, rather than causal inference. However, it is the latter that is of primary interest for inference and decision-making.

Finally, there are a number of sub-structural equations that define the variables in the equations above, which are key to establishing identification in the presence of price endogeneity with the conditional exogeneity assumption described in Section 3.3. Consider the set of structural equations

$$PREF_{g,t} = h_1(DMG_g, U^{PREF}_t) \tag{1.5}$$

$$SHPR_{r,g,t} = h_2(PROM_{r,g,t}, U^{SHPR}_{r,g,t}) \tag{1.6}$$

$$STKPL_{r,g,t} = h_3(TPROM_{r,g,t}, TPROM_{-r,g,t}, U^{STKPL}_{r,g,t}). \tag{1.7}$$

Equation (1.5) proposes that aggregate preferences of consumers depend on the demographics of geographic market g, an observable variable denoted DMG_g, as well as unobserved determinants of preferences, $U^{PREF}_{g,t}$, which contain variables such as national advertising by manufacturers, ADV_t, health trends, $HLTH_t$, and seasonality, $SEAS_t$. Demographics can affect consumers' preferences through age distribution, number of families with children, and race. Health trends and national advertising may affect consumers' preferences as well. Ippolito and Mathios (1990) find that cereal manufacturers' advertising campaigns informing consumers of the health benefits related to consuming fiber found in cereal increased national consumption of fiber-rich cereals. The extent to which these variables are observable or not by the researcher is dictated by the available data set. National advertising and health trends may not be observed, but in the conditional exogeneity framework of Section 3.3, it is appropriate to use proxies, such as variables constructed by exploiting the panel structure of the scanner data. If national advertising, health trends, or any other shock affects consumers' preferences for a particular good, then the quantity purchased of that good ought to be impacted in all geographic markets, denoted $\overline{Q}_t$. Thus, $\overline{Q}_t$ is a response to the elements of U_t^{PREF}, and may be used as a conditioning instrument, W, in the conditional exogeneity assumption $P \perp U^Q|W$ to identify the effect of price on demand. Furthermore, seasonality can be proxied with time fixed effects, FE_t, which may act as a proxy for ADV_t and $HLTH_t$ as well.

Equation (1.6) proposes that the proportion of shoppers in a geographic market that are not store-loyal and that choose to frequent retail chain r during week t is a function of the promotional activity in r, denoted $PROM_{r,g,t}$, and unobservable determinants, $U_{r,g,t}^{SHPR}$. There is evidence that supermarkets within a geographic market compete for the same shoppers through offering various promotions, such as temporary price cuts or sales, local feature advertising, and displays. This is consistent with Lal and Matutes (1994), Pesendorfer (2002), and Hosken and Reiffen (2004a). Although these studies employ different assumptions in their models, they each conclude that supermarkets temporarily promote goods on a week-to-week basis to draw shoppers into stores. The elements that compose $U_{a,g,t}^{SHPR}$ may include factors such as demographic information, which has already been incorporated in (1.5).

Finally, equation (1.7) proposes that the stockpiling behavior of consumers at retailer r is determined by the number of time periods that has elapsed since a promotion occurred at that retailer, denoted $TPROM_{r,g,t}$, the number of time periods that has elapsed since a promotion occurred in rival supermarkets within the same geographic market, denoted $TPROM_{-r,g,t}$, and unobservable determinants of stock-piling behavior, $U^{STKPL}_{r,g,t}$. The dependency of stockpiling on the time that has elapsed since the last promotion is consistent with the findings of Hendel and Nevo (2006) and Pesendorfer (2002). There is also information that can proxy for $U^{STKPL}_{r,g,t}$ by exploiting the panel structure of the scanner data. For example, additional factors that may play a role in the stockpiling of a product is the past history of prices and quantities of that product in the supermarket as well as competing supermarket accounts within the same geographic market. Thus, $P_{r,g,-t}$, $P_{-r,g,-t}$, $Q_{r,g,-t}$, and $Q_{-r,g,-t}$ are proxies for unobserved determinants of stockpiling, and can be potentially useful conditioning instruments. Supermarket fixed effects, $FE_{r,g}$, interacted with lagged quantity, $Q_{r,g,-t}$, may also capture heterogeneity in stockpiling across supermarkets.

Figure 1.13 presents a path diagram for the system of structural equations describing quantity. Figure 1.14 presents a path diagram for the system of structural equations describing prices. Variables outlined in dashed boxes are typically unobservable to the researcher. The diagrams help in determining the confounding variables that prevent structural identification of the parameter that theoretically captures the effect of price on quantity demanded. More specifically, stockpiling behavior, consumer preferences, and consumer habits play a role in determining both demand and price. These confounding variables are not observable, but proxies that drive them or variables that are responses to them may be obtained from the panel structure of the scanner data, as well as publicly available income and demographic data. The following section provides identification assumptions for recovering the effect of price on demand in the presence of price endogeneity.

1.3 Identification

The following sections provide analyses related to identification of price effects in demand systems when (i) Prices are exogenous; (ii) Prices are endogenous, but cost shifters exist and are exogenous; (iii) Prices are endogenous and cost shifters do not exist, but proxies for cost shifters exist and are exogenous; and (iv) Prices are endogenous, but proxies for the confounding effects causing the price endogeneity exist.

Suppose that the functions $d(\cdot)$, $f(\cdot)$, $g(\cdot)$, $h_1(\cdot)$, $h_2(\cdot)$, and $h_3(\cdot)$ in the previous section are standard linear-separable parametric equations. Equation (1.2) is thus given by

$$Q_{r,g,t} = P^{'}_{r,g,t}\beta_o + U^Q_{r,g,t}, \tag{1.8}$$

where again P is a $(1 \times J)$ vector of prices, and β_o is the marginal effect of price on demand. When price exogeneity fails ($P \not\perp U^Q$), identification of the causal effect of price on demand with a regression of quantity on price is not possible; rather, only predictive estimates are possible that may not be causally meaningful. The desired price effect is tantamount to having data available in which the prices of goods in a relevant product segment are randomly changed by reasonable increments across retailers, location, and time. In such an ideal world, even a basic regression of purchased quantities on market prices ought to yield the causal price effects. Because unobservable (to the researcher) factors that drive market price and demand overlap, as detailed in Section 2.3, rendering market prices endogenous, estimating causal effects as if one had data from an ideal randomized experiment becomes problematic. The following sections examine identification assumptions.

1.3.1 Exogenous Prices

If prices are in fact randomly set in the product market in question, then prices are by definition exogenous; that is, $P_{r,g,t} \perp U^Q_{r,g,t}$. Suppressing subscripts, price exogeneity implies the moment condition $E(PU^Q) = 0$, which yields the familiar ordinary least squares estimator

$$\widehat{\beta}^{ols} = (\mathbf{P'P})^{-1}\mathbf{P'Q}, \tag{1.9}$$

where $\mathbf{P}$ is a $N \times K$ matrix[13] of price variables and $\mathbf{Q}$ is a $N \times 1$ vector of quantity data for good j. From the analyses presented in Section 2, there is very little reason to believe that prices are exogenous; thus, an OLS regression of quantity on price does not provide an estimate of the causal price effect. It does, however, present a benchmark case to compare other estimates to.

1.3.2 Exogenous Instruments - Cost Shifters

A popular remedy for price endogeneity in demand systems involves the use of instrumental variables (IV), typically cost shifters. Cost shifters are arguably relevant because a good's price is affected by its production costs, and arguably valid or exogenous since costs do not affect demand for the good except through the good's price. Let Z denote a vector of cost shifters. Consider again the linear demand function given by (1.8), where $P \not\perp U^Q$, but $Z \perp U$. Instrument exogeneity, given by the condition $Z \perp U^Q$, allows for identification of β_o by way of the moment condition $E(ZU^Q) = 0$. The plug-in estimator that results is the familiar instrumental variables estimator

$$\widehat{\beta}^{iv} = (\mathbf{Z}'\mathbf{P})^{-1}\mathbf{Z}'\mathbf{Q}, \tag{1.10}$$

where $\mathbf{Z}$ is a $N \times L$ matrix of instruments or cost shifters, $\mathbf{P}$ is a $N \times J$ matrix of pricing variables, and $\mathbf{Q}$ is a $N \times 1$ vector of quantity data.

The difficulty of the cost shifter approach to IV is that it quickly becomes an infeasible source of instruments for differentiated product markets, primarily because it is difficult for the researcher to procure as many cost shifters as there are endogenous price variables. Therefore, far more often than not, cost shifters are unobservable. This notion is captured by equation (1.3) defining the price-setting process, which shows that marginal costs are unobserved, and are thus captured by $U^P_{r,g,t}$. Finding a few valid cost shifters in practice is difficult; finding over 50 is virtually impossible without further potentially restrictive assumptions.

[13] N contains the panel r, g, t. K contains $j = 1, 2, ..., J$ products, as well as a constant. K may also factor in other demand drivers that are exogenous.

Hausman Instruments

In cases in which the scanner data contains geography, such as city or metro area, as one of its dimensions, instruments that have come to be widely known among IO practitioners as "Hausman instruments" have been used as proxies for the unobserved cost shifters. The main idea of Hausman instruments is that prices in other geographic markets in the panel can be employed as instruments for prices in a particular geographic market. This source of instruments is adapted from Hausman and Taylor's (1981) development of instrumental variables for time invariant characteristics in a panel data setting, and later applied to demand estimation when scanner data is available by Hausman et al. (1994), Hausman (1997), Hausman and Leonard (2002), and Nevo (2001), among other studies. The usefulness of Hausman instruments stems from that fact that all the instruments are contained in the scanner data. Although the J different cost shifters, one for each of the J goods in the system, are unobservable, prices in other geographies, denoted $P_{-g,t}$, can be employed as proxies for the unobserved nationwide costs, C_t.

Although Hausman instruments provide a feasible solution even with several goods in the demand system, it has been pointed out that the assumption underlying the validity of Hausman instruments may be too restrictive, particularly for nationally-branded consumer goods markets. This line of criticism is well-known in empirical IO, and has been presented as a caveat in research involving the use of Hausman instruments.[14] The researcher must make at least two assumptions in order to successfully employ Hausman instruments: (i) The unobserved shocks to product costs affect all geographic markets, and (ii) There are only geography-specific unobservable demand shocks and not nationwide demand shocks.

The first assumption, captures the notion of instrument relevance. That is, $P_{r,g,t}$ and $P_{-g,t}$ are correlated to the extent that underlying changes in the nationwide component of the cost of the good, C_t (unobserved), will affect the price of the good in all geographies, not just geography g. The second assumption is neces-

[14]E.g., Bresnahan (1997), Nevo (2000a), and Nevo (2001).

sary to justify the exogeneity of Hausman instruments. If there exist nation-wide demand shocks, then prices in geographies $-g$ would be correlated with the error term in the demand function for geography g, and Hausman instruments would thus not be exogenous. Hausman and Leonard (2005) write: "The intuition is that prices in each city reflect both underlying product costs and city-specific factors that vary over time as supermarkets run promotions on a particular product. To the extent that the stochastic city-specific factors are independent of each other, prices from one city can serve as instruments for another city."

The underlying exogeneity assumption of Hausman instruments appears to be the primary source of controversy in the demand estimation literature. In his comment on Hausman's (1997) paper, Bresnahan (1997)[15] points out that it is difficult to justify the assumption that there is an absence of nationwide demand shocks in nationally-branded differentiated product markets, such as ready-to-eat cereal. Nevo (2000b) acknowledges this potential drawback as well. For example, consider again national advertising campaigns. Advertising campaigns, such as television promotions, are likely to be nationwide events that are determined by manufacturers. Demand shocks at the national level could violate the assumption that there are only geography-specific (i.e., city-specific) demand shocks. If $P_{-g,t} \perp U^Q_{r,g,t}$ does not hold, then instrumental variables estimates of β_o employing Hausman instruments, will not yield the causal effect of price. This may or may not be a credible concern depending on the extent to which the researcher believes that the prices being set by retailers are driven by national advertising or other shifts in consumers' preferences. Recall that equations (1.3) and (1.4) suggest that prices do theoretically depend on unobserved demand expectations, but the strength of the relationship depends entirely on the retailer's ability to forecast demand when setting prices. Therefore, the correlation between $P_{-g,t}$ and $U^Q_{r,g,t}$

[15] See also Hosken et al. (2002). It appears a debate developed between Jerry Hausman and Timothy Bresnahan regarding this point. Hausman's initial study and Bresnahan's comment on it appeared in *The Economics of New Goods*, Timothy Bresnahan and Robert J. Gordon, eds., NBER Studies in Income and Wealth Number 58, The University of Chicago Press, 1997. Hausman then wrote a note, "Reply to Prof. Bresnahan" (1997), and Bresnahan later responded with "The Apple Cinnamon Cheerios War: Valuing New Goods, Identifying Market Power, and Economic Measurement" (1997).

may be strong or weak depending on the circumstances.

Although the validity of Hausman instruments is typically the source of econometric concern, very little has been said about the potentially important issue that Hausman instruments are weak instruments. Recall that instrument relevance requires $P_{-g,t} \not\perp P_{r,g,t}$, which is justified by the assumption that $P_{-g,t} \not\perp C_t$ and $P_{r,g,t} \not\perp C_t$. Consider structural equation (1.3) without a retailer distinction, $P_{g,t} = f(E(D_{g,t}), U^P_{g,t})$. This shows that price in geography g during time t is determined by expected demand and unobserved price determinants, such as input costs. Hausman instruments need for $U^P_{g,t}$ to include a national cost component, U^P_t, that is common across geographies so that $P_{g,t} \not\perp P_{-g,t}$. While it is economically reasonable to believe that exogenous shocks to underlying costs ought to affect the price of a good in all geographic locations, it is unlikely that underlying aggregate productions costs vary enough to create meaningful correlation between price in geographic market g and other geographic markets $-g$ on a week-to-week basis.

Bound et al. (1995) and Staiger and Stock (1997) explain that if the correlation between the instruments and the endogenous variable is low, i.e. weak instruments, then the IV estimator is biased towards the OLS estimator. This may be a potential explanation of why econometric studies[16] tend to reject the hypothesis of price endogeneity when employing Hausman instruments for the Hausman (1978) specification test. The Hausman specification test requires valid instruments to compare the difference between the IV and OLS estimates. If the difference is sufficiently small, then one may reject the hypothesis that endogeneity is a problem. If the instruments are weak, then the IV estimates may be biased towards the OLS estimates, causing the researcher to incorrectly reject price endogeneity with a Hausman specification test.

Weinberg and Hosken (2008) estimate demand for breakfast syrup using

[16] Hausman (1997) indicates no difference in his IV and OLS results at the brand level for cereal demand, and Scheffman and Spiller (1996) indicate no difference in their IV and OLS results at the brand level for butter/margarine demand. To note, neither study reports the first stage of the IV regression, so it is not possible to assess the weakness of Hausman instruments, which may or may not be an issue for these studies. Staiger and Stock (1997) note that most studies omit reporting the first stage R^2 and F-statistic.

scanner data. In their analysis, they employ Hausman instruments, and find that these instruments are weak as measured by the first-stage partial F-stat (all less than four). Their first-stage results and subsequent estimates of price effects are similar to the estimates for cereal in this chapter. Both are suggestive of instrument weakness. The next section presents an alternate identification strategy.

1.3.3 Conditionally Exogenous Prices

While cost shifters are typically not available for IV estimation, and Hausman instruments may either be invalid or weak, there is still an identification strategy that has yet to be examined in the context of demand. The import of equations (1.2) - (1.7) are they provide guidance on obtaining conditioning instruments necessary to structurally identify price effects with the conditional exogeneity identifying assumption.

The conditional exogeneity (CX) identification framework employed here is contained in the research of White (2006), White and Chalak (2006), White and Chalak (2009), and Chalak and White (2010), to name a few. They provide a unified framework for estimating causal effects in structural systems that encompasses research in statistical theory (e.g., Dawid (1979)), artificial neural networks (e.g., Pearl (2000)), familiar econometric procedures (e.g., instrumental variables), the treatment effects literature (e.g., Rubin (1974) and Rosenbaum and Rubin (1983)), and extensions primarily employed in labor economics (e.g., Heckman, Ichimura, and Todd (1997), and Heckman and Vytacil (2005)). This framework is applied here as an alternate procedure to remedy the statistical dependence between market price and unobserved demand shocks.

The failure of exogeneity is due to the relationship among the variables in structural equations (1.2) - (1.4), but the sub-system of structural equations that further define those variables, given by equations (1.5) - (1.7), provide predictive proxies or responses for the expected demand that the retailer or manufacturer may consider when setting prices. In other words, the collection of observable variables that drive or are driven by the unobservable variables determining both price and quantity in Figure 1.13 and Figure 1.14 are the conditioning instruments defined by $W_{r,g,t} = \{\overline{Q}_t,\ DMG_{g,t},\ FE_t,\ FE_{r,g},\ PROM_{r,g,t},\ TPROM_{r,g,t},\ TPROM_{-r,g,t},$ $P_{r,g,-t},\ P_{-r,g,-t},\ Q_{r,g,-t},\ Q_{-r,g,-t}\}$, the elements of of which were defined in Section 2.3. Because the source of price endogeneity is due to the retailer potentially taking into consideration expected demand when setting price, and because $W_{r,g,t}$ contains the set of variables that may determine expected demand, then $W_{r,g,t}$ is the set of conditioning instruments that identify the effect of price on quantity.

The idea behind CX is that price exogeneity can be achieved by conditioning on a set of instruments that proxy for the potential sources of endogeneity, but these instruments need not be exogenous in the standard sense of instrumental variables regression. In the demand application here, price and the unobserved determinants of demand are exogenous when conditioning on the set of confounding variables or conditioning instruments, $W_{r,g,t}$. The key CX identifying assumption is given by (suppressing subscripts)

$$P \perp U^Q | W.$$

Similar to preceding sections, let $Q_{r,g,t}$ denote the quantity purchased of a good and $P_{r,g,t}$ denote prices, and let $W_{r,g,t}$ denote the vector of conditioning instruments as defined above. Consider again a linear demand function given by $Q_{a,g,t} = P'_{r,g,t}\beta_o + U^Q_{a,g,t}$, where price endogeneity, implied by $P_{r,g,t} \not\perp U^Q_{r,g,t}$, is present, but that conditional exogeneity, $P_{r,g,t} \perp U^Q_{r,g,t} | W_{r,g,t}$, holds. Identification of β_o can be obtained from the conditional exogeneity assumption as follows.[17] The assumption $P \perp U|W$ (suppressing subscripts and superscripts) implies the moment condition

$$E(PU|W) - E(P|W)E(U|W) = 0.$$

Rearranging gives $E\{PU - E(P|W)U\ |W\} = 0$, or alternately $E\{[P - E(P|W)]U\ |W\} = 0$. Substituting the regression representation of $E(P|W)$, given by $E(P|W) = E(PW')E(WW')^{-1}W$, into the preceding expression yields

$$E\{[P - E(PW')E(WW')^{-1}W]U\ |W\} = 0.$$

[17] A similar derivation is given in Chalak and White (2010). Given its relatively recent examination as an identification strategy, it warrants reexamination here.

By the law of iterated expectations, we have $E\{[P - E(PW')E(WW')^{-1}W]U\} = 0$. Substituting for $U = Q - P'\beta_o$ yields $E\{[P - E(PW')E(WW')^{-1}W][Q - P'\beta_o]\} = 0$, and, assuming the appropriate rank conditions to ensure invertibility, β_o is given by

$$\beta_o = [E(PP') - E(PW')R^{-1}E(WP')]^{-1}[E(PQ) - E(PW')R^{-1}E(WQ)].$$

where R is given by $E(WW')$ for notational simplicity. The corresponding plug-in estimator is

$$\widehat{\beta}^{cx} = [\mathbf{P}'\mathbf{P} - \mathbf{P}'\mathbf{W}(\mathbf{W}'\mathbf{W})^{-1}\mathbf{W}'\mathbf{P}]^{-1}[\mathbf{P}'\mathbf{Q} - \mathbf{P}'\mathbf{W}(\mathbf{W}'\mathbf{W})^{-1}\mathbf{W}'\mathbf{Q}] \tag{1.11}$$

where $\mathbf{P}$ is a $N \times J$ matrix of price variables, $\mathbf{W}$ is a $N \times L$ matrix[18] of conditioning instruments, and $\mathbf{Q}$ is a $N \times 1$ column vector of quantity.

Although $W_{r,g,t}$ is not valid or exogenous in the standard sense of classical instrumental variables analysis, it is still instrumental in identifying β_o in the demand system. Hence, the identification strategy derived above is classified as an extended instrumental variable (EIV) estimation procedure by Chalak and White (2010). For comparison purposes, consider again the instrument exogeneity identifying assumption for Hausman instruments to recover structural parameters in the traditional IV procedure given by $P_{-g,t} \perp U^Q_{r,g,t}$. This is fundamentally different from the conditional exogeneity assumption proposed here that $P_{r,g,t} \perp U^Q_{r,g,t}|W_{r,g,t}$. There is no need for the conditioning instruments to be uncorrelated with the unobserved determinants of demand, as is required for Hausman instruments, as long as price is uncorrelated with the demand drivers after conditioning on the instruments W.

Chalak and White provide two insightful interpretations of the CX estimator. First, $\widehat{\beta}^{cx}$ is the coefficient vector associated with $\mathbf{P}$ from a regression of $\mathbf{Q}$ on $\mathbf{P}$ and $\mathbf{W}$. Importantly, the identification assumption $P \perp U^Q|W$ does not require that the coefficient estimates associated with the conditioning instruments

[18] N is the row dimension of the panel data. L need not be the same size as J as in traditional IV analysis. The dimension of the set of conditioning instruments depends on the number of proxies available for the confounding variables. As always, a constant (or vector of ones) ought to be included in the analysis.

have sensible signs or magnitudes because nothing in the assumption suggests that the coefficients associated with W are identified. In other words, the conditioning instruments only serve to identify the structural effect of price on quantity, and the coefficients associated with the conditioning instruments only have a predictive interpretation instead of a structural or causal one. This notion is somewhat analogous to the first stage of traditional IV estimation in that the parameter estimates of the first stage of two stage least squares (TSLS) only serve to identify β in the second stage, but the first stage estimates have no structural or causal meaning. The second interpretation of $\widehat{\beta}^{cx}$ is that it is the second stage estimate of $\mathbf{Q}$ on $\widehat{\mathbf{U}}$, where $\widehat{\mathbf{U}}$ is the residual from the first stage regression of $\mathbf{P}$ on $\mathbf{W}$. In this context, it is readily seen that the conditioning instruments W identify the price effect in the second stage by doing the exact opposite of what standard exogenous instruments Z do in the first stage, which is why conditional exogeneity requires using the fitted error from the first stage as opposed to fitted price. The conditional exogeneity assumption is evaluated in the next section with a simulation.

1.3.4 Simulation

In this section, a simulation is performed based on the system of equations describing price and demand proposed in Section 2.3. The following subsections discuss the data generation process (DGP) of the variables found in equations (1.2) - (1.4), although some of it has been simplified here.

Data Generation Process

First, a time period is generated $t = 1, 2, ..., T$; each unit of time is considered to be a week. Two hypothetical geographic markets are created, and within each geography exists two supermarket retail chains that compete with each other. For the moment, consider just geography one, indexed $g = 1$, and retailer one, indexed $r = 1$. The price set by retailer one in geographic market one, denoted $P_{1,1,t}$, is generated from $U^P_{1,1,t}$ and $E(D_{1,1,t})$ as described by equation (1.3). $U^P_{1,1,t}$ can be thought of as the wholesale price that retailer one must pay to the manufacturer to obtain the good; i.e., the retailer's marginal cost that is unobservable to the researcher. $U^P_{1,1,t}$ contains a baseline manufacturing cost that increases every year (or approximately every 50 weeks for $t = 1, 2, ..., T$) by a percent. The particular a that is used in the simulation plays a significant role in determining the strength or weakness of the Hausman instrument since increases in manufacturing costs gets passed on to the prices set in the supermarkets in both geographic markets. For simplicity, it is assumed that the retailer purchases the good at the production cost of the manufacturer.

Next, retailer one decides which time periods to set the good on sale.[19] To implement this, expected demand in geographic market one, $E(D_{\cdot,1,t})$ is generated as a uniform(0,1) random variable. During time periods in which $E(D_{\cdot,1,t})$ is between v_1 and v_2, where v_1 and v_2 are between zero and one such that the difference in the cumulative uniform distribution at v_1 and v_2 is set to 0.20, which indicates that the good is promoted by retailer one about 20 percent of the time. When the good is promoted, the retail price is cut by 50 percent to indicate the sale price. 1.19 (available upon request from author) presents one draw of the simulation for retail price. The pattern of the price series that emerges is reasonably characteristic of what researchers have documented using actual scanner data. See Hosken and Reiffen (2001), Pesendorfer (2002), Besanko et al. (2005), and Eichenbaum et al. (2008), as well as Figure 1.15 of this chapter. The data generation process for retailer two is similar to the procedure described for retailer one, except temporary price promotions are generated in time periods so that there is some overlap with retailer one. In other words, the variable $E(D_{\cdot,1,t})$, which represents the expected demand that retailers one and two face in geography one, causes for some correlation of prices within a geographic market.

For retailer one in geography one, the aggregate quantity purchased by consumers, $Q_{1,1,t}$, is generated as $Q_{1,1,t} = \alpha + \beta P_{1,1,t} + \gamma E(D_{1,1,t}) + \varepsilon_{1,1,t}$, where $P_{1,1,t}$ and $E(D_{1,1,t})$ are price and expected demand, respectively, and and ε is a

[19]Sales and promotions are a major point of price variation for supermarket retail accounts. See Hosken and Reiffen (2004) and Pesendorfer (2002). It is reasonable that the manufacturer plays a role in retail promotions through trade deals, but as Besanko et al. (2005) point out, the retailer ultimately decides the time period in which to offer the promotions.

$N(0,1)$ random error. Clearly, $P \perp \varepsilon$ since ε is random. However, if $E(D_{1,1,t})$ is unobservable, then the estimating equation becomes $Q_{1,1,t} = \alpha + \beta P_{1,1,t} + u_{1,1,t}$, where $u_{1,1,t} = \gamma E(D_{1,1,t}) + \varepsilon_{1,1,t}$. Since $E(D_{1,1,t})$ played a role in the DGP of $P_{1,1,t}$ as described above, we have the endogeneity condition $P \not\perp u$. Therefore, it is argued here that if the econometrician cannot observe expected demand, thus estimating $Q_{1,1,t} = \alpha + \beta P_{1,1,t} + u_{1,1,t}$ instead of $Q_{1,1,t} = \alpha + \beta P_{1,1,t} + \gamma E(D_{1,1,t}) + \varepsilon_{1,1,t}$, the OLS estimate of β does not yield the structural effect of price on quantity.

The steps for the retail price DGP discussed previously are also employed for two supermarket retail chains in the second geographic market. It is necessary to construct data for a second geography for use in the analysis of Hausman instruments. For purposes here, the same underlying manufacturing cost of the good that was used for the retailers in geographic market one is also used for the retailers in market two. This ultimately becomes the basis for the relevance of the Hausman instrument. More specifically, the Hausman instrument for price in retail chain one in geography one during time t is the average price of the two retail chains in geography two during time t. Finally, a conditioning instrument is constructed based on expected demand. The DGP for the conditioning instrument for retail chain one in geography one is $Z_{1,1,t} = \theta + \pi E(D_{1,1,t}) + \mu_{1,1,t}$ where $\mu_{1,1,t}$ is distributed $N(0, \sigma^2)$. The σ that is chosen for the DGP drives the strength or weakness of the conditioning instrument.

Simulation Results

The simulation results found in Figure 1.9 are based on the following parameters: $\alpha = 1000$, $\beta = -100$, and $\gamma = -50$. The weakness or strength of the Hausman instrument with price depends on a, which takes on a values between 0.05 and 0.15. The weakness or strength of the conditioning instrument with expected demand depends on σ, which takes on a values between 10 and 0.25. The first 10 numbers listed are the results of the first 10 draws of the simulation. The summary statistics below them present the mean and standard deviation of the full 10,000 draws of the simulation. The sample of size is 200 weeks.

The first three columns corresponding to the weak correlation scenario in which the Hausman instrument is weakly correlated with price and the conditioning instrument is weakly correlated with expected demand, the source of the price endogeneity in the DGP. The results indicate that OLS, IV with Hausman instruments, and CX with conditioning instruments do not recover the structural parameter of interest, $\beta = -100$, as expected. The OLS estimate is $\widehat{\beta}^{ols} = -85$, the IV estimate is $\widehat{\beta}^{iv} = -123$, and the CX estimate is $\widehat{\beta}^{cx} = -86$. A noteworthy point is that the standard deviation of the IV estimator is extremely large, a consequence of the weakness of the Hausman instrument as indicated by the very low first stage F-stat of 1.85. Although both the Hausman instrument and the conditioning instrument were constructed to be weak,[20] and thus destined to fail in recovering $\beta = -100$, the conditioning instrument yields far more stable results.

The middle three columns corresponding to moderate correlation still show that the OLS estimator does not recover the true β parameter, but that the IV and CX estimates are getting much closer. The first stage F-stat for the Hausman instrument is 9.38, indicating that the instruments are no longer very weak. To get that result, a was set to 0.10, meaning that nationwide manufacturing costs increase by 10 percent every 50 weeks, and these costs are passed on in the retail prices in both geographic markets perfectly contemporaneously by the supermarket. The final three columns correspond to the case in which the Hausman instrument is strongly correlated with price and the conditioning instrument is strongly correlated with expected demand. Again, the OLS estimate does not recover the true parameter $\beta = -100$, but now the IV and CX estimates are both successful, although the IV estimates yield a relatively high variance.

In practice, the question of whether Hausman instruments are weak or not is an empirical one. The empirical estimates of demand for cereal presented in Figures 1.5 and 1.7, as well as the distribution of first stage F-Stats presented in Figure 1.18, all point to the conclusion that the Hausman instruments derived from the scanner data here are empirically weak. This is consistent with the findings of Weinberg and Hosken (2008). This corresponds to the simulation results in which

[20]For IV, the Hausman instrument is weak if it is uncorrelated with price. For CX, the conditioning instrument is weak if it is uncorrelated with expected demand.

the IV estimator with Hausman instruments failed to recover the structural effect of price on quantity. Furthermore, the empirical instability and unreliability of the IV estimation in Figure 1.5 for the 65 cereal products is very similar to the erratic behavior of the simulation results based on the weak Hausman instruments in Figure 1.9.

1.4 Application to RTEC

In this section, the previous identification strategies and their resulting estimators are applied to estimating the demand for ready-to-eat cereal (RTEC) products. The RTEC industry has received much attention by IO and marketing researchers. Some examples that directly study this industry include Schmalensee (1978), Ippolito and Mathios (1990), Hausman (1997), Rubinfeld (2000), Nevo (2000a, 2001), Nevo and Wolfram (2002), and Shum (2004). Nevo (2000a, 2001) provides an informative background on the history and competitive nature of the RTEC market.

The four major nationally-branded cereal manufacturers operating in the U.S. are General Mills, Kellogg, Post, and Quaker. A fifth important source of cereal sales are store brand products. General Mills and Kellogg are stand-alone, publicly traded companies that also sell a variety of other food products. Kellogg owns the Kashi line of cereals it acquired in June 2000. Until recently, Post was owned by Kraft Foods. In August 2008, Ralston Corp., the primary manufacturer of store brand cereals, acquired the Post cereal business from Kraft for approximately $2.6 billion. Finally, Quaker is owned by Pepsi Co., another publicly traded company that primarily operates in the beverage market. Figure 1.1 presents market share information examined further below.

1.4.1 Data

The data are supermarket scanner data from Information Resources, Inc. (IRI). The data set consists of variables measuring price, quantity, and merchandising and promotional variables for 150 of the top-selling RTEC stock-keeping units (SKUs)[21] for three years beginning January 2005 and ending December 2007. The data has a panel structure, where the time dimension has a weekly frequency and the individual dimension is a particular supermarket retail chain operating in a particular geographic market. Except for cases of missing observations and other irregularities, there are 157 weeks of data for each of the retail chains, and there are 121 retail chains covering 41 geographic markets across the United States.[22]

Although there are some differences, IRI's definition of a geographic market is roughly equivalent to the Census Bureau's metropolitan statistical area (MSA) or combined metropolitan statistical area (CMSA), which is convenient for merging income and demographics data with the scanner data. On average, there are about three major retailers comprising each of the 41 geographic markets in the sample. Figure 1.2 presents a variety of summary statistics for price, quantity, merchandising variables, and distribution[23] for the top-selling product of each of the five manufacturers.

This level of detail in market-level scanner data is typically not found in most IO studies of demand. Hausman's (1997) estimation of demand for RTEC was for aggregated brands, and while the frequency was weekly, Hausman's data did not have supermarket chain detail, only geography-level detail. In other words, that data set aggregated chain-level supermarket prices and quantities up to the city-level. Figure 1.15 demonstrates the information that is potentially lost when data comes in an aggregated form. The first price series is the retail

[21] A stock-keeping unit (SKU) is the most detailed level of a product sold in a supermarket. SKUs consist of the brand, flavor, and size of a particular product. For example, General Mills Cheerios 10oz. and General Mills Cheerios 15oz. are two distinct SKUs. Those two SKUs along with other sizes of General Mills Cheerios comprise the brandline General Mills Cheerios. Many SKUs, even those that made the top 150 best-selling cutoff, have very low sales because they are not widely popular products and are thus not carried in many supermarkets across the country. The full set of SKUs that were ultimately used in the estimation consisted of 65 out of the 150 original products in the data set.

[22] The original dataset contained 172 supermarket retail chains spanning 63 geographic markets. However, due to many missing observations and other irregularities, data cleaning led to the remaining 121 retailers spanning 41 geographies. What is left is still a detailed, comprehensive set of data that covers all major parts of the U.S.

[23] See Little (1998) for a detailed explanation of the merchandising variables and the distribution variable typically found in scanner data.

price of General Mills Cheerios 15oz. in just one of the five supermarket chains in a major metropolitan geographic market. The other price series is a weighted average price of the same product for all five supermarket chains in the same metropolitan geographic market. There is a noticeable difference in pricing. The single retailer series has an average price of \$4.12 and standard deviation of \$0.71, and the five-supermarket aggregated price series has an average price of \$3.27 and standard deviation of \$0.58. The correlation between the two price series is only 0.39. Nevo's (2001) estimation of demand for RTEC was also brand-level price and quantity data aggregated to the city-level, and the frequency was quarterly, not weekly. Hosken et al. (2002) point out that estimating demand systems using aggregated scanner data can cause biased results and incorrect statistical inference. Chapter 2 shows that the aggregation bias results in demand estimates that are much less elastic compared to disaggregated scanner data.

Income and demographic data from the Bureau of Labor Statistics (BLS) and the Census Bureau is merged with the scanner data. IRI's geographic markets are roughly the same as the U.S. government's survey definition of an MSA or CMSA. Average weekly wage data for each geographic market is from the BLS's Quarterly Census of Employment and Wages (QCEW) database. As the name suggests, the wage data is collected on a quarterly basis; thus, while the scanner data contains data at a weekly frequency, the QCEW wage data is only at a quarterly frequency. Demographic information for each of the IRI geographic markets is from the Census Bureau's 2000 census. Variables such as the distribution of age, race, gender, and households with children is merged with the scanner data. Clearly, there is no intertemporal variation in these data, but the lack of such variation in the analyses undertaken here is likely to be of negligible consequence since the scanner data covers only a three-year period, which is not long enough to cause for concerns of meaningful inter-city migration. Figure 1.3 presents a table of descriptive statistics for the income and demographic variables.

Market Share and Pricing

The four major manufacturers of RTEC, as well as several store brand products, are represented in the data. Figure 1.1 presents aggregate market share information on the 150 top-selling products (SKUs). Kellogg and General Mills are the leading nationally-branded manufacturers based on dollar sales and pound (quantity) sales, followed by Post and Quaker. Store brands consists of the private label cereals that sell under the supermarket's brand for a lower price. Figure 1.1 shows that store brand products are priced 32 percent below the total average price of all 150 products. All four of the major manufacturers are priced well above the store brand products, selling at a premium between 32 to 69 percent.

Figure 1.16 displays the monthly market shares of the four RTEC manufacturers as well as store brand products. The store brand market share does not deviate much from five percent on a month-to-month basis. In fact, the average market share of the store brand products across the 36 months ending December 2007 is five percent with a standard deviation of only 0.20 percent. Post and Quaker's market shares vary more than store brands, but are still relatively less volatile than the two biggest manufacturers in the RTEC market, Kellogg and General Mills. The average monthly market share for Post and Quaker are 14.7 percent and 7.6 percent, with standard deviations of 0.7 percent and 0.6 percent, respectively. In contrast, Kellogg and General Mills have market shares that vary more over time than the other manufacturers. Kellogg has an average market share of 37.9 percent with a standard deviation of 1.4 percent, and General Mills has an average market share of 34.9 percent with a standard deviation of also 1.4 percent. Figure 1.16 indicates that not only do the market shares of Kellogg and General Mills vary considerably over time, but that they are nearly mirror images. The correlation of the monthly market share for Kellogg and General Mills is -0.82, indicating that Kellogg's gain in share is General Mills' loss, and vice-versa.

Figure 1.17 presents the weighted average monthly retail prices for the manufacturers between 2005 and 2007. General Mills tends to be sold at a higher price than the other manufacturers, while the store brands products are sold at a deep discount compared to the nationally-branded products. The variation of the manufacturers' prices over time is due to temporary price promotions that are more pronounced in the weekly data. For example, Figure 1.15 presents the weekly price points for General Mills Cheerios 15oz. in a particular supermarket chain located in a major metropolitan geographic market, as well as the weekly price of the same product but at the geographic market level. As pointed out earlier, the difference in the patterns between the two series is indication of the amount of information that is contained in the weekly scanner data that may be lost when researchers use data that is aggregated. This is a particularly important point when the product set being examined is sold through supermarkets in which temporary weekly price cuts are an important source of price variation.[24]

Figure 1.4 presents pricing equations for the top-selling product of each of the five manufacturers. The price equations are an empirical estimation of the structural equation for price proposed in Section 2.3 and Figure 1.14. Here, log price is regressed against the conditioning instruments $W_{r,g,t}$ as well as the Hausman instrument, which is used as a proxy for marginal cost. The results indicate that time since last sale or promotion, lag prices in the same supermarket, lag quantities in the same supermarket, and demographic variables have a statistically significant impact on prices of cereals. Lagged prices and lagged quantities of the product in the next-largest supermarket account within the same geographic market appear to not have much impact on the price-setting behavior of supermarkets.

1.4.2 Demand Estimation

In this section, the OLS, IV, and CX estimators examined in Section 3 are used to estimate demand for RTEC using the almost ideal demand system (AIDS) of Deaton and Muellbauer (1980).[25] It is well documented that AIDS has a high level of econometric flexibility, in that even if the true demand function

[24] See, e.g., Pesendorfer (2002) and Hosken and Reiffen (2004).

[25] A constant elasticity of demand (or log-log demand) form was also estimated. The basic comparative results were the same as the AIDS model.

that describes the data is not AIDS, it will provide an approximate estimate that is near the true demand estimate.

The estimating equation for AIDS is given by

$$s_j = \alpha_j + \sum_k \gamma_{jk} \ln p_k + \beta_j \ln\left(\frac{X}{P}\right) \tag{1.12}$$

for all products indexed $j = 1, 2, ..., J$. The scanner panel subscripts, rgt, which correspond to retailer, geographic market, and time period, respectively, are omitted here for simplicity. The variables s and p denote the expenditure share and price, respectively. X is the total expenditure for the product segment and P is a price index given by the Stone price index

$$\ln P = \sum_i s_i \ln p_i. \tag{1.13}$$

For further details on AIDS, see Deaton and Muellbauer (1980), Chalfant (1987), Green and Alston (1990), Hausman (1997), and Hausman and Leonard (2002). The price elasticities are calculated based on the coefficient estimates from (1.12) as $\eta_{jk} = -\delta_{jk} + \frac{\gamma_{jk}}{s_j} - \beta_j \frac{s_k}{s_j}$ where $\delta_{jk} = 1$ if $j = k$ and 0 else. The expenditure elasticity is given by $\eta_{jx} = 1 + \frac{\beta_j}{s_j}$. See Green and Alston (1990) for a detailed explanation of elasticities in AIDS models.

Figure 1.5 presents a summary of demand elasticity estimates for 65 RTEC products using the three identification schemes described in Section 3. To note, no restrictions, such as symmetry or homogeneity, were imposed in the estimation of the AIDS model. Figures 1.6 through 1.8 present elasticity matrices of product groupings that ought to have high intra-group substitution patterns. These results are examined in the following sections.

Results - OLS

The OLS estimation of AIDS[26] for the 65 cereal products yields fairly reasonable results. Column (1) of Figure 1.5 presents the own price elasticities of

[26]Given the assumption of exogenous regressors (price), only log prices and expenditures were included as RHS variables as presented by the AIDS estimating equation (1.12).

the AIDS. The summary statistics at the bottom of Figure 1.5 indicate that the average elasticity across all 65 products is about -3, corresponding to an average implied price-cost margin (PCM) of 36% found in column (4). The four store brand SKUs in the sample tend to be less elastic, with an average elasticity of about -2, compared with their nationally-branded counterparts. This certainly makes sense since retailers earn higher profits on store brand products.

To note, the price-cost margin for good j in Figure 1.5 is based on the Lerner index, $PCM_j = -1/\eta_j$, where η_j is the own price elasticity of good j. This is a reasonable calculation if each of the 65 goods are priced individually (single-product firm pricing). The merger simulation in Section 5 uses multi-product firm pricing, which is more fitting of an oligopolistic market.

One interesting trend that also emerges is that the own price elasticity of healthy cereals tend to be less elastic. For example, General Mills Fiber One, Kashi Go Lean Crunch, Kellogg All Bran, Kellogg Special K, and Post Shredded Wheat have own price elasticities ranging from -1.22 to -1.95, which are much less elastic than the average of all 65 products in the sample. This yields higher PCMs for these products, indicating profitability tends to be above average for healthy cereal products. Column (7) of Figure 1.5 presents the percentage of cross-price elasticities that are positive. Regardless of whether two cereal products are marketed towards kids or adults, the cross price elasticity is expected to be greater than or equal to zero, indicating the products are either substitutes or unrelated, respectively. On average 60 percent of the cross price elasticities are positive. Finally, column (10) is the summation of the γ parameters on log price from equation (1.12). A zero sum of the γ parameters indicates homogeneity.

Figure 1.6 presents a 13×14 elasticity matrix that is a subset of the full 65×66 elasticity matrix. There are three product groupings: (i) Honey Nut Cheerios; (ii) Raisin Bran; and (3) Corn Flakes. In each product grouping there are several different sizes and a few different manufacturers including at least one store brand version of the good. A few interesting patterns emerge. As expected, intra-group cross price elasticities are high, indicating consumers substitute between different sizes of the same product as well as among different manufacturers of similar products. For example, when the price of General Mills Honey Nut Cheerios 20oz. increases by one percent, the quantity demanded of Honey Nut Cheerios 27oz. increases 0.77 percent, followed by the 14oz. SKU. When the price of General Mills 14oz. Cheerios increases, consumers tend to substitute more heavily to the 14oz. store brand version of Cheerios, and then towards larger size Cheerios SKUs. Another point of interest is the large negative cross-price elasticities between store brand products when the price of store brand Corn Flakes is increased. For example, the OLS estimates suggest that when the price of store brand Corn Flakes decreases by one percent, the quantity demanded of store brand Oat Rings increases by 0.41 percent, and similarly the quantity demanded of store brand Raisin Bran increases by 0.70 percent, indicating these goods are fairly strong complements. The CX estimator fixes this anomaly in Figure 1.8.

Results - IV with Hausman Instruments

Figures 1.5 and 1.7 present results of the AIDS model[27] estimated with IV (TSLS) using Hausman instruments. Included (assumed exogenous) regressors in the model are wages, merchandising and distribution variables, geography-retailer fixed effects, month fixed effects, and trend. The Hausman instrument for the price of product j sold in supermarket retail chain r located in geographic market g during week t is the weighted average[28] price of product j sold in all supermarkets in all geographies not g during week t.

The own price elasticities for the 65 products found in column (2) of Figure 1.5 indicate some disturbing trends. First, none of the elasticities are statistically significant, even at the 10 percent level. Furthermore, the own-price elasticities are very volatile, with several being inelastic or even positive. The average own-price elasticity is reasonable at -3.24, but the standard deviation is 8.92. Very little about the instrumental variables estimation with Hausman instruments is

[27]Supermarket and time fixed effects were included, as well as the merchandising variables, distribution, and a weekly time trend.

[28]I.e., the sum of the dollar sales of product j in all supermarkets located in all other geographic markets is divided by the sum of the pound (quantity) sales of product j in all supermarkets located in all other geographic markets.

reasonable. The implied PCM has a very large range, from -1379 percent to 3409 percent, and column (8) shows that the number of positive cross-price elasticities is about half of the OLS estimation results. Furthermore, the sum of the γ parameters are rarely near zero, indicating homogeneity would likely be rejected in a statistical test. Overall, it appears Hausman instruments have failed to recover the structural effect of price on quantity. These results are similar to Weinberg and Hosken's (2008) demand study of breakfast syrup.

The cross-price elasticities are presented in Figure 1.7. Again, rarely anything makes sense given the unreasonably large negative values for the cross price elasticities and absence of statistically significant results. However, the problem does not appear to stem from invalid or endogenous instruments, but rather instrument weakness. That is, the Hausman instrument does a poor job of predicting the endogenous prices in the first-stage of the 2SLS estimation routine. Figure 1.18 presents a histogram of the first stage F-Stat for the 65 endogenous price regressors that were instrumented with their corresponding Hausman instrument. All but one of the F-Stats were less than 10, with more than half of the first stage F-Stats less than six. According to Staiger and Stock (1997), the instrumental variables results obtained here are unreliable given the extremely poor fit in the first stage.

A comment is in order offering some perspective on why the price of a good at a particular supermarket located within some geographic market during a specific week would be weakly correlated with the average price of that same good in all other geographic markets during the same week. Recall that temporary price promotions or sales are a significant source of retail price variation for many goods found in supermarkets. For example, Hosken and Reiffen (2004b) find that 45 percent of the variation in retail cereal prices is due to temporary downward price promotions, which is supported here by Figure 1.15. The decision to set a good on sale by a supermarket located in a particular city arguably has little to do with the price of the same good set by supermarkets in all other cities, with the exception of common underlying costs to acquire the good from the manufacturer. Such common acquisition costs are the basis for the relevance of Hausman instruments, but common acquisition costs are likely to be a small source of the retail price variation for a good when prices and promotions are being determined on a weekly basis.

One may argue that manufacturers offer the same wholesale price or trade promotion deal to all supermarket chains regardless of geographic location, which ought to make inter-geography prices highly correlated. However, Besanko et al. (2005) find evidence that supermarkets exercise considerable discretion in how and when to offer temporary price promotions resulting from trade deals offered by manufacturers. Thus, even if supermarket chains across the country all face the same cost of acquiring a good, and even if this is a significant source of retail price variation, it is unlikely that supermarkets located in different cities are setting the good on sale at the same time. This would indicate low correlation or weak instruments when the scanner data frequency is at the weekly interval, which is precisely what the simulation conveys in Section 3.4.

Results - CX with Conditioning Instruments

The final set of estimates are for the CX estimator with conditioning instruments. The observed conditioning instruments that correspond to $W_{r,g,t}$ defined in Section 2.3 include the following. The natural log of the sum of quantity for a particular good in all geographic markets not g is used for $\overline{Q}_t$. For each geographic market, the percentage of the population that is white, the percentage that is age 18 or under, the percentage that is age 55 or over, the percentage that is female, and the percentage of the population that are households that are families with children are all employed for $DMG_{g,t}$. Monthly dummy variables are employed for FE_t, and geography-supermarket dummy variables interacted with lagged quantity for each good are employed for $FE_{r,g}$. The merchandising variables given by local advertising only, supermarket display only, and advertising and display only are employed for $PROM_{r,g,t}$.[29] For a particular supermarket retail chain r, the number of weeks that have elapsed since that supermarket last held a

[29]Note that local advertising only, supermarket display only, and advertising and display are three mutually exclusive variables.

temporary price promotion or sale for a particular good is used for $TPROM_{r,g,t}$. For a particular supermarket located within a geographic market, the number of weeks that have elapsed since the largest[30] supermarket located within the same geographic market last held a temporary price promotion for a particular good is used for $TPROM_{-r,g,t}$. The natural log of lagged prices (1-week through 8-week lags) for a particular product in a supermarket is used for $P_{r,g,-t}$. For a supermarket located in a geographic market, the natural log of lagged prices (1-week through 8-week lags) for a particular product in the largest supermarket located within the same geographic market is used for $P_{-r,g,-t}$. The natural log of lagged quantity (1-week through 8-week lags) for a particular product in a supermarket is used for $Q_{r,g,-t}$. For a supermarket retail chain located in a geographic market, the natural log of lagged quantity (1-week through 8-week lags) for a particular product in the largest supermarket located in the same geographic market is used for $Q_{-r,g,-t}$. Finally, also included are the natural log of the average weekly wage for a geographic market, a time trend, and a product's distribution intensity.

Figure 1.5 again presents elasticity estimates for the 65 cereal products. In general, the estimates based on CX found in column (3) are less elastic than the estimates based on OLS in column (1). The average elasticity is -2.53 for CX compared to -2.99 for the OLS estimation. The typical concern when estimating demand using the OLS estimator in the presence of endogeneity is that the own price elasticity is biased towards zero due to simultaneity. Here, however, the own price elasticity for the OLS estimator seems to be too elastic relative to the CX estimates. The reason is that traditional simultaneity concerns appear to not be an issue for differentiated products sold in supermarkets. The bias in this case is that the own price elasticity from OLS is too elastic because of dynamic considerations such as stockpiling. This phenomenon is pointed out by Hosken and Reiffen (2004a), and they suggest that pricing dynamics ought to be taken into consideration in the demand estimation process, which is precisely what the

[30] The definition of largest is based on total dollar-sales of all cereals. If supermarket account r is the largest supermarket within the geographic market, then the next-largest supermarket is used. This definition holds for other variables that are based on "largest" supermarket within a geographic market.

CX estimation framework does with the conditioning instruments that proxy for stockpiling behavior. Thus, the fact that most of the own price elasticities were reduced in column (3) of Figure 1.5 compared to column (1) is good indication that the CX framework based on the conditional exogeneity assumption, given by $P \perp U|W$, is doing its job with regard to stockpiling effects.

The implied PCMs in column (6) are slightly higher than PCMs in column (4), due to the relatively more inelastic results from the CX estimator. Finally, the number of positive cross price elasticities produced by the CX estimator, found in column (9), are often greater than the OLS estimator and much greater than the IV estimator with Hausman instruments.

Figure 1.8 presents the elasticity matrix for selected products based on the CX estimator. Compared to OLS in Figure 1.6, the CX estimates of cross price elasticities are smaller in magnitude, due to accounting for price and demand dynamics, but still comport with *a priori* notions of what the substitution patterns among characteristically similar products ought to be. One problem that was remedied with the conditioning instruments methodology is that many of the large negative cross price elasticities in Figure 1.6 have either turned positive or are no longer statistically significant in Figure 1.8. For example, consider again the cross price elasticity between store brand Corn Flakes and store brand Oat Rings. In Figure 1.6, the cross price elasticity between store brand Corn Flakes (price) and store brand Oat Rings (quantity) was -0.41 from the OLS estimation. In Figure 1.8, that cross price elasticity is reduced to -0.09. In Figure 1.6, the cross price elasticity between store brand Corn Flakes (price) and store brand Raisin Bran was -0.70. In Figure 1.8, that cross price elasticity is reduced to -0.01 and is no longer statistically significant.

The system of equations presented in Section 2.3 shed light on new sources of price endogeneity, such as stockpiling and competition among supermarkets, which are addressed with the CX estimator and conditioning instruments here. Compared to the OLS estimates, the CX estimates yield lower own price elasticities that result from accounting for dynamics with the conditioning instruments W. The CX estimates also produce more positive cross price elasticities, which comports with *a priori* notions of substitution patterns. The OLS and CX results are both reasonable, but the IV estimates with Hausman instruments suffer from a weak instruments problem. In research involving demand estimation for differentiated products with disaggregated, high frequency scanner data, the CX estimator with conditioning instruments proposed here or the simple OLS approach appear to yield the most reasonable results.

1.5 Merger Simulation

The Federal Trade Commission (FTC) granted early termination of the Hart-Scott-Rodino (HSR) waiting period in January 2008 for the Ralston-Post merger. Given the high market concentration in the cereal industry, it appears that regulators had sufficient evidence leading them to conclude that a second request for information was not necessary. One of those considerations may have been that a merger simulation produced post-merger results that were not indicative of enhanced market power in the cereal industry.

Given that the scanner data set employed in this chapter covers the three-year period leading up to the year of the Ralston-Post merger, it may well be what the merging parties produced to antitrust regulators for the HSR filing, or it may be what they would have produced if the FTC initiated a second request. Thus, an important question addressed here is what was it that the FTC saw that may have led to a favorable decision for the merging parties. To address the possibility that the FTC or merging parties' economists did not conduct any in-depth demand analysis and merger simulation for this particular transaction during the initial HSR waiting period, the question then becomes to what extent do the results of a merger simulation performed here comport with the FTC's decision to not further pursue the Ralston-Post merger. The answer, detailed in subsequent sections below, is that the simulated Ralston-Post acquisition results in simulated post-merger cereal industry prices that are only slightly above pre-merger prices.

The merger that transpired in 2008 between Ralston Corp. and Post is most closely approximated here by the merger of store brand products with Post products. According to Ralston Corp.'s 2007 SEC 10-K filing, they are the largest producer of store brand ready-to-eat cereals in the United States.[31] Therefore, the analysis presented here using all store brand products in place of Ralston is only an approximation to the true industry merger that transpired, as Ralston does not produce and set the prices of all store brand products.[32]

The 1997 Merger Guidelines, jointly set forth by the FTC and Department of Justice (DOJ), is a standard reference for assessing competitive effects regarding horizontal mergers. The Guidelines provide economists with a concise overview of the general ideas behind the antitrust agencies' merger reviews, but the particular details of what economists ought to do in analyzing a merger (or how to do it) are virtually nonexistent. The Guidelines do explicitly define the Herfindahl-Hirschman index (HHI), and suggest it as a starting point in determining the competitive effects of mergers. Figure 1.20 (available from author upon request) provides the HHI market concentration calculations for the merger of Post with Kellogg, General Mills, Quaker, and store brands. The underlying market shares based on pounds of cereal sold between 2005-2007 used to calculate the HHIs can be found in Figure 1.1. All the mergers between Post and one of the other firms yields post-merger HHIs and corresponding changes in HHI that "...raise significant competitive concerns" according to the Merger Guidelines.

A deeper examination of the substitution patterns of store brand cereals with Post cereals indicates there may be very little to be concerned about regarding the overlap of Ralston's and Post's portfolio of brands. The largest cross price elasticity between the two firms is for Post Raisin Bran 20oz. and store brand Raisin Bran 20oz., which is not surprising given that these two goods are nearly identical from a product characteristics perspective. According to the CX estimates, a one percent increase in the price of Post Raisin Bran leads to a 0.23 percent increase in the quantity purchased of store brand Raisin Bran. Given

[31] Ashenfelter and Hosken (2008) also point out that Ralston was the "largest private label cereal producer in the U.S." as of 1997, but it's unclear what their specific share of the store brands sales is.

[32] Of course, this inherently overstates the pre and post-merger HHI numbers. This also means that the post-merger simulated prices estimated here are an upper bound that are biased towards being too high.

that this appears to be the only major overlap of the two firms' products in the demand estimation sample, it is not difficult to see why regulators did not pursue the merger examination any further. Yet this still provides a convenient backdrop for performing a merger simulation to determine what might result in this industry post-merger.

Although merger simulation is not directly addressed in the Guidelines, it has become a popular exercise by economists at the U.S. antitrust agencies, as well as by academics and private sector economists often hired by the merging firms' attorneys. For example, Hausman, Leonard, and Zona (1994) examine mergers in the beer industry; Werden and Froeb (1994) examine mergers in the telecommunications industry; Nevo (2000a) examines mergers in the cereal industry; Peters (2006) examines mergers in the airline industry; and Weinberg and Hosken (2008) examine mergers in the motor oil and breakfast syrup industries. These studies show that merger analysis is often the impetus for demand system estimation, which can be thought of as the "first stage" of the merger simulation procedure. The following section details the ensuing "second stage" of the merger simulation using the AIDS estimates.

1.5.1 Industry Conduct

As in virtually all merger simulations, a differentiated Nash-Bertrand static model of oligopoly is assumed here with multiproduct firm pricing. The first order conditions (FOCs) resulting from the industry's profit maximization problem in the pre-merger state of the world are used to back out the implied marginal costs, which are then used along with the post-merger industry ownership structure to simulate what prices will be in the post-merger state of the world. Assuming constant marginal costs, the FOCs that result from profit maximization employing the AIDS are given by

$$\underset{(J\times 1)}{\begin{bmatrix} s_1 \\ s_2 \\ \vdots \\ s_J \end{bmatrix}} + \left(\underset{(J\times J)}{\mathbf{A}} \circ \underset{(J\times J)}{\begin{bmatrix} \eta_{11} & \eta_{12} & \cdots & \eta_{1J} \\ \eta_{21} & \eta_{22} & \cdots & \eta_{2J} \\ \vdots & & \ddots & \vdots \\ \eta_{J1} & \eta_{J2} & \cdots & \eta_{JJ} \end{bmatrix}}' \right) \underset{(J\times 1)}{\begin{bmatrix} (p_1 - c_1)s_1/p_1 \\ (p_2 - c_2)s_2/p_2 \\ \vdots \\ (p_J - c_J)s_J/p_J \end{bmatrix}} = \underset{(J\times 1)}{\begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}} \quad \text{or} \tag{1.14}$$

$$\mathbf{s} + (\mathbf{A} \circ \eta')(\psi \circ \mathbf{s}) = \mathbf{0}, \tag{1.15}$$

where s, p, and c denote budget share, price, and marginal cost, respectively, and $\circ$ denotes the Hadamard product (i.e., element-wise multiplication). Define $\mathbf{\Omega} \equiv \mathbf{A} \circ \eta'$, where $\mathbf{A}$ contains information on the ownership structure of the goods in the industry and η contains the responsiveness of consumers to price changes for those goods. For multiproduct firm pricing, $\mathbf{A}$ is comprised of elements $a_{jk} = 1$ if j, k are owned by the same firm, and $a_{jk} = 0$ otherwise.[33] The second $(J \times J)$ matrix comprising $\mathbf{\Omega}$ is a price elasticity matrix, η, with elements $\eta_{jk} = \frac{\partial q_j/q_j}{\partial p_k/p_k}$. It is important to note here that $\mathbf{s}$ and $\mathbf{\Omega}$ are functions of price, although this is not made explicit above purely for notational simplicity. The (implied) price-cost margins, $\psi_j \equiv (p_j - c_j)/p_j$, are solved for as

$$\psi = -(\mathbf{\Omega}^{-1}\mathbf{s}) \circ \widetilde{\mathbf{s}}, \tag{1.16}$$

where $\widetilde{\mathbf{w}}$ denotes the vector of inverses of the budget shares. Finally, the $(J \times 1)$ vector of (implied) marginal costs is given by

$$\mathbf{c} = [(\mathbf{\Omega}^{-1}\mathbf{s}) \circ \widetilde{\mathbf{s}} \circ \mathbf{p}] + \mathbf{p}, \tag{1.17}$$

where $\mathbf{p}$ denotes the vector of market prices of the J goods in the system.

The pre-merger marginal costs obtained above are assumed to be the same post-merger. This assumption can be augmented to incorporate potential efficiencies attained from the synergistic nature of the merger by reducing the marginal

[33]Note that for single product firm pricing, whereby each product's price is set by a separate firm, $\mathbf{A}$ is the identity matrix. For a joint ownership monopoly, whereby all products' prices are set by one firm in the industry, $\mathbf{A}$ is a matrix of ones.

costs of the merging firms. To determine the simulated post-merger prices, the FOCs from equation (1.14) or (1.15) are again utilized. First, note that the post-merger matrix $\mathbf{A}$ is repopulated to reflect the industry's product ownership structure where now the two merging firms comprise a new firm. Second, because the post-merger $\mathbf{\Omega}$ and $\mathbf{s}$ are functions of price, and given the functional form of AIDS, an analytical solution for the post-merger prices is not available. Thus, only a numerical solution for the post-merger version of equation (1.14) using the pre-merger demand estimates and implied marginal costs is possible. That is, a numerical search is performed to find the prices that make the post-merger version of (1.14) true; the resulting price vector is the post-merger simulated prices for the industry.

1.5.2 Merger Simulation Results

Four mergers are simulated: General Mills-Post, Kellogg-Post, Quaker-Post, and Store Brand-Post. The last merger most closely captures the Ralston-Post merger that actually transpired in 2008. Each of these mergers are simulated based on the AIDS estimates of Section 4.2, which includes OLS, IV with Hausman instruments, and CX with the conditioning instruments proposed earlier.

Figure 1.10 presents the pre-merger and post-merger averages of prices and price-cost margins (PCMs) for each of the manufacturers based on the OLS estimates. Figure 1.11 presents the pre-merger and post-merger results based on the IV estimates with Hausman instruments. Finally, Figure 1.12 presents simulation results based on the CX estimates with conditioning instruments. Where needed, prices and budget shares employed in calculating the elasticities and FOCs are the weighted averages or arithmetic averages across retailer-geography-week for each good. Pre-merger prices for each good are used as starting values for the numerical search of post-merger prices that solves the post-merger version of equation (1.15).

As expected, the simulation results based on the CX demand estimates found in Figure 1.12 are relatively more reasonable than the simulation results based on IV with Hausman instruments and the basic OLS estimates. This is not surprising since the merger simulation is heavily driven by the estimates of price effects, which were largely nonsensical based on the IV estimates. The relatively reasonable demand estimates produced by employing the conditioning instruments have translated into relatively reasonable post-merger simulated prices.

Focusing on the store brands-Post merger in Figure 1.12, the post-merger average prices for Post and store brands products are \$2.86 per pound and \$2.00 per pound, respectively, constituting a three percent and five percent price increase post-merger. The overall price increase across all cereal products is three percent, indicating that the store brands-Post merger appears to not harm consumers of cereal much on the basis of price. A Quaker-Post merger does not increase industry prices much either. However, a merger between General Mills-Post or Kellogg-Post results in post-merger simulated prices that are much higher. Kellogg and General Mills are much larger firms that maintain higher market shares and far more products than Quaker and store brands. Based on these results, it is not difficult to see why the FTC granted early termination of the HSR waiting period for the Ralston-Post merger in 2008 despite the highly concentrated industry. The simulated price increase resulting from the merger using pre-merger data are not much higher than pre-merger prices.

1.6 Conclusions

Confounding effects that lead to price endogeneity in demand estimation are better remedied with conditioning instruments proposed here for use in the conditional exogeneity framework of Chalak and White (2010). The typical remedy for addressing price endogeneity is to exploit the panel structure of the scanner data to obtain Hausman instruments, as suggested by Hausman (1997), and to estimate a traditional instrumental variables procedure. However, Hausman instruments are found to be weak in both empirical estimates of demand for cereal and simulation results. In both cases, the demand parameter estimates of interest are highly unreliable and unstable.

In contrast, the conditional exogeneity (CX) estimates are stable and reasonable. Although the conditioning instruments are also obtained by exploiting the panel structure of the scanner data, they are used in a much different way than standard Hausman instruments. The conditioning instruments are proxies for the confounding variables; thus, they need not be exogenous, yet they are instrumental in identifying the causal effect of price on demand.

When disaggregated, high frequency scanner data is available, the CX estimator and the conditioning instruments proposed here are recommended as an alternative estimation methodology for obtaining the price effects in demand systems that are typically of interest to IO and marketing researchers. In the application to ready-to-eat cereals, the CX framework yielded demand elasticities that were relatively less elastic, though still in the elastic range of -3.7 to -1.2 for the 65 cereal products in the sample. Also important is the fact that many of the large statistically significant cross-price elasticities that were estimated with basic OLS either turned positive or statistically insignificant with conditioning instruments. This result is attributed to incorporating the dynamic pricing and purchasing behavior of consumers, such as stockpiling, in the estimation procedure through the conditioning instruments.

Finally, the demand estimates based on the different identifying assumptions yield merger simulation results that are heavily driven by the underlying demand estimates. Simulations based on IV were nonsensical, while simulations based on the CX estimates were relatively reasonable. The post-merger simulated prices for the store brands-Post merger are only three percent higher than pre-merger cereal prices. This is consistent with the FTC's decision to not further investigate the Ralston-Post merger that transpired in 2008.

1.7 Tables and Figures

	Market Share - DS	Market Share - PS	Number of SKUs	Weighted Avg. Price Per Pound	% Above/Below Total Avg. Price	% Above Store Brand Avg. Price
Kellogg	37.9%	37.9%	55	$2.95	-0.1%	46.7%
General Mills	34.9%	30.3%	45	$3.40	15.2%	69.1%
Post	14.7%	16.0%	25	$2.71	-8.5%	34.4%
Quaker	7.6%	8.4%	13	$2.66	-10.2%	31.9%
Store Brands	5.0%	7.3%	12	$2.01	-31.9%	
Total	100.0%	100.0%	150	$2.96		

Notes
Source: IRI scanner data. Based on 121 city-supermarket chain combinations, 2005 through 2007.
Kellogg includes six Kashi brand SKUs.
Market Share - DS is based on dollar sales.
Market Share - PS is based on pound sales.

Figure 1.1: Market Share and Price

Variable	SKU Description	Mean	Median	StDev.	Kurt.	Skew.	Obs.
Average Price Per Pound	G MILLS CHEERIOS BOX 15OZ	3.52	3.51	0.72	2.45	-0.12	18,987
	KELLOGG FROSTED FLAKES BOX 20OZ	2.91	2.88	0.72	2.35	0.12	18,988
	POST HONEY BUNCHES OF OATS REG BOX 16OZ	2.99	2.98	0.67	2.17	-0.05	18,997
	QUAKER LIFE REGULAR BOX 21OZ	2.85	2.88	0.58	2.21	-0.02	17,781
	STR BDS RAISIN BRAN BOX 20OZ	1.80	1.82	0.34	2.75	0.20	18,524
Pound Sales	G MILLS CHEERIOS BOX 15OZ	2,411	1,187	4,229	100	7.6	18,987
	KELLOGG FROSTED FLAKES BOX 20OZ	1,905	837	3,966	100	7.6	18,988
	POST HONEY BUNCHES OF OATS REG BOX 16OZ	1,619	718	3,170	89	7.1	18,997
	QUAKER LIFE REGULAR BOX 21OZ	1,135	715	1,451	50	5.2	17,781
	STR BDS RAISIN BRAN BOX 20OZ	1,102	654	1,260	55	3.8	18,524
Local Advertising Only	G MILLS CHEERIOS BOX 15OZ	0.15	0.00	0.31	5.07	1.92	18,987
	KELLOGG FROSTED FLAKES BOX 20OZ	0.19	0.00	0.35	3.62	1.54	18,988
	POST HONEY BUNCHES OF OATS REG BOX 16OZ	0.15	0.00	0.32	4.84	1.88	18,997
	QUAKER LIFE REGULAR BOX 21OZ	0.08	0.00	0.24	11.37	3.15	17,781
	STR BDS RAISIN BRAN BOX 20OZ	0.07	0.00	0.24	12.04	3.26	18,524
In-Store Display Only	G MILLS CHEERIOS BOX 15OZ	0.05	0.00	0.15	21.02	4.07	18,987
	KELLOGG FROSTED FLAKES BOX 20OZ	0.06	0.00	0.16	16.44	3.55	18,988
	POST HONEY BUNCHES OF OATS REG BOX 16OZ	0.06	0.00	0.16	14.86	3.36	18,997
	QUAKER LIFE REGULAR BOX 21OZ	0.02	0.00	0.09	56.89	6.67	17,781
	STR BDS RAISIN BRAN BOX 20OZ	0.03	0.00	0.11	32.95	5.11	18,524
Local Ad and Display	G MILLS CHEERIOS BOX 15OZ	0.10	0.00	0.25	8.32	2.57	18,987
	KELLOGG FROSTED FLAKES BOX 20OZ	0.08	0.00	0.23	10.24	2.89	18,988
	POST HONEY BUNCHES OF OATS REG BOX 16OZ	0.07	0.00	0.21	11.35	3.07	18,997
	QUAKER LIFE REGULAR BOX 21OZ	0.03	0.00	0.14	30.12	5.18	17,781
	STR BDS RAISIN BRAN BOX 20OZ	0.01	0.00	0.09	64.43	7.58	18,524
Distribution	G MILLS CHEERIOS BOX 15OZ	0.99	1.00	0.03	274.34	-13.55	18,987
	KELLOGG FROSTED FLAKES BOX 20OZ	0.98	1.00	0.04	67.73	-6.00	18,988
	POST HONEY BUNCHES OF OATS REG BOX 16OZ	0.99	1.00	0.03	75.20	-6.19	18,997
	QUAKER LIFE REGULAR BOX 21OZ	0.98	1.00	0.05	27.36	-4.26	17,781
	STR BDS RAISIN BRAN BOX 20OZ	0.96	1.00	0.07	22.79	-3.76	18,524

Notes
Source: IRI scanner data. Based on 121 city-supermarket chain combinations, 157 weekly obs. 2005 through 2007.
The total possible observations in the sample is 121 retailers x 157 weeks = 18,997. Anything less indicates missing observations.
Merchandising and distribution variables are % all commodity value (on a 0 to 1 scale). See Little (1998) for a detailed explanation of these variables.
Percentiles available upon request from author.

Figure 1.2: Descriptive Stats

	Mean	Median	StDev.	Kurt.	Skew.	Obs.	Percentiles 1%	5%	25%	75%	95%	99%
Avg. Weekly Wage ($)	807.4	786.0	124.3	3.4	1.4	780	607.8	652.0	723.0	868.0	1037.1	1228.4
White (%)	77.0	79.0	10.2	-0.7	-0.4	65	54.3	59.2	69.5	85.2	90.9	93.1
Age <= 19 (%)	28.7	28.8	1.9	1.3	0.4	65	24.6	25.7	27.5	29.6	31.5	33.1
Age >=55 (%)	27.5	27.4	2.4	1.0	0.5	65	22.9	23.3	26.1	28.8	31.6	34.1
Female (%)	51.1	51.3	0.7	-0.3	-0.5	65	49.5	49.8	50.7	51.6	52.2	52.4
Households w/Children (%)	63.8	63.3	4.8	2.8	0.6	65	52.4	57.1	61.5	65.5	71.9	77.3

Notes
Source: Bureau of Labor Statistics, Quarterly Census of Employment and Wages (QCEW), and Census Bureau, 2000 Census Information.
Percentages refer to the percent of the MSA's population.
Avg. Weekly Wage is collected on a quarterly basis (65 MSAs x 12 quarters = 780 obs.).

Figure 1.3: Income and Demographics

	LHS Variable: Ln Price of General Mills Cheerios 15oz.	
RHS Variables	Coeff.	Std. Err.
Weeks Elapsed Since Last Promotion in the Supermarket	-0.002***	0.000
Weeks Elapsed Since Last Promotion in the Next-Largest Supermarket within the Same Geog. Market	0.000	0.000
Ln Average Price in All Other Geographic Markets ("Hausman Instrument")	0.345***	0.016
Lag 1 Week of Ln Price of the Same Supermarket	0.399***	0.015
Lag 2 Week of Ln Price of the Same Supermarket	-0.062***	0.016
Lag 3 Week of Ln Price of the Same Supermarket	0.047***	0.016
Lag 4 Week of Ln Price of the Same Supermarket	0.036**	0.016
Lag 5 Week of Ln Price of the Same Supermarket	0.040**	0.016
Lag 6 Week of Ln Price of the Same Supermarket	0.010	0.016
Lag 7 Week of Ln Price of the Same Supermarket	0.019	0.016
Lag 8 Week of Ln Price of the Same Supermarket	0.104***	0.015
Ln Price of the Next-Largest Supermarket within the Same Geographic Market	-0.021***	0.008
Lag 1 Week of Ln Price of the Next-Largest Supermarket within the Same Geographic Market	0.034**	0.016
Lag 2 Week of Ln Price of the Next-Largest Supermarket within the Same Geographic Market	0.015	0.017
Lag 3 Week of Ln Price of the Next-Largest Supermarket within the Same Geographic Market	0.005	0.018
Lag 4 Week of Ln Price of the Next-Largest Supermarket within the Same Geographic Market	0.013	0.017
Lag 5 Week of Ln Price of the Next-Largest Supermarket within the Same Geographic Market	-0.007	0.017
Lag 6 Week of Ln Price of the Next-Largest Supermarket within the Same Geographic Market	-0.015	0.017
Lag 7 Week of Ln Price of the Next-Largest Supermarket within the Same Geographic Market	0.031*	0.017
Lag 8 Week of Ln Price of the Next-Largest Supermarket within the Same Geographic Market	-0.018	0.016
Lag 1 Week of Ln Quantity of the Same Supermarket	0.027***	0.005
Lag 2 Week of Ln Quantity of the Same Supermarket	0.020***	0.005
Lag 3 Week of Ln Quantity of the Same Supermarket	0.009*	0.005
Lag 4 Week of Ln Quantity of the Same Supermarket	-0.002	0.005
Lag 5 Week of Ln Quantity of the Same Supermarket	0.011**	0.005
Lag 6 Week of Ln Quantity of the Same Supermarket	0.001	0.005
Lag 7 Week of Ln Quantity of the Same Supermarket	-0.002	0.005
Lag 8 Week of Ln Quantity of the Same Supermarket	0.018***	0.005
Lag 1 Week of Ln Quantity of the Next-Largest Supermarket within the Same Geographic Market	0.004	0.005
Lag 2 Week of Ln Quantity of the Next-Largest Supermarket within the Same Geographic Market	0.000	0.005
Lag 3 Week of Ln Quantity of the Next-Largest Supermarket within the Same Geographic Market	0.001	0.005
Lag 4 Week of Ln Quantity of the Next-Largest Supermarket within the Same Geographic Market	0.004	0.005
Lag 5 Week of Ln Quantity of the Next-Largest Supermarket within the Same Geographic Market	-0.002	0.005
Lag 6 Week of Ln Quantity of the Next-Largest Supermarket within the Same Geographic Market	-0.004	0.005
Lag 7 Week of Ln Quantity of the Next-Largest Supermarket within the Same Geographic Market	0.008	0.005
Lag 8 Week of Ln Quantity of the Next-Largest Supermarket within the Same Geographic Market	-0.006	0.005
Ln Average Weekly Wage	0.029	0.042
Percent of Population White	0.003***	0.001
Percent of Population Age <= 19	0.087***	0.012
Percent of Population Age >= 55	-0.011***	0.004
Percent of Population Female	0.022**	0.011
Percent of Households with Children	-0.032***	0.005
Weekly Time Trend	-0.001	0.001
Constant	-2.352***	0.690
Supermarket Fixed Effects	Yes	
Monthly Fixed Effects	Yes	
Observations	18005	
F-Stat	50.23	
R-Squared	0.35	

Notes
***, **, and * denotes 1%, 5%, and 10% statistical significance, respectively.
White (1980) heteroskedasticity-robust standard errors are reported.
Results for other four products available upon request from author.

Figure 1.4: Price Equations

	Own Price Elasticity			Implied Price-Cost Margin			% of Elasticities that are Positive			Sum of Price Parameters		
	OLS (1)	IV (2)	CXR (3)	OLS (4)	IV (5)	CX (6)	OLS (7)	IV (8)	CX (9)	OLS (10)	IV (11)	CX (12)
Min	-4.26	-40.00	-3.74	23%	-1379%	27%	45%	30%	52%	-0.02	-0.96	-0.03
Max	-1.22	23.68	-1.19	82%	3409%	84%	72%	45%	78%	0.03	1.53	0.03
Mean	-2.99	-3.24	-2.53	36%	4%	42%	60%	36%	64%	0.00	0.00	0.00
Median	-3.06	-2.65	-2.54	33%	21%	39%	59%	36%	64%	0.00	-0.02	0.00
StDev.	0.73	8.92	0.53	12%	501%	11%	5%	3%	5%	0.01	0.43	0.01

Notes
***, **, and * denote statistical significance as the 1%, 5%, and 10% level, respectively, based on White (1980) heteroskedasticity robust standard errors.
Implied Price-Cost Margin calculated as negative inverse of elasticity (Lerner index).
Sum of Price Parameters refers to adding up all the γ coefficients from the AIDS model. Sum = 0 indicates homogeneity.
Individual products' results available upon request from author.

Figure 1.5: Summary AIDS Results

	G MILLS CHEERIOS HONEY NUT BOX 14OZ	G MILLS CHEERIOS HONEY NUT BOX 20OZ	G MILLS CHEERIOS HONEY NUT BOX 27OZ	STR BDS OAT RINGS HONEY NUT BOX 14OZ	G MILLS TOTAL RAISIN BRAN BOX 18OZ	KELLOGG CRUNCH RAISIN BRAN BOX 18.2OZ	KELLOGG RAISIN BRAN BOX 20OZ	KELLOGG RAISIN BRAN BOX 25.5OZ	POST RAISIN BRAN BOX 20OZ	STR BDS RAISIN BRAN BOX 20OZ	KELLOGG CORN FLAKES BOX 12OZ	KELLOGG CORN FLAKES BOX 18OZ	STR BDS CORN FLAKES BOX 18OZ	Expenditure
G MILLS CHEERIOS HONEY NUT BOX 14OZ	-3.33***	0.32***	0.28***	0.10***	0.04	-0.03	0.15***	0.01	0.02	-0.09**	-0.02	-0.07**	0.00	1.02***
G MILLS CHEERIOS HONEY NUT BOX 20OZ	0.24***	-3.19***	0.29***	0.25***	-0.04	0.19***	0.13***	0.16***	-0.06**	0.02	-0.08**	-0.10***	-0.02	1.06***
G MILLS CHEERIOS HONEY NUT BOX 27OZ	0.33***	0.77***	-3.61***	-0.01	-0.01	0.11***	0.15***	-0.14***	-0.07**	-0.19***	-0.21***	-0.21***	-0.15***	1.07***
STR BDS OAT RINGS HONEY NUT BOX 14OZ	0.44***	0.04	0.13***	-2.36***	-0.06*	-0.10***	0.03	0.10***	0.09***	0.01	-0.13***	-0.04	-0.41***	0.96***
G MILLS TOTAL RAISIN BRAN BOX 18OZ	0.05*	0.24***	0.07**	0.16***	-2.63***	0.10***	0.15***	0.27***	-0.08**	-0.31***	0.01	-0.21***	-0.17***	1.00
KELLOGG CRUNCH RAISIN BRAN BOX 18.2OZ	0.00	-0.02	-0.11***	0.03	0.00	-3.14***	0.09***	0.34***	-0.07**	-0.05	0.09***	-0.13***	-0.26***	0.97***
KELLOGG RAISIN BRAN BOX 20OZ	0.14***	0.19***	0.06	0.09**	-0.15***	0.27***	-4.06***	0.44***	0.13***	0.25***	0.28***	0.05	0.25***	1.00
KELLOGG RAISIN BRAN BOX 25.5OZ	0.16***	0.23***	-0.06**	0.15***	-0.05	0.20***	0.42***	-4.08***	0.12***	0.03	0.31***	0.05*	-0.07**	0.98***
POST RAISIN BRAN BOX 20OZ	0.11**	-0.08	0.08	0.16***	0.21***	0.20***	0.19***	0.25***	-4.13***	0.38***	0.20***	0.09*	-0.14***	0.98**
STR BDS RAISIN BRAN BOX 20OZ	0.13***	0.01	0.05*	-0.18***	0.15***	0.03	0.28***	0.48***	0.22***	-2.08***	-0.01	-0.03	-0.70***	0.93***
KELLOGG CORN FLAKES BOX 12OZ	0.12***	-0.02	-0.01	-0.14***	0.15***	-0.24***	0.00	0.25***	0.15***	0.25***	-3.06***	0.56***	-0.03	0.97***
KELLOGG CORN FLAKES BOX 18OZ	0.09***	0.14***	-0.15***	0.03	0.01	-0.10***	-0.07***	0.17***	0.20***	0.19***	0.29***	-2.54***	0.23***	1.02***
STR BDS CORN FLAKES BOX 18OZ	0.05	-0.15***	-0.28***	-0.30***	-0.17***	0.02	-0.01	0.07	0.06	0.06	0.16***	0.32***	-2.28***	0.97***

Notes
***, **, and * denote statistical significance at the 1%, 5%, and 10% levels, respectively, using White (1980) heteroskedasticity robust standard errors.
Elasticity(i,j) refers to the percentage change in quantity of good i in response to percentage change in price of good j, where i and j refer to the row and column, respectively.
Estimation based on full 65-product model; only 13 products reported here.

Figure 1.6: AIDS OLS

	G MILLS CHEERIOS HONEY NUT BOX 14OZ	G MILLS CHEERIOS HONEY NUT BOX 20OZ	G MILLS CHEERIOS HONEY NUT BOX 27OZ	STR BDS OAT RINGS HONEY NUT BOX 14OZ	G MILLS TOTAL RAISIN BRAN BOX 18OZ	KELLOGG CRUNCH RAISIN BRAN BOX 18.2OZ	KELLOGG RAISIN BRAN BOX 20OZ	KELLOGG RAISIN BRAN BOX 25.5OZ	POST RAISIN BRAN BOX 20OZ	STR BDS RAISIN BRAN BOX 20OZ	KELLOGG CORN FLAKES BOX 12OZ	KELLOGG CORN FLAKES BOX 18OZ	STR BDS CORN FLAKES BOX 18OZ	Expenditure
G MILLS CHEERIOS HONEY NUT BOX 14OZ	-3.69	-0.59	0.23	-0.91	0.15	1.29	-0.35	-2.47	1.43	0.70	0.08	0.62	0.86	0.30
G MILLS CHEERIOS HONEY NUT BOX 20OZ	0.00	-5.53	0.53	1.31	-0.12	-0.29	0.45	-0.40	-0.69	0.05	-1.00	-0.60	-0.29	1.08
G MILLS CHEERIOS HONEY NUT BOX 27OZ	2.70	0.42	-2.39	-1.00	-0.79	-3.78	-1.38	7.44	0.08	-0.42	-2.99	4.02	-7.19	2.40
STR BDS OAT RINGS HONEY NUT BOX 14OZ	0.67	1.11	-0.23	0.19	-0.16	-4.00	-0.78	4.54	0.31	0.09	-2.20	3.01	-7.43	2.64
G MILLS TOTAL RAISIN BRAN BOX 18OZ	-0.01	0.00	0.31	0.34	-3.46	-0.42	0.56	0.87	0.41	-0.32	0.69	0.21	0.42	1.22
KELLOGG CRUNCH RAISIN BRAN BOX 18.2OZ	-1.98	-5.43	-0.85	1.49	0.00	-4.26	0.36	-2.02	1.84	-0.59	-0.26	-0.79	3.10	0.95
KELLOGG RAISIN BRAN BOX 20OZ	-2.55	-7.18	-1.50	2.99	-0.60	0.33	-3.25	-1.75	0.51	-0.57	-0.60	-2.60	5.36	1.01
KELLOGG RAISIN BRAN BOX 25.5OZ	-2.39	-11.46	0.53	2.83	-2.68	-5.87	-1.64	-1.29	0.67	1.12	0.70	6.80	-7.36	1.98
POST RAISIN BRAN BOX 20OZ	0.52	-3.64	3.83	-3.06	-0.87	0.80	-0.90	3.57	0.16	0.51	2.59	4.70	-3.75	2.37
STR BDS RAISIN BRAN BOX 20OZ	-0.21	-0.74	-0.43	1.17	0.55	-1.52	-0.17	2.49	0.45	-1.49	-1.06	0.97	-2.11	1.28
KELLOGG CORN FLAKES BOX 12OZ	-5.34	-17.62	0.36	5.75	-2.99	-5.99	-1.59	4.56	1.35	2.24	-5.61	5.45	-5.92	0.73
KELLOGG CORN FLAKES BOX 18OZ	-1.78	-3.23	-0.66	2.79	0.34	1.59	1.09	-3.02	1.02	-1.16	-0.40	-6.98	3.62	-0.30
STR BDS CORN FLAKES BOX 18OZ	1.25	4.74	1.32	-1.80	-0.77	-0.41	0.08	1.40	-1.17	-0.19	0.59	1.72	-4.01	1.36

Notes

***, **, and * denote statistical significance at the 1%, 5%, and 10% levels, respectively, using White (1980) heteroskedasticity robust standard errors.

Elasticity(i,j) refers to the percentage change in quantity of good i in response to percentage change in price of good j, where i and j refer to the row and column, respectively.

Estimation based on full 65-product model; only 13 products reported here.

Figure 1.7: AIDS IV w/Haus

	G MILLS CHEERIOS HONEY NUT BOX 14OZ	G MILLS CHEERIOS HONEY NUT BOX 20OZ	G MILLS CHEERIOS HONEY NUT BOX 27OZ	STR BDS OAT RINGS HONEY NUT BOX 14OZ	G MILLS TOTAL RAISIN BRAN BOX 18OZ	KELLOGG CRUNCH RAISIN BRAN BOX 18.2OZ	KELLOGG RAISIN BRAN BOX 20OZ	KELLOGG RAISIN BRAN BOX 25.5OZ	POST RAISIN BRAN BOX 20OZ	STR BDS RAISIN BRAN BOX 20OZ	KELLOGG CORN FLAKES BOX 12OZ	KELLOGG CORN FLAKES BOX 18OZ	STR BDS CORN FLAKES BOX 18OZ	Expenditure
G MILLS CHEERIOS HONEY NUT BOX 14OZ	-2.66***	0.13***	0.19***	0.05	0.01	-0.08**	0.08***	0.01	-0.02	-0.12***	-0.05	-0.01	-0.02	1.03***
G MILLS CHEERIOS HONEY NUT BOX 20OZ	0.13***	-2.54***	0.09***	0.04	0.08***	-0.01	0.05**	0.08***	0.02	0.04**	-0.09**	0.07***	0.03	0.77***
G MILLS CHEERIOS HONEY NUT BOX 27OZ	0.29***	0.26***	-3.44***	0.17***	0.11***	0.02	0.05**	-0.10***	-0.03**	-0.02	-0.06***	0.04	0.01	0.56***
STR BDS OAT RINGS HONEY NUT BOX 14OZ	0.37***	0.16***	0.15***	-2.07***	-0.05*	0.01	-0.06***	-0.05***	0.03	-0.06*	-0.04*	0.05*	-0.09**	0.40***
G MILLS TOTAL RAISIN BRAN BOX 18OZ	0.05***	0.12***	-0.01	0.04	-2.25***	0.00	0.00	0.04***	0.01	0.01	-0.04	0.05	-0.05***	0.54***
KELLOGG CRUNCH RAISIN BRAN BOX 18.2OZ	0.05	0.02	0.02	-0.09***	0.08***	-2.64***	0.04*	0.22***	0.02	0.08***	0.01	-0.13***	-0.01	0.66***
KELLOGG RAISIN BRAN BOX 20OZ	0.12***	-0.01	0.05	0.09**	-0.17***	0.19***	-3.42***	0.28***	0.19***	0.24***	0.15***	-0.01	-0.02	1.04***
KELLOGG RAISIN BRAN BOX 25.5OZ	0.13***	0.05	0.00	0.00	-0.04	0.18***	0.21***	-3.58***	0.05***	0.09*	0.11***	0.05*	-0.10**	0.84***
POST RAISIN BRAN BOX 20OZ	0.05	0.04	-0.02	0.00	0.05	0.06	0.26***	0.22***	-3.51***	0.28***	0.09**	-0.07*	-0.04	0.85
STR BDS RAISIN BRAN BOX 20OZ	0.07**	0.08***	0.00	0.04	0.01	0.10***	0.30***	0.25***	0.23***	-1.85***	0.07***	0.00	-0.01	0.33***
KELLOGG CORN FLAKES BOX 12OZ	0.08*	-0.01	-0.02	-0.01	0.01	-0.06***	0.02	0.03	-0.02	0.08***	-2.68***	0.14***	0.03	0.53***
KELLOGG CORN FLAKES BOX 18OZ	0.05	0.02	0.00	0.03	-0.09***	0.00	-0.03	0.03	0.07***	-0.05**	0.09***	-2.55***	0.03	0.68***
STR BDS CORN FLAKES BOX 18OZ	0.04	-0.04**	0.02	0.11	0.09**	0.00	-0.01*	-0.01	0.04	0.06***	0.17***	0.25***	-2.07***	0.24***

Notes

***, **, and * denote statistical significance at the 1%, 5%, and 10% levels, respectively, using White (1980) heteroskedasticity robust standard errors.

Elasticity(i,j) refers to the percentage change in quantity of good i in response to percentage change in price of good j, where i and j refer to the row and column, respectively.

Estimation based on full 65-product model; only 13 products reported here.

Figure 1.8: AIDS CX

First 10 Draws of the Simulation	Weak Correl. (a = 0.05, $\sigma = 10$) OLS	IV	CX	Moderate Correl. (a = 0.10, $\sigma = 2.5$) OLS	IV	CX	Strong Correl. (a = 0.15, $\sigma = 0.25$) OLS	IV	CX
1	-86.5	-875.6	-86.6	-88.4	-90.5	-92.9	-90.6	-97.2	-99.8
2	-85.0	-80.8	-85.6	-88.5	-107.1	-94.3	-91.1	-94.3	-99.7
3	-85.1	-12.7	-86.3	-88.3	-105.8	-94.1	-90.3	-99.9	-100.1
4	-86.1	-34.1	-86.1	-87.9	-98.7	-93.4	-91.4	-106.3	-100.0
5	-84.4	-112.9	-84.5	-87.7	-98.2	-92.8	-91.8	-102.8	-99.9
6	-85.9	-202.3	-86.3	-89.0	-107.2	-94.7	-91.7	-102.1	-100.0
7	-85.5	-271.4	-85.6	-86.3	-101.6	-92.0	-90.2	-102.5	-99.8
8	-85.9	-91.2	-86.1	-87.8	-100.6	-92.0	-91.7	-99.1	-99.8
9	-85.5	-4.8	-86.1	-87.9	-94.9	-92.9	-90.8	-98.8	-100.0
10	-86.4	-93.3	-86.6	-87.7	-91.0	-92.2	-91.4	-96.8	-100.0
Summary Statistics - Based on Full 10,000 Draws of the Simulation									
Mean β (True β = -100)	-85.5	-123.2	-86.1	-88.5	-102.4	-93.7	-91.2	-100.4	-99.9
Std Dev. β	0.70	2205.9	0.75	0.76	55.12	0.82	0.75	3.66	0.15
Mean R-Sqrd 1st Stage of IV		0.01			0.04			0.12	
Mean F-Stat 1st Stage of IV		1.85			9.38			28.54	
Mean R-Sqrd Cond Instr.			0.08			0.57			0.99
Mean F-Stat Cond Instr.			17.75			268.87			26823

Notes

Sample size is 200 weeks. LHS variable is quantity; RHS variable is Price. True Beta to recover is -100.

1st Stage of IV is linear regression of price (LHS var) on the Hausman instrument (RHS var).

The R-Sqrd and F-Stat reported for the conditioning instrument is from a linear regression of Expected Demand (LHS var) on the conditioning instrument (RHS var).

Figure 1.9: Simulation

Merger: General Mills/Post	Pre-Merger			Post-Merger			% Change	
	Price	MC	PCM	Price	MC	PCM	Price	PCM
General Mills	***$3.38***	***$1.60***	***53%***	***$4.31***	***$1.60***	***57%***	***27%***	***9%***
Kellogg	$3.17	$1.17	61%	$4.94	$1.17	64%	54%	10%
Post	***$2.80***	***$1.38***	***46%***	***$3.03***	***$1.38***	***59%***	***9%***	***33%***
Quaker	$2.85	$1.83	35%	$3.21	$1.83	43%	15%	24%
Store Brands	$1.92	$1.14	39%	$2.52	$1.14	42%	30%	-151%
Average of All 65 Products	$3.10	$1.38	54%	$4.22	$1.38	58%	35%	4%

Merger: Kellogg/Post	Pre-Merger			Post-Merger			% Change	
	Price	MC	PCM	Price	MC	PCM	Price	PCM
General Mills	$3.38	$1.60	53%	$3.61	$1.60	55%	7%	-12%
Kellogg	***$3.14***	***$1.24***	***59%***	***$3.76***	***$1.24***	***63%***	***18%***	***-2%***
Post	***$2.68***	***$1.76***	***33%***	***$3.70***	***$1.76***	***49%***	***39%***	***53%***
Quaker	$2.85	$1.83	35%	$2.74	$1.83	27%	-6%	-28%
Store Brands	$1.92	$1.14	39%	$2.22	$1.14	47%	17%	-58%
Average of All 65 Products	$3.08	$1.46	51%	$3.55	$1.46	55%	16%	-3%

Merger: Quaker/Post	Pre-Merger			Post-Merger			% Change	
	Price	MC	PCM	Price	MC	PCM	Price	PCM
General Mills	$3.38	$1.60	53%	$3.41	$1.60	54%	1%	-1%
Kellogg	$3.17	$1.17	61%	$3.27	$1.17	61%	2%	2%
Post	***$2.80***	***$1.38***	***46%***	***$3.22***	***$1.38***	***47%***	***13%***	***8%***
Quaker	***$2.85***	***$1.83***	***35%***	***$3.21***	***$1.83***	***41%***	***11%***	***13%***
Store Brands	$1.92	$1.14	39%	$2.04	$1.14	42%	6%	-20%
Average of All 65 Products	$3.10	$1.38	54%	$3.24	$1.38	54%	4%	1%

Merger: Store Brands/Post	Pre-Merger			Post-Merger			% Change	
	Price	MC	PCM	Price	MC	PCM	Price	PCM
General Mills	$3.38	$1.60	53%	$3.32	$1.60	53%	-1%	0%
Kellogg	$3.17	$1.17	61%	$3.24	$1.17	61%	3%	-1%
Post	***$2.80***	***$1.38***	***46%***	***$2.59***	***$1.38***	***48%***	***-6%***	***-2%***
Quaker	$2.85	$1.83	35%	$2.93	$1.83	38%	4%	9%
Store Brands	***$1.92***	***$1.14***	***39%***	***$2.33***	***$1.14***	***41%***	***22%***	***-96%***
Average of All 65 Products	$3.10	$1.38	54%	$3.11	$1.38	54%	1%	-6%

Figure 1.10: Merger Sim OLS

Merger: General Mills/Post	Pre-Merger			Post-Merger			% Change	
	Price	MC	PCM	Price	MC	PCM	Price	PCM
General Mills	***$3.38***	***$6.49***	***-3%***	***$21.67***	***$6.49***	***-86%***	***538%***	***-71%***
Kellogg	$3.17	$4.64	-37%	$8.72	$4.64	-24%	188%	-188%
Post	***$2.80***	***$2.86***	***29%***	***$5.64***	***$2.86***	***46%***	***88%***	***79%***
Quaker	$2.85	$2.72	7%	$9.59	$2.72	69%	218%	68%
Store Brands	$1.92	$1.58	27%	$10.36	$1.58	21%	519%	23%
Average of All 65 Products	$3.10	$4.77	-10%	$13.01	$4.77	-29%	320%	-85%

Merger: Kellogg/Post	Pre-Merger			Post-Merger			% Change	
	Price	MC	PCM	Price	MC	PCM	Price	PCM
General Mills	$3.38	$6.49	-3%	$13.50	$6.49	251%	309%	23%
Kellogg	***$3.17***	***$4.64***	***-37%***	***$13.75***	***$4.64***	***-23%***	***283%***	***-50%***
Post	***$2.80***	***$2.86***	***29%***	***$8.01***	***$2.86***	***2%***	***217%***	***4%***
Quaker	$2.85	$2.72	7%	$5.00	$2.72	42%	74%	-1%
Store Brands	$1.92	$1.58	27%	$4.01	$1.58	25%	128%	-2%
Average of All 65 Products	$3.10	$4.77	-10%	$11.86	$4.77	84%	264%	-11%

Merger: Quaker/Post	Pre-Merger			Post-Merger			% Change	
	Price	MC	PCM	Price	MC	PCM	Price	PCM
General Mills	$3.38	$6.49	-3%	$2.85	$6.49	-45%	-15%	30%
Kellogg	$3.17	$4.64	-37%	$2.88	$4.64	-52%	-8%	27%
Post	***$2.80***	***$2.86***	***29%***	***$3.11***	***$2.86***	***21%***	***11%***	***1%***
Quaker	***$2.85***	***$2.72***	***7%***	***$2.83***	***$2.72***	***7%***	***-1%***	***-8%***
Store Brands	$1.92	$1.58	27%	$1.92	$1.58	28%	0%	-6%
Average of All 65 Products	$3.10	$4.77	-10%	$2.84	$4.77	-32%	-7%	21%

Merger: Store Brands/Post	Pre-Merger			Post-Merger			% Change	
	Price	MC	PCM	Price	MC	PCM	Price	PCM
General Mills	$3.38	$6.49	-3%	$5.12	$6.49	99%	53%	-9%
Kellogg	$3.17	$4.64	-37%	$3.38	$4.64	-43%	5%	27%
Post	***$2.80***	***$2.86***	***29%***	***$3.58***	***$2.86***	***29%***	***30%***	***-32%***
Quaker	$2.85	$2.72	7%	$2.39	$2.72	-16%	-18%	-116%
Store Brands	***$1.92***	***$1.58***	***27%***	***$2.92***	***$1.58***	***23%***	***62%***	***11%***
Average of All 65 Products	$3.10	$4.77	-10%	$3.95	$4.77	22%	28%	-2%

Figure 1.11: Merger Sim IV w/Haus

Merger: General Mills/Post	Pre-Merger			Post-Merger			% Change	
	Price	MC	PCM	Price	MC	PCM	Price	PCM
General Mills	***$3.36***	***$1.28***	***61%***	***$6.54***	***$1.28***	***77%***	***96%***	***34%***
Kellogg	$3.14	$0.82	71%	$4.65	$0.82	60%	44%	-37%
Post	***$2.84***	***$1.15***	***54%***	***$5.05***	***$1.15***	***84%***	***92%***	***78%***
Quaker	$2.85	$1.80	36%	$3.90	$1.80	52%	36%	46%
Store Brands	$1.92	$0.75	60%	$2.93	$0.75	70%	54%	18%
Average of All 65 Products	$3.08	$1.06	63%	$5.19	$1.06	69%	68%	10%

Merger: Kellogg/Post	Pre-Merger			Post-Merger			% Change	
	Price	MC	PCM	Price	MC	PCM	Price	PCM
General Mills	$3.38	$1.19	64%	$3.45	$1.19	62%	3%	1%
Kellogg	***$3.18***	***$0.72***	***74%***	***$3.40***	***$0.72***	***69%***	***7%***	***8%***
Post	***$2.68***	***$1.55***	***41%***	***$5.17***	***$1.55***	***63%***	***86%***	***53%***
Quaker	$2.85	$1.80	36%	$2.50	$1.80	23%	-14%	-41%
Store Brands	$1.92	$0.75	60%	$2.09	$0.75	64%	11%	6%
Average of All 65 Products	$3.09	$1.05	64%	$3.52	$1.05	63%	15%	9%

Merger: Quaker/Post	Pre-Merger			Post-Merger			% Change	
	Price	MC	PCM	Price	MC	PCM	Price	PCM
General Mills	$3.38	$1.19	64%	$3.38	$1.19	64%	0%	0%
Kellogg	$3.17	$0.66	76%	$3.17	$0.66	76%	0%	0%
Post	***$2.80***	***$1.22***	***52%***	***$2.95***	***$1.22***	***52%***	***4%***	***1%***
Quaker	***$2.85***	***$1.80***	***36%***	***$2.89***	***$1.80***	***37%***	***2%***	***3%***
Store Brands	$1.92	$0.75	60%	$1.93	$0.75	60%	0%	0%
Average of All 65 Products	$3.10	$0.98	66%	$3.12	$0.98	66%	1%	0%

Merger: Store Brands/Post	Pre-Merger			Post-Merger			% Change	
	Price	MC	PCM	Price	MC	PCM	Price	PCM
General Mills	$3.38	$1.19	64%	$3.40	$1.19	65%	1%	0%
Kellogg	$3.17	$0.66	76%	$3.31	$0.66	77%	4%	2%
Post	***$2.80***	***$1.22***	***52%***	***$2.86***	***$1.22***	***55%***	***3%***	***9%***
Quaker	$2.85	$1.80	36%	$2.83	$1.80	36%	0%	1%
Store Brands	***$1.90***	***$0.93***	***51%***	***$2.00***	***$0.93***	***52%***	***5%***	***1%***
Average of All 65 Products	$3.12	$1.00	65%	$3.19	$1.00	67%	3%	2%

Figure 1.12: Merger Sim CX

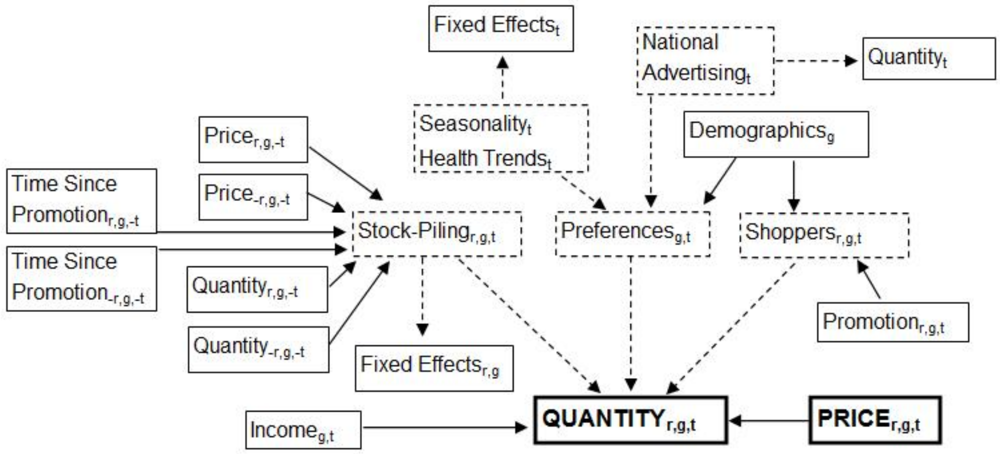


Figure 1.13: Path Diagram for Qty

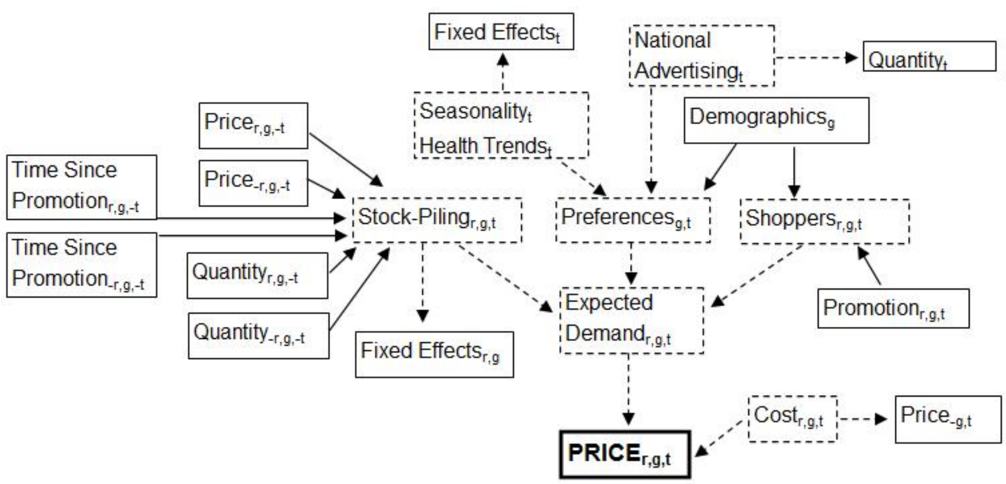
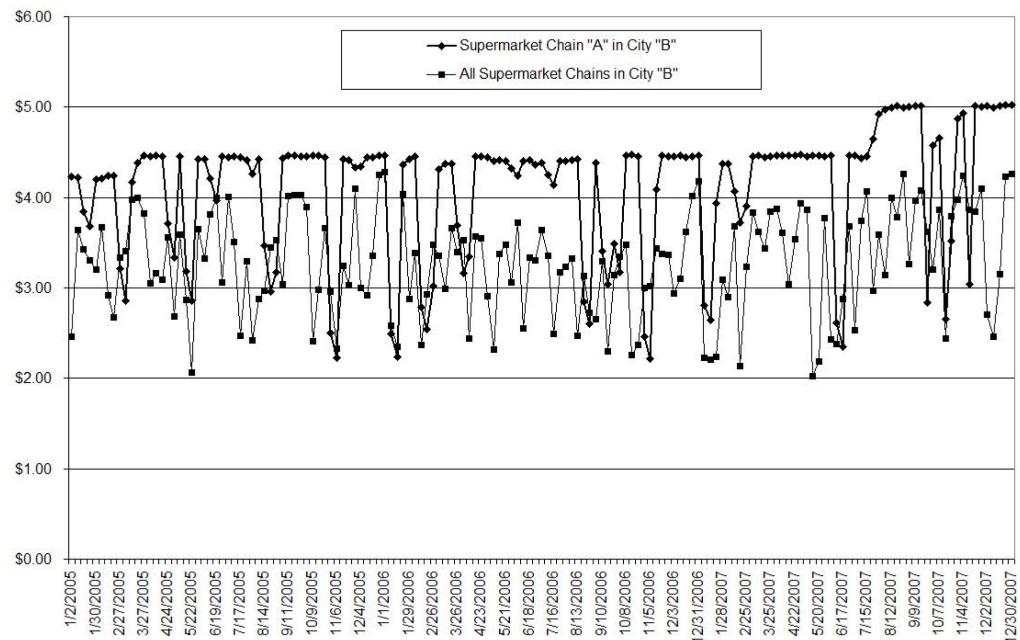
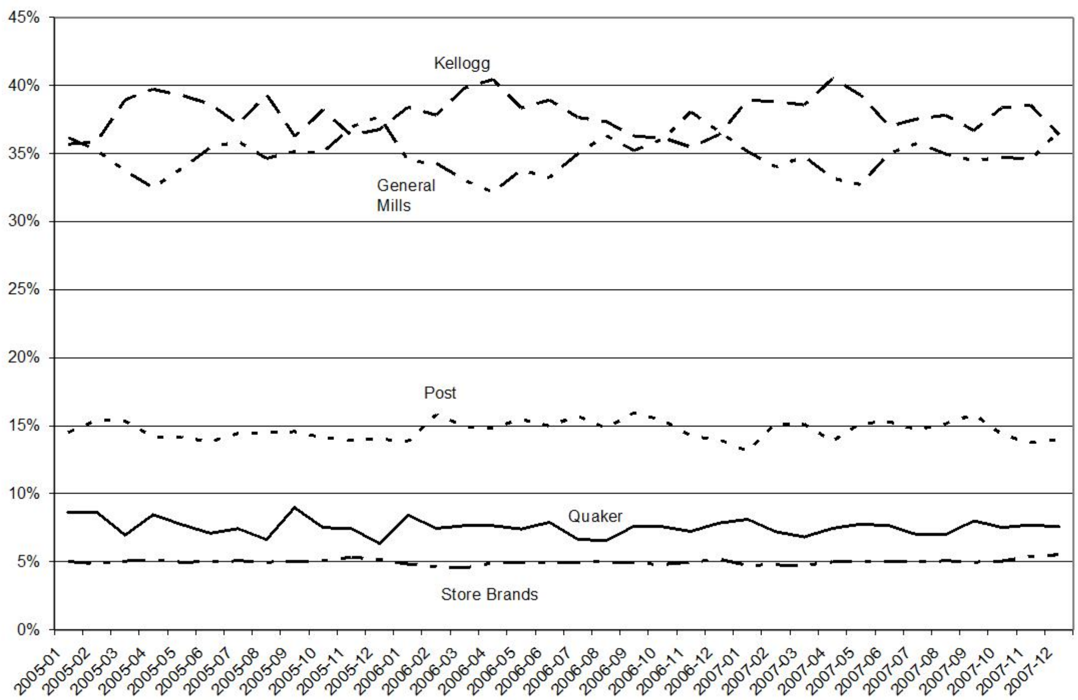


Figure 1.14: Path Diagram for Price



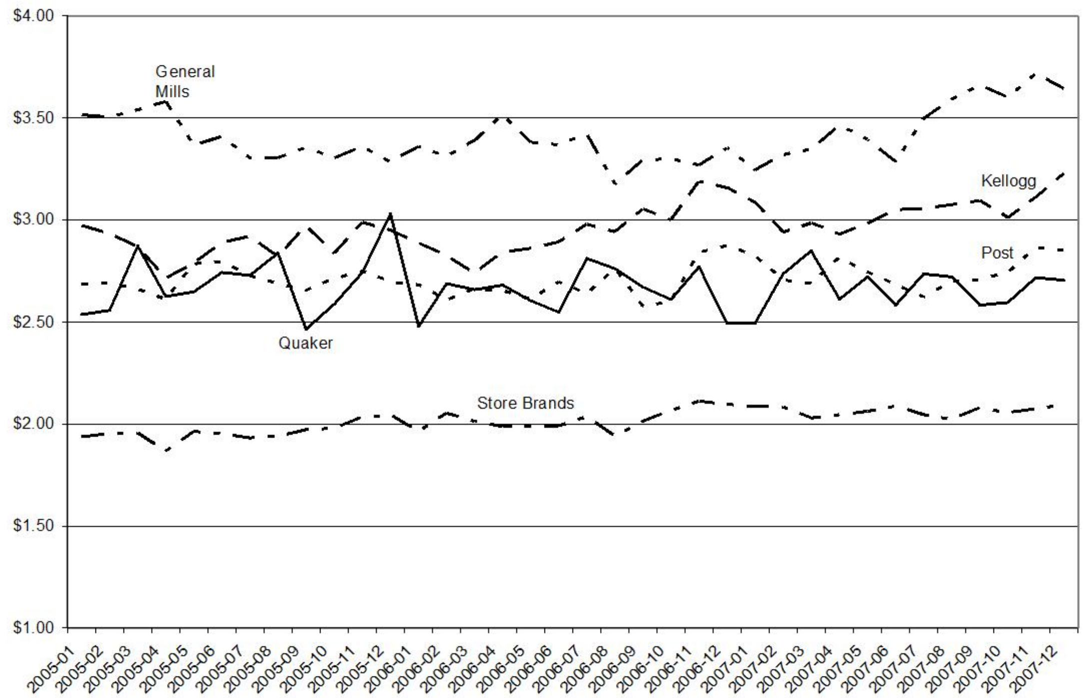
Source: IRI Scanner Data.

Figure 1.15: GM Cheerios - Price



Notes: IRI RTEC data aggregated to monthly frequency.
Market share is based on dollar sales for the 150 top-selling SKUs between 2005 and 2007.

Figure 1.16: Cereal Market Share



Notes: IRI RTEC data aggregated to monthly frequency.
Price is a weighted average dollar per pound measure based on the 150 top-selling SKUs between 2005 and 2007.

Figure 1.17: Cereal Price

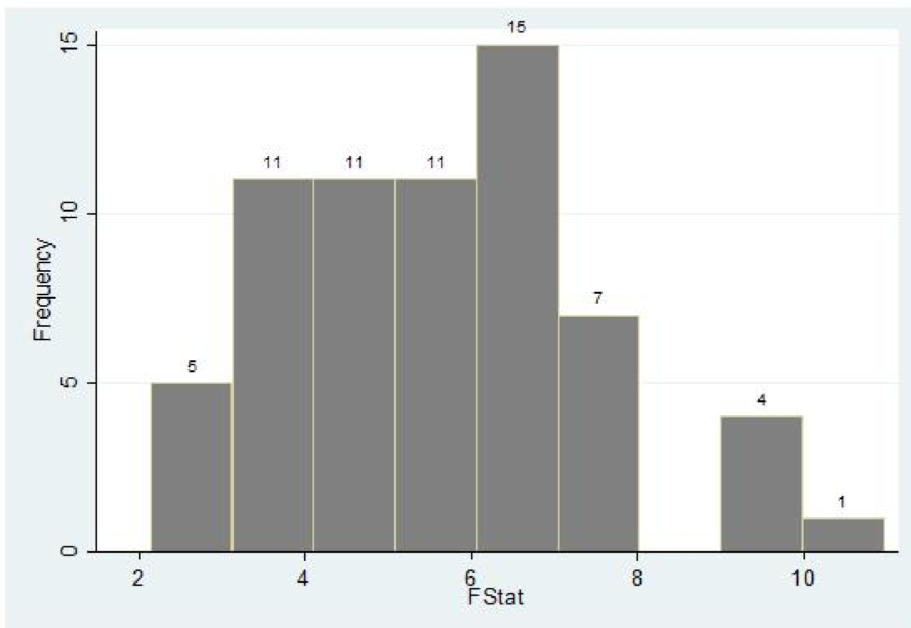


Figure 1.18: Distribution of F-Stat

Chapter 2

Identification of Price Effects in Discrete Choice Models of Demand for Differentiated Products

2.1 Introduction

The specification and estimation of demand for differentiated products is of central importance for research in industrial organization (IO) for assessing competitive effects and market conduct.[1] The primary concern is to obtain reliable econometric estimates of price effects, defined here as the effect of (exogenous) changes in market price on consumer demand. A conditional form of exogeneity is exploited here to estimate price effects in discrete choice models of demand, particularly the logit class.[2] Results based on the conditional exogeneity identification assumption are compared to results from the standard instrument exogeneity

[1]See, for example, Baker and Bresnahan (1985), Scheffman and Spiller (1996), Hausman (1997), and Rubinfeld (2000). See also Nevo (2000b), footnote 1, for a list of other important IO studies.

[2]For an analysis of price endogeneity in a non-discrete choice framework, see, for example, Hausman (1997) and Chapter 1, both of which estimate demand for ready-to-eat cereals using the almost ideal demand system.

assumption found in demand studies with market-level data, such as Berry, Levinsohn, and Pakes (1995), Hausman (1997), and Nevo (2001). To determine the efficacy of the two identification strategies, price elasticities and implied price-cost margins based on assumed market conduct are evaluated for ready-to-eat cereals.

When price exogeneity fails, the researcher is not able to identify the causal effect of price on demand, but rather is only able to obtain predictive estimates that may not be causally meaningful. The desired price effect is tantamount to having data available in which the prices of goods in a relevant product segment are randomly changed by reasonable increments across retailers, location, and time. In such an ideal world, even a basic regression of purchased quantities on market price ought to yield the causal price effect. Because unobservable (to the researcher) factors that drive market price and demand overlap, rendering market price endogenous, estimating causal effects as if one had data from an ideal randomized experiment becomes problematic. In attempting to address price endogeneity, research in IO and marketing appears to have gone no further than employing assumed exogenous instruments at some point in the identification framework. This chapter thus provides a new perspective by obviating the need to rely on outside exogenous instruments, such as rival's product characteristics and prices in other cities.

The conditional exogeneity identification framework employed here is contained in the research of White (2006), White and Chalak (2006), White and Chalak (2009), and Chalak and White (2010), to name a few. They provide a unified framework for estimating causal effects in structural systems that encompasses research in statistical theory (e.g., Dawid (1979)), artificial neural networks (e.g., Pearl (2000)), familiar econometric procedures (e.g., instrumental variables), the treatment effects literature (e.g., Rubin (1974) and Rosenbaum and Rubin (1983)), and extensions primarily employed in labor economics (e.g., Heckman, Ichimura, and Todd (1997), and Heckman and Vytacil (2005)). This framework is applied here to remedy the statistical dependence[3] between market price and

[3]Throughout this chapter, $\perp$ and $\not\perp$ denote independence (zero correlation) and dependence (non-zero correlation), respectively.

unobserved product characteristics or unobserved demand shocks.

When applied to demand, the conditional exogeneity assumption maintains that market price is uncorrelated with unobservable demand drivers conditional on a proposed set of conditioning instruments. That is, $p \perp \xi \,|\mathbf{w}$, where p, ξ, and $\mathbf{w}$, are market price, unobservable drivers of demand, and a set of conditioning instruments, respectively. In contrast, the standard ordinary least squares (OLS) assumption is $(p, \mathbf{x}) \perp \xi$, where $\mathbf{x}$ includes observed product characteristics. To the extent that $\mathbf{x}$ is contained in $\mathbf{w}$, conditional exogeneity $p \perp \xi \,|\mathbf{w}$ is a weaker condition, as it is implied by $(p, \mathbf{x}) \perp \xi$. Furthermore, consider standard instrumental variables (IV) requiring the assumption $(z, \mathbf{x}) \perp \xi$, where z is an assumed exogenous instrument for price. To the extent that z is not exogenous, and is more appropriately classified as a conditioning instrument described below, conditional exogeneity is an alternate identification strategy that is more appealing than standard exogeneity assumptions needed for OLS and IV. Specifically, the conditioning instruments need not be exogenous as is required of the instruments for identification in standard IV estimation, but the conditioning instruments are nevertheless instrumental in identifying the price effect of interest.

Properly choosing the conditioning instruments, $\mathbf{w}$, requires a structural set of equations that describes the data generation processes for market price and consumer demand. This system sheds light on the confounding variables that are potential sources of price endogeneity in the demand system. Drivers of or responses to unobservable variables common to both demand and market price are classified as conditioning instruments. The conditioning instruments proposed here to remedy product-level price endogeneity include observed product characteristics and the baseline price. The baseline price is defined as the average price of a good that varies across products but not markets. To the extent that unobserved product characteristics are being driven by some of the same factors that are driving product price, such as marginal cost (unobserved), the baseline price may be a viable response or predictive proxy for such factors. When employed as a conditioning instrument, empirical estimates show that just a few product characteristics and the baseline price can identify the marginal effect of market price on demand just as well as when including nearly 40 product characteristics or product fixed effects.

Importantly, this identification framework obviates the need to assume that observed product characteristics are uncorrelated with the error in the demand equation. This is a particularly useful contribution, especially in light of Berry's (1994) comment noting that the reliance on assuming exogeneity of observed product characteristics is a "defect" in nearly all applied discrete choice models of demand. By treating observed product characteristics as indirect drivers of utility, they need not be exogenous when employed as conditioning instruments. The conditioning instruments only serve to identify the structural or causal effect of market price on demand, which is the focus of this chapter. Coefficient estimates associated with the conditioning instruments may not be causally meaningful. This notion is analogous to the first stage coefficient estimates in the two stage least squares (TSLS) implementation of traditional IV. The coefficient estimates in the first stage of TSLS only serve to identify the coefficients in the second stage, but the first stage estimates do not contain any structural or causal meaning.

The conditional exogeneity identification scheme is also employed to remedy market-level sources of price endogeneity, such as unobserved demand shocks. Conditioning instruments are created by exploiting the panel structure of the data. For example, the number of weeks elapsed since a product was set on sale by a retailer is used as a driver of stockpiling behavior by consumers. The quantity of the good sold contemporaneously in all other cities is used as a response or predictive proxy for national advertising. The price of the good sold in other retailers within the same city is used to remedy confounding effects arising from retailers competing for the same pool of consumers. Moreover, other variables contained in the scanner data are classified as drivers of unobserved demand factors, making them appropriate conditioning instruments.

The empirical application undertaken here is estimating product-level demand elasticities and implied price-cost margins for ready-to-eat cereals. Demand estimation for cereal products has been the subject of econometric identification concerns in past research in IO; most notably, Hausman (1997) and Nevo (2001).

A new, highly detailed set of supermarket scanner data that differs from the two aforementioned papers is employed here. It contains supermarket chain-level detail within each city at a weekly frequency, as well as stock-keeping unit detail for each of the cereal products in the sample. Information on prices, quantities, and promotional variables are available.

To get a sense of the potential aggregation bias that may occur when researchers use scanner data that is aggregated at the city-quarter level, the disaggregated data here is artificially aggregated from the retailer-city-week detail up to the city-quarter level so that estimates from the two data sets may be compared. Finally, to thoroughly analyze and address the price endogeneity resulting from unobserved product characteristics, a highly detailed set of nearly 40 product characteristics for each brand of cereal is collected. This allows for a variety of experiments to be performed in order to test the efficacy of the instruments and identification assumption proposed here versus the instruments and identification assumptions proposed in previous research.

Estimates from traditional instrumental variables procedures employing instruments suggested by Berry (1994) and Hausman (1997) yield price elasticities that are inelastic and implied price-cost margins that are unreasonably high. In contrast, the conditional exogeneity estimates employing the conditioning instruments proposed here produce price elasticities and margins that make relatively more sense and are stable across specifications. Estimates from the aggregated city-quarter scanner data are inelastic relative to the same specification but using the disaggregated retailer-city-week data. This suggests that studies examining demand estimation with aggregated data are potentially overestimating price-cost margins due to aggregation bias. Moreover, it suggests that past research using aggregated data may have incorrectly attributed inelastic price effects to simultaneity rather than aggregation bias. Another important finding here is that omitting merchandising and local advertising variables leads to appreciably large changes in demand estimates, causing elasticities to be too high and margins to be too low due to confounding effects from stockpiling and consumer preferences.

Finally, during the sample period of the scanner data employed here, General Mills, one of five major ready-to-eat cereal manufacturers, increased the price on nearly all of its products by decreasing the amount of cereal in the box. This "hidden" price change primarily took effect during July 2007, and appears to be the result of an increase in General Mills' input costs (e.g. wheat). I estimate demand before and after this event, and determine that General Mills' price-cost margin increased by nearly four percent, indicating that a meaningful proportion of consumers did not notice the hidden price change. The phenomenon of hidden price changes in differentiated product markets is not uncommon, although there is very little research in IO that thoroughly examines it. Preliminary estimates are provided here, and a comprehensive empirical examination of hidden price changes can be found in Chapter 4.

The organization of this chapter is as follows. Section 2 presents the logit model of demand and outlines sources of price endogeneity. Section 3 examines the conditional exogeneity estimator based on the identifying assumption $p \perp \xi|\mathbf{w}$. Section 4 proposes a structural set of equations describing the data generation processes for demand and market price that sheds light on sources of confounding and provides guidance on selecting the conditioning instruments. Section 5 presents the ready-to-eat cereal scanner data and estimation results. Section 6 concludes.

2.2 Demand and Endogeneity

Throughout this chapter, the primary interest is to estimate the causal effect of market price on demand; obtaining causally meaningful estimates of effects from non-price product characteristics is not a direct concern here. In the sections to follow, the conditional logit discrete choice model of demand is examined in the context of market-level data. The conditional logit demand model suffers from the well-known independence of irrelevant alternatives (IIA) problem examined by McFadden (1974), Hausman and McFadden (1984), and Train (2003). However, despite its potential drawbacks, the logit model does offer a convenient, tractable framework to describe and apply the conditional exogeneity identification assumption. The price endogeneity concerns that are addressed here are not driven by the particular demand model, as they are generally applicable to most demand specifications, such as the almost ideal demand system (see Chapter 1). Extensions to probit, nested logit, and the random coefficients logit demand models is possible, but tractability and expositional convenience is lost.

2.2.1 Demand

The consumer's indirect utility is the starting point for estimating a structural demand system in a discrete choice framework. Let $k = 1, 2, ..., J$ index the goods in the choice set, let $m = 1, 2, ..., M$ index the markets in the sample, and let $i = 1, 2, ..., N_m$ index the consumers in market m.[4] Consumer i's indirect utility from consuming good k in market m takes the general form

$$u_{ikm} = d(p_{km}, \mathbf{x}_{km}, \xi_{km}) + \epsilon_{ikm}. \tag{2.1}$$

The variable p_{km} is the price of good k in market m, $\mathbf{x}_{km}$ is a vector of observed product characteristics, ξ_{km} represents unobserved (to the researcher) product characteristics and demand shocks, and ϵ_{ikm} is a mean-zero independently and identically distributed (iid) type-I extreme value disturbance term. Equation (3.1), is typically represented by the linear-separable parametric form $d(\cdot) = \mathbf{x}_{km}\beta_o + \alpha_o p_{km} + \xi_{km}$. In the context here, the coefficient α_o is the primary focus of interest, where the *a priori* expectation is $\partial u_{ikm}/\partial p_{km} = \alpha_o < 0$. The fact that the coefficients α_o and β_o are not indexed by i, which is the conditional logit model's primary point of differentiation with the random coefficients logit model, indicates that consumers are identical in their tastes. By convention and for ease of exposition, ϵ_{ikm} is assumed to be independent of $\mathbf{x}_{km}$, p_{km}, and ξ_{km}.

Typically in IO, consumer-level data is not available, but market-level data is. In the conditional logit framework here, this is accommodated by aggregating equation (3.1) to the market-level representative consumer, given by

$$u_{km} = d(p_{km}, \mathbf{x}_{km}, \xi_{km}) + \epsilon_{km}. \tag{2.2}$$

[4] Except where noted, a market is defined here as a supermarket chain-city-week combination.

The representative consumer in market m chooses good j over good k for all $k \neq j$ if $u_{jm} > u_{km}$ $\forall k \neq j$. The probability that this event occurs is given by $s_{jm} = \Pr(u_{jm} > u_{km}, \forall k \neq j)$, where s_{jm} is the aggregate choice probability or market share of good j. Defining the mean utility for good k as $\delta_{km} \equiv d(p_{km}, \mathbf{x}_{km}, \xi_{km})$, the market share for good j is given by $\Pr(\epsilon_{km} < \delta_{jm} - \delta_{km} + \epsilon_{jm}, \forall k \neq j)$. Given that ϵ_{km} is iid mean-zero type-I extreme value, the market share of good j in market m is given by the familiar choice probability

$$
\begin{aligned}
s_{jm} &= \int_{-\infty}^{\infty} \left[\prod_{k \neq j} e^{-e^{(\delta_{jm} - \delta_{km} + \epsilon_{jm})}} \right] dF(\epsilon_{jm}) \\
&= \frac{e^{d(p_{jm}, \mathbf{x}_{jm}, \xi_{jm})}}{\sum_{k=0}^{K} e^{d(p_{km}, \mathbf{x}_{km}, \xi_{km})}},
\end{aligned}
\tag{2.3}
$$

where the outside good is represented by $k = 0$. To arrive at the estimating equation, the mean utility for the outside good, $\delta_{0m} \equiv d(p_{0m}, \mathbf{x}_{0m}, \xi_{0m})$, is normalized to zero. Following Berry (1994), dividing s_{jm} by the share of the outside good and taking logs, the equation of interest is given by

$$\widetilde{s}_{jm} = d(p_{jm}, \mathbf{x}_{jm}, \xi_{jm}), \tag{2.4}$$

where $\widetilde{s}_{jm} \equiv \ln(s_{jm}) - \ln(s_{0m})$. The linear-separable parametric version is

$$\widetilde{s}_{jm} = \mathbf{x}_{jm}\beta_o + \alpha_o p_{jm} + \xi_{jm}, \tag{2.5}$$

which is a standard linear equation with JM observations. This provides a convenient framework to address price endogeneity, first by reviewing existing remedies, and then by examining the alternate remedy applied here, conditional exogeneity.

2.2.2 Price Endogeneity

The problem at hand with estimating α_o in equation (2.5) is the statistical dependence or correlation of p_{jm} and ξ_{jm}. Two sources of price endogeneity are classified here: product-level and market-level.

Product-Level Price Endogeneity

In order to separately examine product-level and market-level sources of price endogeneity, it is useful to decompose the error as $\xi_{jm} = \xi_j + \Delta\xi_{jm}$, where ξ_j captures product-level unobservable variables that do not vary across markets, while $\Delta\xi_{jm}$ captures unobservable market-level factors that do vary across markets. The case in which unobserved product characteristics, captured by ξ_j, are correlated with price is defined here as product-level price endogeneity, denoted $\xi_j \not\perp p_{jm}$.

Consider the following example. In the context of Berry, Levinsohn, and Pakes (1995),[5] suppose data for certain product characteristics are not observable, such as whether an automobile is equipped with power windows,[6] but that some characteristics are observable, such as horsepower. Moreover, suppose product characteristics, whether observed or unobserved, drive marginal cost. It is reasonable to believe that horsepower (observed by BLP) affects the marginal cost of producing an automobile through the increased cost of the engine, as do power features (unobserved by BLP). Marginal cost drives price through the profit maximization behavior of the price setter. Therefore, both observable and unobservable product characteristics are correlated with price through a common underlying variable, marginal cost. The path structure that results is $(\xi_j, \mathbf{x}_j) \to c_j \to p_{jm}$, where c_j denotes marginal cost, indicating that $(\xi_j, \mathbf{x}_j) \perp p_{jm}$ is not reasonable.

Because product quality is inherently difficult to quantify, it is a major source of product-level price endogeneity in discrete choice models of demand. For example, Trajtenberg (1989) obtains positive price effects for CT scanners. Trajtenberg's primary finding is that some portion of product quality remained unaccounted for in the logit model that was clearly correlated with price to the extent that products with higher unobserved quality values tend to have higher prices. From Monte Carlo simulations, Berry (1994) finds that estimation procedures that do not properly address price endogeneity due to unobserved product

[5]Hereafter, referred to as BLP (1995).

[6]BLP's study included the 1970s and 1980s, a time period during which power features were not necessarily standard in automobiles. Thus, power features may have been an important consideration in consumers' decisions.

characteristics, primarily product quality, yield misleading results; in particular, a resulting positive bias in estimates of price effects. As another example, the conditional logit demand model in BLP (1995) produces a price effect of -0.089 without accounting for price endogeneity, while their IV estimates yields a price effect of -0.216. They attribute this large change to a successful fix of the problem that automobiles with higher unmeasured quality components sell for higher prices.

Market-Level Price Endogeneity

The second source of price endogeneity is defined here as market-level price endogeneity, stemming from the correlation of price with $\Delta\xi_{jm}$, denoted by the condition $\Delta\xi_{jm} \not\perp p_{jm}$. For example, consider a shock that affects both price and demand in a city, such as advertising or trends in consumer preferences, which are typically not observed by the researcher. Lack of data on national advertising, or, more importantly, lack of data on the efficacy of the national advertising, may be grounds for $\Delta\xi_{jm} \perp p_{jm}$ to fail. Bresnahan's (1997) comment on Hausman's (1997) demand study of the cereal segment points to national advertising as a phenomenon that may affect both demand and price, thus rendering Hausman's identification scheme invalid. Nevo (2001) presents this caveat as well. However, for national advertising and prices to be correlated, price-setters would need credible information on the efficacy of the advertising with enough time to change prices. As noted by Hausman (1997) and Rubinfeld (2000), supermarkets typically set prices well in advance of effective dates. It is thus unclear that prices set by retailers are correlated with the national advertising campaigns of manufacturers. Regardless, the conditioning instruments of Section 4 address Bresnahan's concerns.

Section 4 proposes a structural system of equations describing the data generation processes for demand and price. The analysis of that section sheds light on the confounding variables that drive both product-level and market-level price endogeneity. Further market-level concerns that are addressed include pricing dynamics and price competition among retailers in the same city.

A Survey of Existing Remedies

When market level data are available in a panel structure, so that prices and market shares for each good vary across retailers and/or geographic locations and time, the most prevalent proposals to remedy estimation concerns due to price endogeneity include the use of product-level fixed effects and outside exogenous instruments. If the sample period is short enough, then it is reasonable to believe that certain product characteristics do not vary temporally or even geographically, while other product characteristics do. For example, in the context of cereal, physical product characteristics, such as a good's calories or fiber content, are typically constant across markets, while other product characteristics, such as local advertising by retailers, may vary considerably across markets. With the error decomposition $\xi_{jm} = \xi_j + \Delta\xi_{jm}$, the source of concern about $\xi_{jm} \not\perp p_{jm}$ due to ξ_j may be remedied with product-level fixed effects. See, for example, Besanko et al. (1998), Villas-Boas and Winer (1999), Nevo (2001), and Villas-Boas (2007).

The use of outside exogenous instruments to address price endogeneity concerns in demand estimation dates back to at least Wright's (1928) study of supply and demand elasticities. In the context of a discrete choice model of demand, rivals' product characteristics are proposed by Berry (1994) and BLP (1995) to instrument for price due to product-level endogeneity concerns. Rivals' product characteristics are exogenous if one is willing to assume that observed product characteristics are exogenous. Berry (1994) points out that the assumption that observed product characteristics are predetermined or exogenous is a "defect" that is common in most applications of discrete choice models. For example, if there is reason to believe that higher quality automobiles are more likely to have more horsepower, then unobservable and observable product characteristics may be correlated. Research since Berry (1994) that directly or indirectly relies on this assumption includes BLP (1995), Besanko et al. (1998), Villas-Boas and Winer (1999), Nevo (2001), Villas-Boas (2007), and Berry and Haile (2010). A case in which product characteristics, $\mathbf{x}_j$, are correlated with product-level unobservable variables, ξ_j, is when product characteristics determine a product's quality. For example, an automobile's horsepower influences its perceived quality.

The product characteristics of good k are a relevant instrument for the price of good j due to the profit-maximization conditions borne from the oligopoly structure of the industry. For simplicity, consider single product firms indexed by j, where k denotes firms not j. Following a differentiated Nash-Bertrand game, firm j's price reaction function for its good resulting from profit maximization is (omitting unobserved demand and cost shocks) $p_j = f_j(p_k, c_j)$, where p and c denote price and marginal cost, respectively. Further suppose that the marginal cost of production is a function of both observed and unobserved product characteristics, represented by $c_j = g_j(\mathbf{x}_j, \xi_j)$. Substituting the cost function into the price reaction function shows that the Nash equilibrium requires the mutual consistency condition

$$p_j = f_j\{f_k[p_j, g_k(\mathbf{x}_k, \xi_k)], g_j(\mathbf{x}_j, \xi_j)\}.$$

This shows that firm j's price is a function of the observed product characteristics of its rivals k. Thus, p_j and $\mathbf{x}_k$ are, at least theoretically, correlated.

Ackerberg and Crawford (2008) address the concern that observed product characteristics are not exogenous by assuming outside exogenous instruments are uncorrelated with the inside observed product characteristics. Thus, if the researcher can find instruments z_{jm} such that $(\xi_{jm}, \mathbf{x}_{jm}) \perp z_{jm}$, identifying α_o is possible even if the included observed product characteristics are endogenous. Practitioners typically agree that finding outside exogenous instruments is a daunting task. The fact that this approach requires the instrument for market price to now satisfy three conditions[7] instead of two makes the task of finding a valid instrument extremely difficult, especially if $\mathbf{x}_{jm}$ consists of several product characteristics.

The proposed remedies above are for price endogeneity due to product-level sources. Price endogeneity due to market-level sources (i.e., due to $\Delta\xi_{jm}$) has also been addressed with instrumental variables in past research, where again instrument exogeneity is needed for identification of price effects. Two problems

[7]That is, the instrument needs to be uncorrelated with the error, uncorrelated with the observed product characteristics, and correlated with price.

typically arise: instruments are not available and/or it is difficult to justify instrument exogeneity.

Cost shifters are a strong theoretical source of instruments for endogenous market price in demand studies. A problem inherent in the study of differentiated product markets is that the researcher must obtain at least as many outside exogenous variables to create the instrument set as there are products in the choice set. In differentiated product markets with many goods, such as cereal and automobiles, cost shifters quickly become an infeasible source of instruments.[8]

Because cost shifters are typically unobservable when the choice set of products is large, Hausman (1997) and Nevo (2001) rely on prices in other cities as a source of instruments. When the data set contains geographic location and time as its panel dimensions, instruments that have come to be widely known among IO practitioners as "Hausman instruments" have been used as proxies for the unobserved cost shifters. The main idea of Hausman instruments is that prices in other geographic markets in the panel can be employed as instruments for prices in a particular geographic market,[9] as underlying cost shocks ought to affect prices in all geographic locations. This scheme is adapted from Hausman and Taylor's (1981) development of instrumental variables for time invariant characteristics in a panel data setting, and later applied to demand estimation by Hausman, Leonard, and Zona (1994), Hausman (1997), and Nevo (2001), among others.

Hausman instruments provide a feasible solution even with several goods in the choice set. However, two assumptions are necessary in order to successfully employ Hausman instruments: (i) The unobserved shocks to product costs affect all geographic markets, and (ii) There are only geography-specific unobservable demand shocks and not nationwide demand shocks. The first assumption, captures the notion of instrument relevance, while the second assumption is necessary for instrument exogeneity. Bresnahan (1997) points out that the instrument exogeneity assumption underlying Hausman-type instruments may be unreasonable, particularly for nationally-branded consumer goods markets. The argument is

[8]Two studies that do employ cost shifters are Besanko et al. (1998) and Villas-Boas (2007).

[9]For example, the average price of good j in Chicago and Seattle is the instrument for price in Los Angeles.

that unobserved shocks affecting national demand, such as the effects of national advertising campaigns, render the instruments endogenous if prices in all cities are being set in response to the demand shock.

Finally, Villas-Boas and Winer (1999) employ lagged price as an exogenous instrument for market price. They point out that lagged prices are readily available in the data, thus circumventing the dimensionality problem of cost shifters much like Hausman instruments. While it is reasonable to believe that price in time period t and price in $t-1$ are correlated,[10] it is difficult to believe that price in $t-1$ is uncorrelated with demand in time t, thus rendering lagged price an invalid instrument. As the authors correctly caution, stockpiling could be a reason for correlation between current demand and lagged price, which makes lagged price an appropriate conditioning instrument as opposed to an exogenous instrument for IV. This point is addressed in Section 4.

In each of the IV solutions described above, the instrument exogeneity assumption is the source of identification of α_o, the marginal effect of market price on demand. In many circumstances, the instrument exogeneity assumption may not be reasonable. Moreover, it may be difficult to procure as many outside exogenous instruments as there are goods in the demand system. The next section presents an alternate identification framework that does not require outside instruments and product characteristics to be exogenous, and does not require as many instruments as there are goods in the choice set.

2.3 Conditional Exogeneity (CX)

The previous section describing standard instrumental variables estimation requires exogeneity of product characteristics and outside instruments, such as rival firms' product characteristics and prices in other cities. This potentially restrictive assumption can be overcome with an alternate identification strategy that only requires a conditional form of exogeneity. The conditional exogeneity (CX) assumption and resulting estimator examined here does not require exogenous in-

[10] See, e.g., Pesendorfer (2002) and Chapter 1.

struments, nor does it require that the number of instruments to be at least as many as the number of goods in the demand system. The conditional exogeneity assumption requires only that price is exogenous conditional on the conditioning instruments, $\mathbf{w}_{jm}$. That is, $p_{jm} \perp \xi_{jm}|\mathbf{w}_{jm}$, where several of the variables comprising $\mathbf{w}_{jm}$ are suggested in Section 4 based on the structural equations describing the confounding effects. In short, $\mathbf{w}_{jm}$ consists of observable drivers of or responses to the common unobservable determinants of demand and market price.

The conditional exogeneity assumption for estimating causal effects in structural systems of equations is examined extensively by White (2006), White and Chalak (2006), White and Chalak (2009), and Chalak and White (2010), among others. Their theory provides a unified framework for identifying causal effects in economic structures that encompasses research in statistical theory (e.g., Dawid (1979)), artificial neural networks (e.g., Pearl (2000)), familiar textbook econometric procedures (e.g., IV), the treatment effects literature (e.g., Rubin (1974) and Rosenbaum and Rubin (1983)), and extensions primarily employed in labor economics (e.g., Heckman, Ichimura, and Todd (1997), and Heckman and Vytacil (2005)). Because research in IO and marketing appears to have gone no further than assuming instrument and regressor exogeneity at some point in the identification framework, the conditional exogeneity assumption examined here is useful, as it provides a feasible alternative to standard IV assumptions. Much of the derivations below are based on Chalak and White (2010).[11]

In deriving the estimator resulting from the conditional exogeneity assumption, consider the linear-separable parametric equation[12]

$$\widetilde{s}_{jm} = \alpha_o p_{jm} + \xi_{jm}.$$

Assume a set of conditioning instruments, $\mathbf{w}_{jm} \equiv \left[w_{jm}^{(1)},\ w_{jm}^{(2)},\ ...,\ w_{jm}^{(l)}\ \right]'$, are available so that the conditional exogeneity assumption $p_{jm} \perp \xi_{jm}|\mathbf{w}_{jm}$ holds. Let

[11]Chalak and White (2010) refer to this case as conditionally exogenous causes given conditioning instruments.

[12]Note that observed product characteristics, $\mathbf{x}_{jm}$, are not included, as there is no reason to believe that they are exogenous, and, more importantly, they are classified as conditioning instruments ($\mathbf{w}_{jm}$) in Section 4.

$w_{jm}^{(1)} = 1\ \forall j, m$. The CX assumption implies the conditional covariance moment condition (suppressing subscripts)

$$E(p\xi|\mathbf{w}) - E(p|\mathbf{w})E(\xi|\mathbf{w}) = 0.$$

Rearranging gives $E\{p\xi - E(p|\mathbf{w})\xi\ |\mathbf{w}\} = 0$, or alternatively $E\{[p - E(p|\mathbf{w})]\xi\ |\mathbf{w}\} = 0$. Substituting the regression representation of $E(p|\mathbf{w})$, given by $E(p|\mathbf{w}) = E(p\mathbf{w}')E(\mathbf{w}\mathbf{w}')^{-1}\mathbf{w}$, into the preceding expression yields

$$E\{[p - E(p\mathbf{w}')E(\mathbf{w}\mathbf{w}')^{-1}\mathbf{w}]\xi\ |\mathbf{w}\} = 0.$$

Taking expectations on both sides, and by the law of iterated expectations, we have $E\{[p - E(p\mathbf{w}')E(\mathbf{w}\mathbf{w}')^{-1}\mathbf{w}]\xi\} = 0$. Substituting $\xi = \widetilde{s} - \alpha_o p$ yields $E\{[p - E(p\mathbf{w}')E(\mathbf{w}\mathbf{w}')^{-1}\mathbf{w}][\widetilde{s} - \alpha_o p]\} = 0$, and, assuming the appropriate rank conditions to ensure invertibility, α_o is given by

$$\begin{aligned}\alpha_o &= [E(pp') - E(p\mathbf{w}')E(\mathbf{w}\mathbf{w}')^{-1}E(\mathbf{w}p')]^{-1} \times \\ &\quad [E(p\widetilde{s}) - E(p\mathbf{w}')E(\mathbf{w}\mathbf{w}')^{-1}E(\mathbf{w}\widetilde{s})].\end{aligned}$$

The corresponding plug-in estimator is given by

$$\widehat{\alpha}^{CX} = [\mathbf{p}'\mathbf{p} - \mathbf{p}'\mathbf{W}(\mathbf{W}'\mathbf{W})^{-1}\mathbf{W}'\mathbf{p}]^{-1}[\mathbf{p}'\widetilde{\mathbf{s}} - \mathbf{p}'\mathbf{W}(\mathbf{W}'\mathbf{W})^{-1}\mathbf{W}'\widetilde{\mathbf{s}}], \qquad (2.6)$$

where $\mathbf{p}$ is a $(JM \times 1)$ vector of market price for the $j = 1, 2, ..., J$ inside goods, $\widetilde{\mathbf{s}}$ is a $(JM \times 1)$ vector of the difference of log share of good j and the outside good, and $\mathbf{W}$ is a $(JM \times l)$ matrix of conditioning instruments.

White and Chalak (2006) and Chalak and White (2010) provide several insights worth noting here. Rearranging equation (2.6) gives

$$\widehat{\alpha}^{CX} = [\mathbf{p}'(\mathbf{I} - \mathbf{W}(\mathbf{W}'\mathbf{W})^{-1}\mathbf{W}')\mathbf{p}]^{-1}[\mathbf{p}'(\mathbf{I} - \mathbf{W}(\mathbf{W}'\mathbf{W})^{-1}\mathbf{W}')\widetilde{\mathbf{s}}].$$

$\widehat{\alpha}^{CX}$ is the coefficient from a regression of $\widetilde{\mathbf{s}}$ on $\widehat{\nu}$, where $\widehat{\nu} \equiv (\mathbf{I} - \mathbf{W}(\mathbf{W}'\mathbf{W})^{-1}\mathbf{W}')\mathbf{p}$ are the residuals from the regression of $\mathbf{p}$ on $\mathbf{W}$. That is, $\widehat{\alpha}^{CX}$ can be estimated from a TSLS procedure using $\widehat{\nu}$ as the instrument for price. Note that substituting $\widehat{\nu}$ into $\widehat{\alpha}^{CX}$ gives $\widehat{\alpha}^{CX} = (\widehat{\nu}'\mathbf{p})^{-1}\widehat{\nu}'\widetilde{\mathbf{s}}$, which takes the form of a standard IV/TSLS estimator. Whereas standard IV requires finding exogenous instruments to address the endogeneity in market price by using the fitted price from the first stage, the conditional exogeneity identification framework requires finding conditioning instruments to address the endogeneity in market price by using the fitted residuals from the first stage. Moreover, $\widehat{\alpha}^{CX}$ can be estimated from a regression of $\widetilde{\mathbf{s}}$ on $\mathbf{p}$ and $\mathbf{W}$, where $\widehat{\alpha}^{CX}$ is the coefficient estimate associated with $\mathbf{p}$. Note here that the coefficient estimates *not* associated with market price only have a predictive meaning, rather than a causal or structural one. That is, the conditioning instruments only serve to recover α_o via the CX assumption $p_{jm} \perp \xi_{jm}|\mathbf{w}_{jm}$.

The conditional exogeneity framework is different from an exogenous regressors (OLS) framework, since $\xi_{jm} \perp (\mathbf{w}_{jm}, p_{jm})$ implies $p_{jm} \perp \xi_{jm}|\mathbf{w}_{jm}$, making the exogenous regressors assumption stronger than the conditional exogeneity assumption. Moreover, note that $p_{jm} \perp (\mathbf{w}_{jm}, \xi_{jm})$ also implies $p_{jm} \perp \xi_{jm}|\mathbf{w}_{jm}$, again making the conditional exogeneity assumption a weaker condition. See Dawid (1979). Finally, if the conditioning instruments are poorly chosen, such that $E(p_{jm}\mathbf{w}'_{jm}) = 0$, then the CX estimate from (2.6) reduces to the OLS estimate of $\widetilde{\mathbf{s}}$ on $\mathbf{p}$.

Importantly, this framework obviates the need to assume that observed product characteristics, which now compose $\mathbf{w}_{jm}$, be exogenous. This is a much needed departure from the standard assumption in discrete choice models of demand that Berry (1994) describes as a "defect." For comparison purposes, consider again the instrument exogeneity identification assumption $(\mathbf{x}_{jm}, z_{jm}) \perp \xi_{jm}$ needed for standard IV. This too is different from the conditional exogeneity assumption $p_{jm} \perp \xi_{jm}|\mathbf{w}_{jm}$. There is no need for the conditioning instruments to be uncorrelated with the unobserved determinants of demand, as is required for identification with standard instruments, as long as price is uncorrelated with the unobservable variables when conditioned on the instrument set, $\mathbf{w}_{jm}$. In fact, in this framework, one hopes that the conditioning instruments are correlated with the unobservable variables since the conditioning instruments are drivers of or responses to the unobservable confounding factors that are causing the endogeneity.

Finally, following theorem 4.2 in Chalak and White (2010), the distribution of $\widehat{\alpha}^{CX}$ is as follows. Assume conditions for identification of α_o hold with (suppressing jm subscripts) $E(p|\mathbf{w}) = E(p\mathbf{w}')[E(\mathbf{w}\mathbf{w}')]^{-1}\mathbf{w}$ where $E(p\mathbf{w}')$ and $E(\mathbf{w}\mathbf{w}')$ exist and are finite and $E(\mathbf{w}\mathbf{w}')$ is nonsingular. Further assume that

$$(JM)^{-1}\mathbf{p}'(\mathbf{I} - \mathbf{W}(\mathbf{W}'\mathbf{W})^{-1}\mathbf{W}')\mathbf{p} \xrightarrow{p} Q = E(pp') - E(p\mathbf{w}')[E(\mathbf{w}\mathbf{w}')]^{-1}E(\mathbf{w}p')$$

and

$$(JM)^{-\frac{1}{2}}\sum_{jm}[p_{jm} - E(p_{jm}|\mathbf{w}_{jm})]\xi_{jm} \xrightarrow{d} N(0, V),$$

where V is finite and positive definite. Then the asymptotic distribution of the CX estimator is

$$(JM)^{\frac{1}{2}}(\widehat{\alpha}^{CX} - \alpha_o) \xrightarrow{d} N(0, Q^{-1}VQ^{'-1}).$$

The last piece of information that is necessary to carry out the CX estimation is determining the appropriate conditioning instruments to employ in the estimator given by equation (2.6). This requires a structural set of equations, and is examined in detail in the next section.

2.4 A Structural System and Confounding Effects

A structural system of equations describing the data generation processes for market price and the accompanying consumption response is proposed in this section. The system of equations explicitly states the potential confounding variables causing price endogeneity problems that prevent identification of price effects in the demand model. Moreover, the system provides insight on the drivers of or responses to the confounding variables, which are classified as conditioning instruments in the conditional exogeneity identification scheme of the previous section. More formally, unobservable confounding variables have drivers and responses; to the extent that observable drivers and predictive proxies are available, these compose the conditioning instrument set $\mathbf{w}_{jm}$ that permits recovery of α_o from the conditional exogeneity assumption $p_{jm} \perp \xi_{jm}|\mathbf{w}_{jm}$.

2.4.1 Demand

Consider the demand equation derived from the conditional logit model given by equation (3.6)

$$\widetilde{s}_{jm} = d(p_{jm}, \xi_{jm}), \tag{2.7}$$

where $\widetilde{s}_{jm} \equiv \ln(s_{jm}) - \ln(s_{0m})$. Here, observed product characteristics, $\mathbf{x}_{jm}$, are omitted for three reasons. First, it may not reasonable to assume that observed product characteristics are exogenous. Second, observed product characteristics may not be a direct driver of consumer demand. Third, observed product characteristics are more appropriately classified as conditioning instruments in the analysis below. Consider the decomposition of ξ_{jm} giving the alternate expression

$$\widetilde{s}_{jm} = d(p_{jm}, \xi_j, \Delta\xi_{jm}). \tag{2.8}$$

In the framework presented here, ξ_j captures product-specific unobservable variables that directly drive consumers' preference or demand for good j, and $\Delta\xi_{jm}$ captures market-specific unobservable demand shocks that directly drives consumer demand for good j.

Consider the ready-to-eat cereal product segment that is the focus of the empirical application in the next section. Among other attributes, Nevo (2001) includes sugar content in his vector of observed product characteristics. It is not clear that consumers *directly* take sugar content into consideration when choosing a cereal, or if they are even aware of the sugar content of the products in their choice set. However, it is reasonable to believe that consumers take into direct consideration some broad notion of the sweetness of a cereal in their decision-making process, which is heavily driven by the sugar content of the product. Thus, sugar content is more appropriately classified as a driver of sweetness; that is, sweetness is captured by ξ_j while sugar content is a driver of ξ_j. Although it varies based on the application at hand, it seems more reasonable to treat observed physical product characteristics as indirect drivers of utility or demand, rather than direct drivers of utility or demand.[13]

[13] As a further example, consider the length times width ($l \times w$) used in BLP (1995) regarding

In the context just provided, ξ_j is defined here as the attribute appeal of the good. The attribute appeal of a good depends on the the product segment in question, but generally is a set of unobservable product-level factors that do not vary across markets and are likely determined by the manufacturer rather than retailers. In the context of cereals, ξ_j may consist of sweetness, texture or consistency, healthfulness, and product quality. In the context of automobiles, ξ_j may consist of performance, safety, size, design, and quality. Many physical product characteristics observed by the researcher drive the attribute appeal of the good, ξ_j. For example, a cereal's sugar content, fiber content, vitamin and mineral content, and fat content, among other observable characteristics, drives the sweetness and healthfulness of a cereal, and a cereal's density, shape, and ingredients may drive the texture or consistency of the cereal.

The attribute appeal of good j can be described by the structural equation

$$\xi_j = r_1(\mathbf{w}_j^d), \tag{2.9}$$

where $\mathbf{w}_j^d$ is a vector of variables that drives the attribute appeal of the good, including observed product characteristics of good j that do not vary by market, as well as other variables believed to be drivers of ξ_j. To the extent that an observable driver of ξ_j is not available due to data constraints, a predictive proxy for that component may be employed, where the predictive proxy is defined as a response to ξ_j. An example of a predictive proxy for the unobserved product quality of good j is the baseline price of good j, a measure of central tendency of the price of good j that is obtained by exploiting the panel structure of the data. The idea here is that observable and unobservable attributes of the good that are responsible for its quality drive the marginal cost of production of good j, which drives the baseline price, p_j. Thus, p_j is a response to product quality, and may thus act as a predictive proxy. The baseline price is examined in Section 4.3.1 below.

Next, consider the market-specific unobserved drivers of utility denoted by $\Delta\xi_{jm}$ in equation (2.8). Analogous to the previous setup for ξ_j, $\Delta\xi_{jm}$ is defined

---

automobile size. The variable $l \times w$ would be classified here as a driver of some broad notion of the size of an automobile, rather than a direct driver of utility.

here as the market appeal of the good. The market appeal of a good is a set of factors that varies across markets, and is more likely to be determined by retailers rather than manufacturers. $\Delta\xi_{jm}$ may consist of consumers' awareness of good j in market m[14] and consumers' propensity to stockpile. Observed non-physical product characteristics that vary across markets drive the market appeal of the good, $\Delta\xi_{jm}$. For example, the promotional intensity of a good by a retailer in a given city during a particular week, such as local feature advertising, is likely to drive consumers' awareness of a product.

The market appeal of a good can be described by the structural equation

$$\Delta\xi_{jm} = r_2(\Delta\mathbf{w}^d_{jm}), \tag{2.10}$$

where $\Delta\mathbf{w}^d_{jm}$ is a vector that includes drivers of the market appeal of good j that vary by market. An example of a driver of consumers' propensity to stockpile good j in a particular geographic market during time t is the number of weeks that has elapsed since the last promotion for good j in that market. See, for example, Pesendorfer (2002) and Chapter 1. Other direct drivers may include promotional intensity in the market. Again, to the extent that observable drivers are not available, observable responses to $\Delta\xi_{jm}$ may be employed as predictive proxies.

2.4.2 Price

The equation describing the data generation process for market price is given special attention here because of the importance placed on identifying the marginal effect of market price on demand. Consider the structural equation describing the market price of good j in market m

$$p_{jm} = s(\zeta_{jm}). \tag{2.11}$$

Researchers typically do not observe information on marginal costs and other factors that play a role in determining the price of a good. Equation (2.11) simply

[14] Recall that a market m can be thought of as a retailer-geography-time (rgt) combination in this chapter.

states that the price of good j in market m is driven by a variety of factors that are not observable by the researcher, such as marginal cost and other supply-side considerations. Again, unobservable drivers of market price can be separated into a part that is product-specific, ζ_j, and a part that is a market-specific deviation, $\Delta\zeta_{jm}$, resulting in

$$p_{jm} = s(\zeta_j,\ \Delta\zeta_{jm}). \tag{2.12}$$

The product-specific driver of price, ζ_j, captures manufacturer-level factors, such as the marginal cost to produce good j. ζ_j is described by

$$\zeta_j = h_1(\mathbf{w}_j^s), \tag{2.13}$$

where $\mathbf{w}_j^s$ is a vector of drivers of ζ_j that may include, among other variables, observed product characteristics of good j that do not vary by market. If observed drivers are not available, predictive proxies may be employed. For example, in the context of ready-to-eat cereals, $\mathbf{w}_j^s$ may include the sugar, fiber, and vitamin content, as well as the primary ingredient and shape or density of the cereal. Each of these observed product characteristics likely plays a role in determining the marginal cost of producing the good. Product quality is also a factor that affects the marginal cost of production, and should thus be included in $\mathbf{w}_j^s$. Because product quality is unobservable, and because the product characteristics that the researcher actually observes may not sufficiently capture product quality, the baseline price is again proposed here as a predictive proxy as it is a response to ζ_j. Of course product fixed effects may also be thought of as a predictive proxy for ζ_j; product fixed effects are addressed in the empirical application in Section 5.

Finally, the unobservable market factor, $\Delta\zeta_{jm}$, that causes the market price of good j to fluctuate from market to market, which are likely determined by the retailer is given by

$$\Delta\zeta_{jm} = h_2(\Delta\mathbf{w}_{jm}^s), \tag{2.14}$$

where $\Delta\mathbf{w}_{jm}^s$ is a vector that includes drivers of market-specific variations in the price of good j. Studies in IO and marketing suggest that $\Delta\zeta_{jm}$ may capture temporary price promotions to exploit consumer types, competition among retailers for consumers, coordination with local advertising, and demographic considerations regarding the underlying consumer base in the market. See Pesendorfer (2002), Hosken and Reiffen (2004a,b), Besanko et al. (2005), and Chintagunta (2002). These unobservable factors, $\Delta\zeta_{jm}$, are driven by $\Delta\mathbf{w}^s_{jm}$, which may include the number of weeks elapsed since the last sale, competitors' prices, local advertising variables, and demographic variables, such as income, race, and family composition.

2.4.3 Confounding Effects and Conditioning Instruments

The import of equations (2.8) through (2.14) is that they provide a structural framework to determine the confounding effects that prevent identification. The primary setback in estimating price effects in equation (2.8) is that unobservable factors that influence demand also influence price. The confounding effects preventing identification come in two flavors: product-level ($\mathbf{w}_j$) and market-level ($\Delta\mathbf{w}_{jm}$). Taken together, $\mathbf{w}_{jm} \equiv \{\mathbf{w}_j, \Delta\mathbf{w}_{jm}\}$ is the set of conditioning instruments needed for identification of $\partial\widetilde{s}_{jm}/\partial p_{jm}$ with the CX assumption $p_{jm} \perp \xi_{jm}|\mathbf{w}_{jm}$.

Product-Level Confounding Effects

Consider again product-level price endogeneity. In the structural framework presented, product quality is a component of ξ_j, the attribute appeal of good j, in equation (2.8). Factors that drive product quality, such as certain observed product characteristics, are contained in $\mathbf{w}^d_j$ in equation (2.9) and $\mathbf{w}^s_j$ in equation (2.13). Equations (2.8), (2.9), (2.12), and (2.13) show the link between unobserved product quality, product characteristics, market price, and consumer demand, and thus outline the confounding effects preventing identification of $\partial\widetilde{s}_{jm}/\partial p_{jm}$. That is, $p_{jm} \not\perp \xi_{jm}$ because $\mathbf{w}_j$, the common elements of $\mathbf{w}^d_j$ and $\mathbf{w}^s_j$, is driving both ξ_{jm} and ζ_{jm}.

In the ready-to-eat cereal product segment, for example, product quality is a determinant of consumer demand, and certain product attributes, such as the vitamin/mineral content or type of ingredients composing the cereal, drive the marginal cost of production, which in turn determines the price of the good. Simply put, higher quality cereals cost more to produce, thus have higher price, but at the same time the quality of the cereal affects consumers' preferences, which is the source of confounding that prevents causal identification of $\partial\widetilde{s}_{jm}/\partial p_{jm}$ in equation (2.8).

More formally, let $\mathbf{w}_j$ be defined as common elements of $\mathbf{w}_j^d$ and $\mathbf{w}_j^s$. That is, $\mathbf{w}_j$ is a vector that contains observable variables that drive ξ_j and ζ_j or predictive proxies for them. The elements of $\mathbf{w}_j$ are conditioning instruments defined by Chalak and White (2010). In the estimating equation given by (2.8), the conditional exogeneity assumption $p_{jm} \perp \xi_j|\mathbf{w}_j$ is sufficient for identification of $\partial\widetilde{s}_{jm}/\partial p_{jm}$ in the absence of market-level price endogeneity. In the linear-separable parametric version of equation (2.8), given by $\widetilde{s}_{jm} = \alpha_o p_{jm} + \xi_j + \Delta\xi_{jm}$, the conditional exogeneity condition $p_{jm} \perp \xi_j|\mathbf{w}_j$ is sufficient for identification of α_o in the absence of market-level confounding variables.

Several candidates for $\mathbf{w}_j$ are proposed here. A starting point is to determine the observed product characteristics that drive both ξ_j and ζ_j. A plausible case is that all observed product characteristics drive both ξ_j and ζ_j. For example, sugar content, fiber, density, ingredients, vitamin/mineral content, shape, among other attributes, all drive the attribute appeal of the good, ξ_j, such as sweetness, texture, healthfulness, and quality. On the pricing side, all the listed product characteristics may also drive the marginal cost of production, captured by ζ_j, which drives market price, p_{jm}. Therefore, one starting point is that all observed product characteristics ought to comprise the product-specific conditioning instruments, $\mathbf{w}_j$.

What if the researcher has not collected a sufficiently detailed set of product characteristics? This setback may be well-remedied with a variable that acts as a predictive proxy for the information contained in a comprehensive set of product characteristics. The predictive proxy proposed here is the baseline price of a good. The baseline price of a good is defined as a measure of central tendency for the price of good j, $p_j = M^{-1}\sum_m p_{jm}$. That is, the overall or long-run average price of good j ought to adequately capture the information contained in the components of ξ_j and ζ_j according to the structural equations describing demand and price in the previous sections. Consider decomposing the price of good j in market m into two parts: $p_{jm} = p_j + \Delta p_{jm}$, where p_j does not vary across markets but Δp_{jm} does. Averaging over markets yields

$$M^{-1} \sum_m p_{jm} = M^{-1} \sum_m p_j + M^{-1} \sum_m \Delta p_{jm},$$

and rearranging gives $p_j = M^{-1} \sum_m p_{jm}$. The baseline price captures inter-product variation in product characteristics, both observed and unobserved. Because the attribute appeal of a good (ξ_j), such as sweetness, texture, healthfulness, and quality, is driven by product characteristics, and because product characteristics drive marginal cost (captured by ζ_j), the baseline price ought to be a reasonable conditioning instrument as it captures approximately the information contained in the complete set of product characteristics and/or product-level fixed effects.

Moreover, both observed product characteristics and the baseline price can together comprise the conditioning instrument, $\mathbf{w}_j$. To the extent that the available observable product characteristics do not adequately explain ξ_j and ζ_j, the baseline price may contain additional information contributing to identification. Similarly, if the baseline price does not adequately explain ξ_j and ζ_j, then certain observed product characteristics may prove useful in identifying the marginal impact of market price on demand. In fact, this appears to be the case as detailed in Section 5. When the baseline price and a few key observed product characteristics are used together as the conditioning instrument set, $\mathbf{w}_j$, the estimate of α_o is approximately the same as an estimate of α_o when product-specific dummy variables are used to control for ξ_j. See Figure 2.5.

Market-Level Confounding Effects

The second source of price endogeneity to consider is at the market level. The concern here is that market-level unobservable factors that determine consumers' demand for good j also influence the price of good j in that market. A market m is defined to be a retailer-geography-time (rgt) combination. Factors that affect market demand for a good, such as brand awareness and consumer preferences, are contained in $\Delta\xi_{jm}$ in equations (2.8) and (2.10). To the extent that these factors also drive $\Delta\zeta_{jm}$ in equations (2.12) and (2.14), they are the market-level confounding effects that prevent identification of $\partial\widetilde{s}_{jm}/\partial p_{jm}$.

Let $\Delta\mathbf{w}_{jm}$ be defined as common elements of $\Delta\mathbf{w}_{jm}^d$ and $\Delta\mathbf{w}_{jm}^s$. That is, $\Delta\mathbf{w}_{jm}$ is a vector that contains observable variables that drive both $\Delta\xi_{jm}$ and $\Delta\zeta_{jm}$ or predictive proxies for them if observable drivers are not available. The elements of $\Delta\mathbf{w}_{jm}$ are candidates for conditioning instruments, but this time for addressing market-level sources of price endogeneity. In estimating equation (2.8), the conditional exogeneity assumption $p_{jm} \perp \Delta\xi_{jm}|\Delta\mathbf{w}_{jm}$ is sufficient for identification of $\partial\widetilde{s}_{jm}/\partial p_{jm}$ in the absence of product-level price endogeneity. In the linear-separable parametric version of equation (2.8), given by $\widetilde{s}_{jm} = \alpha_o p_{jm} + \xi_j + \Delta\xi_{jm}$, $p_{jm} \perp \Delta\xi_{jm}|\Delta\mathbf{w}_{jm}$ is sufficient for identification of α_o in the absence of product-level confounding variables.

Three potential sources of market-level price endogeneity are considered. First, consumers' preferences in a particular market are expected to depend on the demographics of geographic location g, local advertising, national advertising, seasonal demand shocks, and consumption trends. Demographics can affect consumers' preferences through age distribution, family composition, and race. Health trends and advertising may affect consumers' preferences as well.[15] The extent to which these variables are observable or not by the researcher is dictated by the available data set. In the context of scanner data, local advertising variables may be available, but data related to health trends and national advertising are typically not observed, although they can be proxied with responses obtained by exploiting the panel structure of the scanner data. For example, if national advertising, health trends, or any other aggregate demand shock affect consumers' preferences, then the quantity purchased of the product segment (or individual good) ought to be affected in all geographic locations. The panel structure of the data, where market m is composed of a retailer-geography-time (rgt) combina-

[15] For example, Ippolito and Mathios (1990) find that cereal manufacturers' advertising campaigns informing consumers of the health benefits related to fiber in cereal increased national consumption of fiber-rich cereals during the 1980s.

tion, contains information on consumers' response to the aggregate demand shock; namely, $Q_{rgt} = \sum_{(-g)} q_{(-g)t}$, where q is quantity and $-g$ denotes all geographies not g. In other words, Q_{rgt} is a response to aggregate demand shocks, and may be used as an element of $\Delta\mathbf{w}_{jm}$ due to its status as a predictive proxy for national advertising and other demand shocks.

A second source of market-level price endogeneity includes cases in which different retailers compete for the same consumers within a geographic location through promotional pricing and corresponding local advertising. This is consistent with Lal and Matutes (1994), Chintagunta (2002), Pesendorfer (2002), Shankar and Bolton (2004), and Hosken and Reiffen (2004a). Each of these studies conclude that supermarkets temporarily promote goods temporally to draw shoppers into stores. Consumers' demand for a good offered by a retailer within a geographic location may depend on the pricing and promotional activity of that retailer relative to other retailers within the geographic market, and the retailers may be setting prices and promotional activity based on that to draw in consumers. This potential confounding effect can be remedied by including the price of the good sold by retailer r's local competitor during the same time period as a conditioning instrument. Moreover, the time elapsed since the last promotion of a rival retailer is likely a driver of $\Delta\xi_{jm}$, and is thus an appropriate conditioning instrument. Finally, to the extent that local advertising information is available, this too ought to be included in $\Delta\mathbf{w}_{jm}$.

Finally, consider dynamics in pricing and consumer demand. For example, if the retailer temporarily reduces the price of a good periodically to capture welfare from low-valuation consumers, and if consumers tend to be the type to stockpile the good in low-price time periods, then confounding may exist. The dependency of stockpiling on the time that has elapsed since the last price promotion is consistent with the findings of Hendel and Nevo (2006) and Pesendorfer (2002). A few variables that drive stockpiling (or non-stockpiling) behavior of consumers and retailers' pricing decisions is the time that has elapsed since the last price promotion, local advertising, and lagged prices and quantities. These variables would also be included in the conditioning instrument set, $\Delta\mathbf{w}_{jm}$, as they are common drivers of or responses to the market appeal of the good, $\Delta\xi_{jm}$, in equation (2.10) and unobservable market-level determinants of price, $\Delta\zeta_{jm}$, in equation (2.14). Note that what makes lagged price an ideal conditioning instrument in this framework makes it an invalid exogenous instrument in the traditional instrumental variables framework proposed by Villas-Boas and Winer (1999). That is, to the extent that lagged price affects current price and demand,[16] lagged price would fail as an exogenous instrument, but would succeed as a conditioning instrument.

Path Diagram

Figure 2.14 presents a path diagram for the system of structural equations describing demand. Figure 2.15 presents a path diagram for the system of structural equations describing price. Variables outlined in dashed boxes are unobservable, and variables in the solid boxes are observable. Similarly, dashed arrows represent directional links from unobserved variables, and solid arrows represent directional links from observed variables. The diagrams help to determine the confounding variables that prevent structural identification of price effects in the demand system. That is, there are common unobservable factors that determine both demand and market price in the figures; these are the confounding variables. The drivers of them or the responses to them are appropriate conditioning instruments, $\mathbf{w}_{jm} \equiv \{\mathbf{w}_j, \Delta\mathbf{w}_{jm}\}$.

Observable variables that are responses to unobserved drivers of demand in Figure 2.14 are appropriate conditioning instruments to the extent they are also responses of market price in Figure 2.15. Examples include the baseline price, product segmentation variables, product fixed effects, total quantity in other cities, and time fixed effects. These overlapping predictive proxies are conditioning instruments to be included in $\mathbf{w}_{jm}$. In general, the predictive proxies are the observable variables that move away from the unobservable variables (in dashed boxes) in the path diagrams, as they are responses, rather than drivers.

[16]Pesendorfer (2002) finds that the probability of a temporary price cut for ketchup by a retailer is an increasing function of the time elapsed since the last price cut. Chapter 1 presents price equations showing that the current price of cereal depends on lagged prices as well as the time elapsed since the last promotion.

Observable variables that are drivers of unobserved demand in Figure 2.14 are also appropriate conditioning instruments to the extent they are also drivers of market price in Figure 2.15. Examples include observed product characteristics, promotional activity, local advertising, lagged prices, and demographics. These overlapping drivers comprise $\mathbf{w}_{jm}$. The observed drivers are the observable variables that move towards the unobservable variables in the path diagrams, as they are drivers, rather than responses.

2.5 Data and Estimation Results

The ready-to-eat (RTE) cereal industry has received considerable attention by IO and marketing researchers. Studies that examine this industry include Schmalensee (1978), Ippolito and Mathios (1990), Hausman (1997), Rubinfeld (2000), Nevo (2001), Nevo and Wolfram (2002), and Shum (2004). See Nevo (2000a, 2001) for an informative background on the history and competitive nature of the RTE cereal industry.

Currently, the four major nationally-branded cereal manufacturers operating in the U.S. are General Mills, Kellogg, Post, and Quaker. A fifth important source of cereal sales are store brand products that are low-price versions of branded cereals. General Mills and Kellogg are publicly traded companies that also sell a variety of other food products. Kellogg acquired the Kashi line of cereals in June 2000. In August 2008, Ralston Corp., the primary manufacturer of store brand cereals, acquired the Post cereal business from Kraft for approximately $2.6 billion.[17] Finally, Quaker is owned by Pepsi Co., a publicly traded company that primarily operates in the beverage market. Figure 2.1 presents market share information examined further in subsequent sections.

[17] The Federal Trade Commission granted early termination of the HSR waiting period in January 2008.

2.5.1 Data

The primary data set employed here is supermarket scanner data from Information Resources, Inc. (IRI). The data consists of variables measuring price, quantity, and merchandising and local advertising[18] variables for 150 of the top-selling RTE cereal stock-keeping units (SKUs) for January 2005 through December 2007. The data has a panel structure, where the time dimension consists of a weekly frequency and the individual dimension consists of a supermarket retail chain located in a geographic market. Except for missing observations and other irregularities, there are 157 weeks of data for 121 retailer-geography combinations. The sample spans about 41 geographic markets, which are similar to the Census Bureau's definition of a metropolitan statistical area (MSA) or combined metropolitan statistical area (CMSA). On average, there are three major supermarket chains comprising each of the 41 geographic markets in the sample. Figure 2.2 presents summary statistics for the top-selling product of each of the five firms.

This level of detail in market-level scanner data is typically not found in most IO studies of demand. Hausman's (1997) estimation of demand for RTE cereal was based on scanner data aggregated up to the city level. Nevo's (2001) estimation of demand for RTE cereal was based on scanner data aggregated at the city-quarter level. Figures 2.16 and 2.17 demonstrate the information that is potentially lost when data comes in an aggregated form. Figure 2.16 presents the price and market share for General Mills Cheerios 15oz. in just one of the five supermarket chains in a major metropolitan geographic market. Figure 2.17 displays the weighted average price and market share for General Mills Cheerios 15oz. aggregated at the geography-quarter level, where aggregation of geography includes the five supermarket chains that operate in that geographic market. Much of the variation in price and market share is lost in the aggregated data, which is particularly important when the product segment being examined is sold at supermarkets where temporary price promotions account for a significant part of the

[18]See Little (1998) for a detailed explanation of the local advertising and merchandising variables.

series' variance.[19] To get a sense of how much aggregation affects demand estimates, certain results in subsequent sections are reported using both disaggregated and aggregated data.

Income and demographics data gathered from the Bureau of Labor Statistics and the Census Bureau are merged with the scanner data based on matching IRI's geographic market with a MSA or CMSA. Average weekly wage data for each geographic market is collected from the BLS's Quarterly Census of Employment and Wages (QCEW) database. Demographic information for each of the IRI geographic markets is gathered from the Census Bureau's 2000 census. Figure 2.3 presents descriptive statistics for the income and demographics variables.

Finally, a detailed set of product characteristics is gathered from cereal manufacturer web sites and boxes. As in Nevo's studies, the product characteristics includes calories, fat, fiber, sugar, and product segmentation (kid, adult, family); but unlike Nevo's data, several other product characteristics are available, such as nutrient content, shape and type (e.g., flake, ring, biscuit), whether the product is frosted, contains raisins or nuts, the primary and secondary ingredient, volume, and density, among other attributes. One reason extra effort is placed on gathering such a comprehensive set of product characteristics is to determine if it's even practically possible to have a set of product characteristics that delivers the same estimate of α_o as product fixed effects. The answer is yes, and will be examined subsequently (see Figure 2.5).

Figure 2.4 provides summary statistics for a subset of 40 product characteristics for the 64 cereal products that comprise the inside goods of the choice set. The average weight is about one pound of cereal per box, with an average density of approximately 1.5 ounces per cup. Fourteen percent of the cereals contain raisins and nine percent are frosted. By far, flakes are the most common type of cereal, followed by rings. Most cereals in the sample primarily contain wheat or corn, followed by oat or rice. Finally, 50 percent of the 64 products compose what is known as the adult segment, with the kid and family segments comprising the remaining 50 percent.

[19]See, e.g., Pesendorfer (2002) and Hosken and Reiffen (2004b).

Market Shares and Prices

The four major manufacturers of RTE cereal and several store brand products are represented in the data. Figure 2.1 presents aggregate market share information on the 150 top-selling SKUs. Kellogg and General Mills are the leading nationally-branded manufacturers based on dollar sales and quantity (pound) sales, followed by Post, Quaker, and store brands. All four of the major manufacturers price well above the store brand products, selling at a 32 to 69 percent premium, while enjoying higher market shares.

Figure 1.16 of Chapter 1 presents the monthly market shares based on quantity purchases in pounds. The average market share of the store brand products during the 36 months ending December 2007 is 7.4 percent with a standard deviation of only 0.40 percent. Post and Quaker's market shares vary more than store brands, but are still relatively less volatile than the two biggest manufacturers in the market, Kellogg and General Mills. The average monthly market share for Post and Quaker are 16 percent and 8.4 percent, with standard deviations of 1.1 percent and 1.0 percent, respectively. In contrast, Kellogg and General Mills have market shares that vary more over time than the other manufacturers. Kellogg has an average market share of 37.9 percent with a standard deviation of 2.1 percent, and General Mills has an average market share of 30.3 percent with a standard deviation of 2.0 percent.

Figure 1.16 of Chapter 1 indicates that not only do the market shares of Kellogg and General Mills vary considerably over time, but that they are nearly mirror images. The correlation of the monthly market share for Kellogg and General Mills is -0.78, indicating that Kellogg's gain in market share is typically General Mills' loss, and vice-versa. Figure 1.17 presents the weighted average monthly retail price per pound for the manufacturers between 2005 and 2007. General Mills products tend to be sold at a higher price than the other manufacturers, while the store brands products are sold at a deep discount compared to the nationally-branded products. The temporal variation of prices is due to temporary price promotions that are more pronounced in the weekly data. See

Figure 2.16.

2.5.2 Estimation Results

Figures 2.5 through 2.10 provide estimates of α_o employing various specifications and identification assumptions. The dependent variable is $\widetilde{s}_{jm} \equiv \ln(s_{jm}) - \ln(s_{0m})$ for the $J = 64$ inside cereal products in the choice set. The outside good is the remaining 86 RTE cereal products in the sample that either had low market share, were products that had high price correlations with one or more of the 64 inside goods, or were products that had data quality issues.[20] The market share for each good is calculated based on quantity purchases in pounds, $s_{jm} \equiv q_{jm}/Q_m$, where q_{jm} is the quantity of good j in market m, and Q_m is the total quantity of cereal purchased (including the 86 outside goods) in market m. Market price, p_{jm}, is the price per pound of good j in market m. Except where noted, a market is a supermarket chain-geographic location-week combination.

Product-Level

Figures 2.5 and 2.6 provide estimates of α_o ignoring market-level endogeneity concerns. Column (1) of Figure 2.5 provides an estimate of -0.447 based on the assumption that price is exogenous. The corresponding own price elasticities for the top-selling product of each manufacturer are also presented in the table, along with the corresponding price-cost margin (PCM). The elasticities of the 64 products are evaluated at their respective median price and market share, and the PCMs reported are based on single-product firm pricing; i.e., the Lerner index. Section 5.2.4 (Figures 2.11 and 2.12), presents margins based on multi-product firm pricing. The OLS estimate of -0.447 yields elasticities that appear too inelastic and resulting PCMs that are too high (68 percent average). Of course,

[20] This differs from Nevo's (2001) definition of an outside good. In that study, an assumption is made regarding the number of servings of cereal an individual consumes in a given quarter, which is multiplied by the population to obtain the market size. It is perhaps at least as reasonable to use total cereal quantity as the market size, making the implicit assumption that the quantity of cereal consumed in each market is approximately proportional to the true size of the market. This obviates the need to rely on how many servings of cereal the population consumes.

assuming market price is exogenous is unreasonable, as the analysis of Section 4 described the confounding variables rendering market price endogenous. However, the estimates of column (1) in Figure 2.5 provide a benchmark case to compare subsequent results.

Column (2) of Figure 2.5 employs as conditioning instruments calories, fat, sugar, fiber, density, and product segmentation dummy variables.[21] This produces an estimate of -0.489, which is not much different from the -0.447 in column (1). Column (3) adds to the seven conditioning instruments an additional 30 product characteristics, including dummy variables for ingredients, vitamin and mineral content, dummy variables for the shape of the cereal, the volume of cereal in the box, the weight of the cereal in the box, dummy variables for additions to the cereal (e.g., raisins, nuts, oat clusters, frosting), and manufacturer dummy variables. The resulting CX estimate of α_o is -0.636, indicating that the comprehensive set of 37 product characteristics used as conditioning instruments has increased the average price elasticity from -1.55 to -2.21. Column (4) only includes the baseline price as a conditioning instrument. The baseline price by itself brings the CX estimate of α_o to -0.596. Remarkably, when just the seven product characteristics are used in conjunction with the baseline price in column (5), the CX estimate is -0.625, almost identical to the estimate of -0.636 with all 37 product characteristics. This suggests that the baseline price contains as much information as the 30 additional product characteristics, lending credibility to the product-level structural system of equations in Section 4.

Finally, column (6) of Figure 2.5 presents the estimate of α_o using product fixed effects, which ought to capture inter-product variation in characteristics that do not vary by market, including unobserved product characteristics. The resulting estimate is -0.647. Thus, the 37 product characteristics available here do about the same job as product fixed effects. More importantly, seven product characteristics and the baseline price provide a similar estimate of α_o compared to the product fixed effects or 37 product characteristics.

[21] These seven product characteristics are roughly the same as those used by Nevo (2001) in Table V columns (i) and (iv).

Figure 2.6 presents instrumental variables estimates (TSLS) of α_o using rivals' product characteristics as assumed exogenous instruments suggested by Berry (1994) and BLP (1995). For each product, the closest rival product is determined by using a variety of non-price and non-size product characteristics (see Figure 2.4) to determine a good's closest neighbor. This is done by specifying a distance measure to determine how far good j is in terms of observed product characteristics from all other goods not produced by j's manufacturer. The distance between good j and good k, where k indexes all products not produced by good j's manufacturer, is calculated as

$$\varphi_{jk} = \sum_{\forall s} \left| \frac{r_{js} - r_{ks}}{\sigma_s} \right|,$$

where s indexes the product characteristics, r_{js} is the value of characteristic s for product j, r_{ks} is the value of characteristic s for product k, and σ_s is the standard deviation of product characteristic s across all products in the sample. For product j, φ_{jk} is calculated for all k products, and the closest rival for j is the k that yields the smallest φ_{jk} value. The k that yields the second smallest φ_{jk} value is product j's second closest rival, and so on.[22]

For the top rival and the top five rivals, the sugar content, fiber, fat, calories, and density are used as instruments in a TSLS procedure. The resulting estimates of α_o are found in Figure 2.6. The difference among the columns in Figure 2.6 are due to differing product characteristics being included as (assumed exogenous) regressors and different definitions of rivals (top rival versus average of top five rivals) for the instruments. What is most noticeable here is that none of the estimates come close to recovering -0.647 from column (6) in Figure 2.5. In fact, some of the IV estimates are positive, with the rest having magnitudes that make very little sense. It appears that using rivals' product characteristics as valid instruments has performed poorly in identifying the causal effect of market price on demand.

The conditional exogeneity estimates in Figure 2.5 appear to do a much

[22]This method yielded very sensible results that comports with *a priori* suspicions of product sets that are similar. E.g., Kellogg Raisin Bran and Post Raisin Bran.

better job in recovering sensible estimates of α_o compared to the TSLS estimates employing rivals' product characteristics. In general, the AIC statistics for the CX estimates in Figure 2.5 are smaller than the ones reported for the TSLS estimates in Figure 2.6, indicating that the CX specification has a better model fit. The next section addresses market-level endogeneity.

Product-Level and Market-Level

Figures 2.7 and 2.9 provide estimates of α_o addressing both product-level and market-level price endogeneity. The product-level conditioning instruments include those described previously for Figure 2.5. The set of market-level conditioning instruments here are variables that are either predictive proxies (responses) or drivers of the unobservable confounding variables that determine both demand and market price, as described in Section 4. This includes: (i) Variables typically found in scanner data, such as in-store display, distribution, and local advertising (see Figure 2.2); (ii) Demographic and income variables obtained from the Census Bureau and BLS, including the distribution of age, percentage of households with children, race, and average weekly wage (see Figure 2.3); and (iii) A variety of variables obtained by exploiting the panel structure of the scanner data, such as the price of the good in the next largest retailer in the same city, the number of weeks that has elapsed since the good had a temporary price promotion, the number of weeks that has elapsed since the good had a temporary price promotion in the next largest retailer in the same city, lagged prices, lagged quantities, the total amount of the good sold in all other cities, linear trend, retailer-city fixed effects, and month fixed effects.

Columns (1) - (6) in Figure 2.7 are directly comparable to columns (1) - (6) in Figure 2.5, the difference being that the market-level conditioning instruments described above are included in Figure 2.7. Column (1) of Figure 2.7 provides the CX estimate of α_o using only the market-level conditioning instruments. Regarding columns (2) - (6), what is notable is that the conditional exogeneity estimates of the price effects are relatively stable regardless of the product-level conditioning instruments used, as they are all between -0.599 and -0.575. This yields an average price elasticity for all 64 products in the sample between -2.08 and -2.00, and corresponding average price-cost margins between 51 and 53 percent, all of which are reasonable estimates. It thus appears that accounting for market-level price endogeneity dominates product-level price endogeneity since the estimates hinge only moderately on the product-level conditioning instruments used. Nonetheless, CX estimates of α_o in columns (4) and (5), which employ the baseline price as a conditioning instrument, are extremely close to the CX estimate of α_o employing product fixed effects.

Column (7) reproduces column (6) but excluding merchandising and local advertising from the market-level conditioning instrument set. This causes the CX estimate of α_o to go from -0.599 to -0.846, a substantial change in magnitude. By omitting merchandising and local advertising, the price elasticity of General Mills Cheerios 15oz., the top-selling cereal product in the sample, changes from -2.07 to -2.92 with a corresponding margin change from 48 percent to 34 percent. Recall that merchandising and local advertising were proposed as drivers of stockpiling in Section 4, which explains why omitting these variables would cause the price effect to be more elastic. By not including them as conditioning instruments, the estimate of α_o of -0.846 is including stockpiling effects, which is expected to result in price effects that are overestimated. Note that column (7) reports the largest AIC, indicating poor model fit compared to the other specifications that do include merchandising and local advertising variables.

Figure 2.9 presents TSLS estimates employing as market-level exogenous instruments the weighted average price of the good in other geographic markets suggested by Hausman (1997) and lagged prices suggested by Villas-Boas and Winer (1999). In all the specifications, the following market-level variables are included as assumed exogenous regressors: merchandising and local advertising, the demographics and income variables described earlier, time trend, retailer-city fixed effects, and month fixed effects. In general, the specifications in the first four columns produce estimates of α_o that are too small in magnitude, yielding many inelastic price elasticities and impossibly high PCMs over 100 percent. Recall that the basic OLS estimate of α_o without any instruments or regressors is -0.447

found in column (1) of Table 2.5, which yields more reasonable results than most of the TSLS estimates from Figure 2.9.

Column (5) of Figure 2.9 employs product fixed effects and includes lagged prices as assumed exogenous instruments. The estimate of -0.438 yields margins for the top-selling products that are still implausibly high. The same specification except employing prices in other cities as exogenous instruments is found in column (6), which again produces estimates of PCMs that are too high. This suggests that lagged prices and prices in other cities do not appear to be doing a good job of addressing price endogeneity. An important point here is that lagged price appears to be more appropriately classified as a conditioning instrument (Figure 2.7) rather than an exogenous instrument (Figure 2.9). That is, lagged price certainly has a role to play in the identification process, but the economic system proposed in Section 4 suggests its role is to identify α_o through $p_{jm} \perp \xi_{jm}|\mathbf{w}_{jm}$ rather than $z_{jm} \perp \xi_{jm}$.

Finally, column (7) reproduces column (6) but excludes merchandising and local advertising from the set of market-level regressors. This causes the TSLS estimate of α_o to go from -0.495 to -0.781, again a substantial change in magnitude. By omitting merchandising and local advertising from the included set of regressors, the average price elasticity changes from -1.72 to -2.71, with a corresponding margin change from 61 percent to 39 percent. By omitting merchandising and local advertising, the PCMs are now similar to those obtained by Nevo (2001).[23] The AIC value for the specification in column (7) is the largest compared to all the other IV specifications, indicating relative misspecification.

Aggregated Data

Figures 2.8 and 2.10 present estimates using aggregated scanner data. The detailed retailer-geography-week data set is "artificially" aggregated into a geography quarter (i.e., city-quarter) data set. See Figures 2.16 and 2.17. Figure 2.8 is

[23] See Nevo's Table VIII, pg. 333. His median elasticity is in the 32 to 36 percent range for single product pricing. It appears Nevo did not have access to merchandising and local advertising variables in his scanner data. The specification of column (7) in Figure 2.9 here is roughly comparable to Nevo's Table V columns (vii) and (ix).

directly comparable to the CX estimates of Figure 2.7, and Figure 2.10 is directly comparable to the TSLS estimates of Figure 2.9. In general, the estimates are less elastic with the aggregated data in the "-A" tables, which is largely due to the loss of information with the city-quarter data.

For example, the CX estimate of α_o using the disaggregated data found in column (5) of Figure 2.7 is -0.595, yielding an average elasticity and PCM of -2.07 and 51 percent, respectively. In contrast, the CX estimate of α_o employing the aggregated data found in column (5) of Figure 2.8 is -0.472, yielding an average elasticity and PCM of -1.43 and 74 percent, respectively. As another example, the TSLS estimate of α_o using the disaggregated data found in column (6) of Figure 2.9 is -0.495, yielding an average elasticity and PCM of -1.72 and 61 percent, respectively. In contrast, the TSLS estimate of α_o employing the aggregated data found in column (6) of Figure 2.10 is -0.261, yielding an average elasticity and PCM of -0.79 and 134 percent, respectively.

These results have two important implications. First, disaggregated scanner data ought to be employed in demand studies, as they contain important variation lost in the aggregated data. Hosken et al. (2002) point out that estimating demand systems using aggregated scanner data can cause biased results and incorrect statistical inferences. This appears to be the case here. Moreover, the results suggest that demand studies that obtain inelastic estimates when employing aggregated data may be incorrectly attributing the problem to simultaneity rather than aggregation bias. If simultaneity is the problem, and if the instruments that are often employed in demand studies are valid, then the estimates from Figures 2.9 and 2.10 would have been similar.

Price-Cost Margins and Hidden Price Changes

Figure 2.11 presents the implied price-cost margins for the estimates of α_o from column (6) in Figure 2.7 (the CX estimate) and column (6) in Figure 2.9 (the TSLS estimate). The PCMs that are calculated include the following: (i) Single-product firm (SPF) pricing, defined here as the case where each of the $J = 64$ cereal products' prices are set separately by 64 firms; (ii) Multi-product firm (MPF) pricing, defined here as the case in which each of the cereal firms sets the price for its own goods; and (iii) Joint ownership monopoly (JOM) pricing, defined here as the case in which one firm sets the price of all 64 products jointly. It is assumed that prices are set based on a differentiated products Nash-Bertrand oligopoly model with a unique (pure strategy) equilibrium supported by positive prices. The general system of first order conditions is given by

$$\underset{(J\times 1)}{\begin{bmatrix} s_1 \\ s_2 \\ \vdots \\ s_J \end{bmatrix}} + \left(\underset{(J\times J)}{\mathbf{A}} \circ \underset{(J\times J)}{\begin{bmatrix} \frac{\partial s_1}{\partial p_1} & \frac{\partial s_2}{\partial p_1} & \cdots & \frac{\partial s_J}{\partial p_1} \\ \frac{\partial s_1}{\partial p_2} & \frac{\partial s_2}{\partial p_2} & \cdots & \frac{\partial s_J}{\partial p_2} \\ \vdots & & \ddots & \vdots \\ \frac{\partial s_1}{\partial p_J} & \frac{\partial s_2}{\partial p_J} & \cdots & \frac{\partial s_J}{\partial p_J} \end{bmatrix}} \right) \underset{(J\times 1)}{\begin{bmatrix} p_1 - mc_1 \\ p_2 - mc_2 \\ \vdots \\ p_J - mc_J \end{bmatrix}} = \underset{(J\times 1)}{\begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}} \quad \text{or} \tag{2.15}$$

$$\mathbf{s} + \mathbf{\Omega}\psi = \mathbf{0}, \tag{2.16}$$

where s, p, and mc denote market share, price, and marginal cost, respectively, and $\circ$ denotes the Hadamard product (i.e., element-wise multiplication). What distinguishes SPF from MPF from JOM in the system of equations in (2.15) and (2.16) is the matrix $\mathbf{A}$ that comprises $\mathbf{\Omega}$. The elements of $\mathbf{A}$ are ones and zeros depending on the assumed industry conduct. For SPF, $\mathbf{A}$ has elements $a_{jk} = 1$ if $j = k$, and 0 otherwise; i.e., $\mathbf{A}$ is the identity matrix. For MPF, $\mathbf{A}$ has elements $a_{jk} = 1$ if j, k are owned by the same firm, and $a_{jk} = 0$ otherwise. For JOM, $\mathbf{A}$ has elements $a_{jk} = 1$ for all j, k; i.e., $\mathbf{A}$ is a matrix of ones. For each of the industry structures, the $(J \times 1)$ vector of price-cost margins, $(p_j - mc_j)/p_j$, is given by

$$\text{PCM} = \psi \circ \widetilde{\mathbf{p}} = \left(-\mathbf{\Omega}^{-1}\mathbf{s}\right) \circ \widetilde{\mathbf{p}}, \tag{2.17}$$

where $\widetilde{\mathbf{p}}$ is a $(J \times 1)$ vector of the inverse of prices, and $\circ$ again denotes element-wise multiplication.

The PCMs are calculated based on (2.17) for all 64 products. Figure 2.11 reports the PCMs for the top-selling product of the five RTE cereal manufacturers, as well as the average PCM for the five firms. The margins are all evaluated at the median price and market share for each of the products. First, note that the margins increase going from single-product firm pricing to multi-product to joint ownership, as expected. That is, the less firms in the market that control the 64 products, the higher the margins earned on those products. Next, the store brand products enjoy higher margins than their nationally-branded counterparts. For example, the average store brand PCM based on the CX estimate is 86 percent in the MPF column, while the branded firms earn an average of 51 to 61 percent. Given the less elastic estimate of α_o from the 2SLS procedure, the PCMs resulting from that estimate are higher compared to the CX estimate. For example the average PCM for Kellogg's products based in CX is 61 percent in the multi-product firm column, compared to 74 percent based on TSLS.

Figure 2.12 presents PCMs based on the CX and TSLS estimates from column (7) in Figures 2.7 and 2.9, respectively. These estimates exclude merchandising and local advertising variables. Compared to Figure 2.11, the PCMs in Figure 2.12 are much lower. For example, the average PCM of all 64 RTE cereal products based on CX estimation with multi-product firm pricing is 41 percent in Figure 2.12, compared to 58 percent in Figure 2.11. While both sets of PCMs based on the CX estimates of α_o appear reasonable, lower PCMs from omitting merchandising and local advertising variables in the demand estimation stage suffer from the confounding effects due to stockpiling behavior. If merchandising and local advertising variables are not available, researchers may overestimate elasticities and underestimate corresponding price-cost margins.

Finally, Figure 2.13 presents the price-cost margins for the five cereal manufacturers estimated separately before and after July 15, 2007. General Mills increased the price of nearly all of its products by decreasing the amount of cereal in the box, which took effect in supermarkets near mid-July 2007 according to the scanner data. This phenomenon of increasing the price of a good by changing a product characteristic, defined a "hidden" price change, is the subject of Chapter 4. Figure 2.13 presents preliminary estimates of the change in profitability before and after the hidden price change. The CX specification from Figure 2.7 column (6) is employed here with three modifications. First, market price is interacted with product dummy variables in order to estimate a product-specific idiosyncratic deviation from the overall price effect. The mean and standard deviation of the deviations are reported. Second, demand is estimated separately before and after mid-July 2007. Third, the elasticities and margins are evaluated at the weighted average price and market share for each good.

For the sample period prior to July 15, the average marginal price effect is -0.723, and the mean product-specific deviation across the 64 products is 0.073. For General Mills, this results in an average price-cost margin of 66.6 percent, compared to an average PCM of 70.2 percent in the sample period after they reduced their box size. That is, General Mills' average PCM increased by 3.6 percentage points, consistent with the notion that some fraction of consumers did not notice the hidden price change. As of 2008, which is outside of the sample period, Kellogg also increased the prices of its cereal products by decreasing the amount of cereal in the box. This comports with the margin calculations in Figure 2.13, as Kellogg lost nearly 2.8 percentage points after General Mills committed to their price increase. That is, Kellogg lost profits after General Mills moved first, and now they have followed suit. Moreover, the PCMs for the other three manufacturers are either flat or have increased; this potentially explains why they have not followed General Mills and Kellogg in committing to a hidden price change. These issues are examined in Chapter 4.

2.6 Conclusions and Further Research

This chapter provides an alternate approach to classifying and remedying price endogeneity in a discrete choice model of demand. The structural system of equations describing demand and market price proposed in Section 4 help to determine the confounding variables that potentially lead to price endogeneity in the demand system. The proposed structural equations also help in determining the conditioning instruments to identify price effects with the conditional exogeneity (CX) assumption $p_{jm} \perp \xi_{jm} | \mathbf{w}_{jm}$. This framework obviates the need to assume that observed product characteristics are exogenous, as observed product characteristics are more appropriately classified as conditioning instruments employed in the CX estimator examined in Section 3.

Several conditioning instruments are proposed, such as the baseline price and other variables obtained by exploiting the panel structure of the scanner data. The CX identification framework is used to estimate demand for a large sample of ready-to-eat cereal products. The estimates of price effects resulting from conditional exogeneity are more reasonable and stable compared to standard instrumental variables (IV). Thus, CX is an alternate identification strategy that may be more palatable in many circumstances.

Estimates based on scanner data aggregated up to the city-quarter level yield inelastic estimates compared to estimates from disaggregated data. These differences are not mere discrepancies, but are potentially meaningful differences suggesting that past demand studies using aggregated scanner data may have incorrectly attributed inelastic estimates to simultaneity rather than aggregation bias. Moreover, omitting merchandising and local advertising variables has a dramatic impact on both the CX and IV estimates of demand, causing elasticity estimates to increase appreciably due to confounding effects from unobserved stockpiling effects, which leads to significantly lower implied price-cost margins. Finally, margins are estimated before and after General Mills increased the price of its cereals by decreasing the amount of cereal in the box. The estimates support what has transpired in the cereal industry, whereby Kellogg followed suit in engaging in "hidden" price changes, while other manufacturers did not. A detailed empirical study of the hidden price change phenomenon can be found in Chapter 4.

In order to incorporate consumer heterogeneity and flexible substitution patterns, the next revision of this chapter will include estimates from a random coefficients logit model along the lines of BLP (1995) and Nevo (2001). The structural equations, conditioning instruments, and identification schemes presented here are general enough to apply to any of the logit models of demand. This research is currently in progress. Further investigation includes examining nonparametric estimation that relaxes the assumption that utility (demand) is linearly-separable. That is, rather than assuming that the conditional logit model takes the form $\widetilde{s}_{jm} = \alpha_o p_{jm} + \xi_{jm}$, it will take the more flexible form $\widetilde{s}_{jm} = d(p_{jm},$

ξ_{jm}), where the average marginal effect of price on demand, $\partial d(\cdot)/\partial p$, is a function of price rather than a coefficient estimate. This will provide much-needed insight regarding information that is lost by assuming a linear-separable parametric utility for the logit class of demand models. See Chapter 3.

2.7 Tables and Figures

	Market Share - DS	Market Share - PS	Number of SKUs	Weighted Avg. Price Per Pound	% Above/Below Total Avg. Price	% Above Store Brand Avg. Price
Kellogg	37.9%	37.9%	55	$2.95	-0.1%	46.7%
General Mills	34.9%	30.3%	45	$3.40	15.2%	69.1%
Post	14.7%	16.0%	25	$2.71	-8.5%	34.4%
Quaker	7.6%	8.4%	13	$2.66	-10.2%	31.9%
Store Brands	5.0%	7.3%	12	$2.01	-31.9%	
Total	100.0%	100.0%	150	$2.96		

Source: IRI scanner data. Based on 121 supermarket chain-geography combinations from 2005-2007. Kellogg Includes six Kashi brand products. Market Share - DS is based on dollar sales. Market Share - PS is based on pound sales.

Figure 2.1: Market Share and Price

Variable	SKU Description	Mean	Median	StDev.	Kurt.	Skew.	Obs.
Price	G MILLS CHEERIOS BOX 15OZ	3.52	3.51	0.72	2.45	-0.12	18,987
	KELLOGG FROSTED FLAKES BOX 20OZ	2.91	2.88	0.72	2.35	0.12	18,988
	POST HONEY BUNCHES OF OATS REG BOX 16OZ	2.99	2.98	0.67	2.17	-0.05	18,997
	QUAKER LIFE REGULAR BOX 21OZ	2.85	2.88	0.58	2.21	-0.02	17,781
	STR BDS RAISIN BRAN BOX 20OZ	1.80	1.82	0.34	2.75	0.20	18,524
Market Share	G MILLS CHEERIOS BOX 15OZ	2.5%	1.7%	2.4%	20.91	3.50	18,987
	KELLOGG FROSTED FLAKES BOX 20OZ	2.1%	1.2%	2.7%	29.18	4.26	18,988
	POST HONEY BUNCHES OF OATS REG BOX 16OZ	1.7%	1.1%	1.8%	21.94	3.31	18,997
	QUAKER LIFE REGULAR BOX 21OZ	1.3%	1.0%	1.1%	33.07	4.41	17,781
	STR BDS RAISIN BRAN BOX 20OZ	1.3%	1.1%	1.0%	19.99	2.71	18,524
Local Ad Only	G MILLS CHEERIOS BOX 15OZ	0.15	0.00	0.31	5.07	1.92	18,987
	KELLOGG FROSTED FLAKES BOX 20OZ	0.19	0.00	0.35	3.62	1.54	18,988
	POST HONEY BUNCHES OF OATS REG BOX 16OZ	0.15	0.00	0.32	4.84	1.88	18,997
	QUAKER LIFE REGULAR BOX 21OZ	0.08	0.00	0.24	11.37	3.15	17,781
	STR BDS RAISIN BRAN BOX 20OZ	0.07	0.00	0.24	12.04	3.26	18,524
In-Store Display Only	G MILLS CHEERIOS BOX 15OZ	0.05	0.00	0.15	21.02	4.07	18,987
	KELLOGG FROSTED FLAKES BOX 20OZ	0.06	0.00	0.16	16.44	3.55	18,988
	POST HONEY BUNCHES OF OATS REG BOX 16OZ	0.06	0.00	0.16	14.86	3.36	18,997
	QUAKER LIFE REGULAR BOX 21OZ	0.02	0.00	0.09	56.89	6.67	17,781
	STR BDS RAISIN BRAN BOX 20OZ	0.03	0.00	0.11	32.95	5.11	18,524
Local Ad and Display	G MILLS CHEERIOS BOX 15OZ	0.10	0.00	0.25	8.32	2.57	18,987
	KELLOGG FROSTED FLAKES BOX 20OZ	0.08	0.00	0.23	10.24	2.89	18,988
	POST HONEY BUNCHES OF OATS REG BOX 16OZ	0.07	0.00	0.21	11.35	3.07	18,997
	QUAKER LIFE REGULAR BOX 21OZ	0.03	0.00	0.14	30.12	5.18	17,781
	STR BDS RAISIN BRAN BOX 20OZ	0.01	0.00	0.09	64.43	7.58	18,524
Distribution	G MILLS CHEERIOS BOX 15OZ	0.99	1.00	0.03	274.34	-13.55	18,987
	KELLOGG FROSTED FLAKES BOX 20OZ	0.98	1.00	0.04	67.73	-6.00	18,988
	POST HONEY BUNCHES OF OATS REG BOX 16OZ	0.99	1.00	0.03	75.20	-6.19	18,997
	QUAKER LIFE REGULAR BOX 21OZ	0.98	1.00	0.05	27.36	-4.26	17,781
	STR BDS RAISIN BRAN BOX 20OZ	0.96	1.00	0.07	22.79	-3.76	18,524

Source: IRI scanner data. Statistics are based on 121 supermarket chain-geographic markets during 157 weeks from 2005 through 2007.
Merchandising and Distribution variables are % all commodity value (on a 0 to 1 scale). See Little (1998) for a detailed explanation of these variables.
Percentiles available from author upon request.

Figure 2.2: Descriptive Statistics

							Percentiles					
	Mean	Median	StDev.	Kurt.	Skew.	Obs.	1%	5%	25%	75%	95%	99%
Avg. Weekly Wage ($)	807.4	786.0	124.3	3.4	1.4	780	607.8	652.0	723.0	868.0	1037.1	1228.4
White (%)	77.0	79.0	10.2	-0.7	-0.4	65	54.3	59.2	69.5	85.2	90.9	93.1
Age <= 19 (%)	28.7	28.8	1.9	1.3	0.4	65	24.6	25.7	27.5	29.6	31.5	33.1
Age >=55 (%)	27.5	27.4	2.4	1.0	0.5	65	22.9	23.3	26.1	28.8	31.6	34.1
Female (%)	51.1	51.3	0.7	-0.3	-0.5	65	49.5	49.8	50.7	51.6	52.2	52.4
Households w/Children (%)	63.8	63.3	4.8	2.8	0.6	65	52.4	57.1	61.5	65.5	71.9	77.3

Source: Bureau of Labor Statistics, Quarterly Census of Employment and Wages (QCEW), and Census Bureau, 2000 Census Information.
Percentages refer to the percent of the MSA's population.
Avg. Weekly Wage is collected on a quarterly basis (65 MSAs x 12 quarters = 780 obs.).

Figure 2.3: Descriptive Statistics - Demographics

Variable	Mean	StDev	Min	Max
SKU Box Weight	16.77	3.73	8.56	27.00
SKU Box Volume	12.01	3.46	5.87	20.50
Density	1.51	0.53	0.75	4.09
Nutritional Info				
Calories	155.43	50.04	86.67	400.00
Total Fat	1.59	1.51	0.00	9.33
Total Carbs	34.38	11.88	18.67	96.00
Fiber	3.85	4.50	0.00	28.00
Sugar	10.94	5.87	0.00	20.00
Protein	3.54	2.39	1.00	13.33
Additions				
Raisin	0.14	0.35	0	1
Oat Cluster	0.14	0.35	0	1
Frosted	0.09	0.29	0	1
Nuts	0.03	0.17	0	1
Shape/Type				
Flake	0.39	0.49	0	1
Ring	0.16	0.36	0	1
Biscuit	0.06	0.24	0	1
Primary Ingredient				
Wheat	0.33	0.47	0	1
Corn	0.27	0.44	0	1
Oat	0.22	0.41	0	1
Rice	0.14	0.35	0	1
Sugar	0.05	0.21	0	1
Category				
Adult	0.50	0.50	0	1
Kid	0.30	0.46	0	1
Family	0.20	0.40	0	1

Source: Manufacturer websites and cereal boxes collected during Q4-2008.
Summary statistics are based on 64 cereal products.
SKU box weight is ounces per box. SKU box volume is cups per box.
Density is calculated as ounces per cup.
Nutritional info are measured in grams per cup. Calories are per cup.

Figure 2.4: Descriptive Statistics - Product Characteristics

Dep Var: $\ln(S_{jm}) - \ln(S_{0m})$	(1)	(2)	(3)	(4)	(5)	(6)
Market Price	-0.447***	-0.489***	-0.636***	-0.596***	-0.625***	-0.647***
(Standard Error)	(0.0007)	(0.0008)	(0.0009)	(0.0011)	(0.0010)	(0.0009)
Conditioning Instruments						
Product Fixed Effects	No	No	No	No	No	Yes
7 Product Characteristics	No	Yes	Yes	No	Yes	No
30 Additional Product Characteristics	No	No	Yes	No	No	No
Baseline Price	No	No	No	Yes	Yes	No
Observations	1,186,098	1,186,098	1,186,098	1,186,098	1,186,098	1,186,098
R-Squared	0.276	0.356	0.493	0.308	0.385	0.556
RMSE	0.716	0.675	0.599	0.700	0.660	0.561
AIC/Obs	2.171	2.053	1.814	2.125	2.007	1.681
	Price Elasticity					
GM Cheerios 15oz.	-1.54	-1.69	-2.20	-2.06	-2.16	-2.23
KG Frosted Flakes 20oz.	-1.27	-1.39	-1.81	-1.70	-1.78	-1.84
PO Honey Bunches of Oats Reg 16oz.	-1.32	-1.44	-1.87	-1.76	-1.84	-1.91
QK Life Regular 21oz.	-1.28	-1.40	-1.82	-1.70	-1.78	-1.85
SB Raisin Bran 20oz.	-0.80	-0.88	-1.14	-1.07	-1.12	-1.16
Avg of All 64 Products	-1.55	-1.70	-2.21	-2.07	-2.17	-2.25
	Price-Cost Margin					
GM Cheerios 15oz.	65%	59%	46%	49%	46%	45%
KG Frosted Flakes 20oz.	79%	72%	55%	59%	56%	54%
PO Honey Bunches of Oats Reg 16oz.	76%	69%	53%	57%	54%	52%
QK Life Regular 21oz.	78%	72%	55%	59%	56%	54%
SB Raisin Bran 20oz.	125%	114%	88%	93%	89%	86%
Avg of All 64 Products	68%	62%	48%	51%	49%	47%

***, **, * denote significance at 1%, 5%, and 10%, respectively, based on White (1980) heteroskedasticity-robust standard errors.

Figure 2.5: OLS and CX Estimates, Product-Level

Dep Var: $\ln(S_{jm})$ - $\ln(S_{0m})$	(1)	(2)	(3)	(4)	(5)	(6)
Market Price	-0.163***	-0.382***	0.765***	0.046***	-0.184***	-1.833***
(Standard Error)	(0.0021)	(0.0034)	(0.0178)	(0.0026)	(0.0044)	(0.0207)
Included Product Characteristics						
Product Fixed Effects	No	No	No	No	No	No
7 Product Characteristics	No	Yes	Yes	No	Yes	Yes
30 Additional Product Characteristics	No	No	Yes	No	No	Yes
Instuments						
Top Rival's Characteristics	Yes	Yes	Yes	No	No	No
Top 5 Rivals' Characteristics	No	No	No	Yes	Yes	Yes
Observations	1,186,098	1,186,098	1,186,098	1,186,098	1,186,098	1,186,098
F	33,118	8,904	1,776	24,461	6,705	922
AIC/Obs	2.477	2.408	2.252	2.486	2.399	2.228
			Price Elasticity			
GM Cheerios 15oz.	-0.56	-1.32	2.64	0.16	-0.64	-6.33
KG Frosted Flakes 20oz.	-0.46	-1.09	2.18	0.13	-0.52	-5.22
PO Hny Bunches Oats Reg 16oz.	-0.48	-1.13	2.25	0.14	-0.54	-5.40
QK Life Regular 21oz.	-0.47	-1.09	2.18	0.13	-0.53	-5.23
SB Raisin Bran 20oz.	-0.29	-0.69	1.37	0.08	-0.33	-3.29
Avg of All 64 Products	-0.57	-1.33	2.66	0.16	-0.64	-6.37
			Price-Cost Margin			
GM Cheerios 15oz.	178%	76%	-38%	-630%	157%	16%
KG Frosted Flakes 20oz.	215%	92%	-46%	-763%	191%	19%
PO Hny Bunches Oats Reg 16oz.	208%	89%	-44%	-738%	184%	19%
QK Life Regular 21oz.	215%	92%	-46%	-761%	190%	19%
SB Raisin Bran 20oz.	341%	146%	-73%	-1210%	303%	30%
Avg of All 64 Products	186%	79%	-40%	-660%	165%	17%

***, **, * denote significance at 1%, 5%, and 10%, respectively, based on White (1980) heteroskedasticity-robust standard errors.
The F statistic reported tests that the first stage coefficients for the excluded instruments are jointly equal to zero.
AIC/Obs is for the reduced form.

Figure 2.6: IV Estimates, Product-Level

Dep Var: ln(S_{jm}) - ln(S_{0m})	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Market Price	-0.576***	-0.575***	-0.582***	-0.598***	-0.595***	-0.599***	-0.846***
(Standard Error)	(0.0012)	(0.0012)	(0.0012)	(0.0012)	(0.0012)	(0.0012)	(0.0012)
Product-Level Conditioning Instruments							
Product Fixed Effects	No	No	No	No	No	Yes	Yes
7 Product Characteristics	No	Yes	Yes	No	Yes	No	No
30 Additional Product Chars	No	No	Yes	No	No	No	No
Baseline Price	No	No	No	Yes	Yes	No	No
Market-Level Conditioning Instruments	Yes	Yes	Yes	Yes	Yes	Yes	Yes[a]
Observations	1,100,564	1,100,564	1,100,564	1,100,564	1,100,564	1,100,564	1,100,564
R-Squared	0.718	0.728	0.752	0.726	0.733	0.771	0.669
RMSE	0.444	0.436	0.416	0.438	0.432	0.400	0.481
AIC/Obs	1.215	1.178	1.084	1.186	1.160	1.007	1.376
	Price Elasticity						
GM Cheerios 15oz.	-1.99	-1.99	-2.01	-2.06	-2.05	-2.07	-2.92
KG Frosted Flakes 20oz.	-1.64	-1.64	-1.66	-1.70	-1.69	-1.71	-2.41
PO Hny Bunches Oats Reg 16oz.	-1.70	-1.69	-1.72	-1.76	-1.75	-1.77	-2.49
QK Life Regular 21oz.	-1.65	-1.64	-1.66	-1.71	-1.70	-1.71	-2.42
SB Raisin Bran 20oz.	-1.03	-1.03	-1.05	-1.07	-1.07	-1.08	-1.52
Avg of All 64 Products	-2.00	-2.00	-2.02	-2.08	-2.07	-2.08	-2.94
	Price-Cost Margin						
GM Cheerios 15oz.	50%	50%	50%	48%	49%	48%	34%
KG Frosted Flakes 20oz.	61%	61%	60%	59%	59%	59%	41%
PO Hny Bunches Oats Reg 16oz.	59%	59%	58%	57%	57%	57%	40%
QK Life Regular 21oz.	61%	61%	60%	59%	59%	58%	41%
SB Raisin Bran 20oz.	97%	97%	96%	93%	94%	93%	66%
Avg of All 64 Products	53%	53%	52%	51%	51%	51%	36%

***, **, * denote significance at 1%, 5%, and 10%, respectively, based on White (1980) heteroskedasticity-robust standard errors.
[a]Excludes merchandising and local advertising variables.

Figure 2.7: CX Estimates, Product-Level and Market-Level

Dep Var: $\ln(S_{jm})$ - $\ln(S_{0m})$	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Market Price	-0.398***	-0.433***	-0.499***	-0.465***	-0.472***	-0.553***	-0.668***
(Standard Error)	(0.0091)	(0.0093)	(0.0088)	(0.0092)	(0.0092)	(0.0084)	(0.0093)
Product-Level Conditioning Instruments							
Product Fixed Effects	No	No	No	No	No	Yes	Yes
7 Product Characteristics	No	Yes	Yes	No	Yes	No	No
30 Additional Product Chars	No	No	Yes	No	No	No	No
Baseline Price	No	No	No	Yes	Yes	No	No
Market-Level Conditioning Instruments	Yes	Yes	Yes	Yes	Yes	Yes	Yes[a]
Observations	28,848	28,848	28,848	28,848	28,848	28,848	28,848
R-Squared	0.759	0.768	0.808	0.768	0.776	0.845	0.791
RMSE	0.299	0.293	0.267	0.293	0.288	0.240	0.279
AIC/Obs	0.427	0.386	0.202	0.387	0.351	-0.015	0.287
				Price Elasticity			
GM Cheerios 15oz.	-1.20	-1.30	-1.50	-1.40	-1.42	-1.66	-2.01
KG Frosted Flakes 20oz.	-0.93	-1.01	-1.16	-1.08	-1.10	-1.29	-1.56
PO Hny Bunches Oats Reg 16oz.	-1.00	-1.09	-1.25	-1.17	-1.18	-1.39	-1.67
QK Life Regular 21oz.	-0.98	-1.07	-1.23	-1.15	-1.17	-1.37	-1.65
SB Raisin Bran 20oz.	-0.64	-0.70	-0.80	-0.75	-0.76	-0.89	-1.08
Avg of All 64 Products	-1.20	-1.31	-1.51	-1.41	-1.43	-1.67	-2.02
				Price-Cost Margin			
GM Cheerios 15oz.	84%	77%	67%	72%	71%	60%	50%
KG Frosted Flakes 20oz.	108%	99%	86%	92%	91%	78%	64%
PO Hny Bunches Oats Reg 16oz.	100%	92%	80%	86%	85%	72%	60%
QK Life Regular 21oz.	102%	93%	81%	87%	86%	73%	61%
SB Raisin Bran 20oz.	156%	143%	124%	133%	131%	112%	93%
Avg of All 64 Products	88%	81%	70%	75%	74%	63%	52%

***, **, * denote significance at 1%, 5%, and 10%, respectively, based on White (1980) heteroskedasticity-robust standard errors.
[a]Excludes merchandising and local advertising variables.

Figure 2.8: CX Estimates, Product-Level and Market-Level, Aggregated

Dep Var: $\ln(S_{jm}) - \ln(S_{0m})$	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Market Price	-0.256***	-0.269***	-0.304***	-0.388***	-0.438***	-0.495***	-0.781***
(Standard Error)	(0.0007)	(0.0084)	(0.0007)	(0.0013)	(0.0014)	(0.0044)	(0.0038)
Included Product-Level Variables							
Product Fixed Effects	No	No	No	No	Yes	Yes	Yes
7 Product Characteristics	No	Yes	Yes	Yes	No	No	No
30 Additional Product Chars	No	No	No	Yes	No	No	No
Included Market-Level Variables	Yes	Yes	Yes	Yes	Yes	Yes	Yes[a]
Product-Level Instruments							
Top 5 Rivals' Characteristics	No	Yes	Yes	Yes	No	No	No
Market-Level Instruments							
Average Price in Other Cities	Yes	Yes	No	Yes	No	Yes	Yes
Lagged Price (weeks t-1 to t-8)	No	No	Yes	Yes	Yes	No	No
Observations	1,186,098	1,186,098	1,118,316	1,118,316	1,118,316	1,186,098	1,186,098
F	1,400,000	180,000	200,000	56,572	74,898	33,565	53,318
AIC/Obs	1.804	1.695	1.618	1.439	1.374	1.457	2.092
	Price Elasticity						
GM Cheerios 15oz.	-0.88	-0.93	-1.05	-1.34	-1.51	-1.71	-2.70
KG Frosted Flakes 20oz.	-0.73	-0.77	-0.87	-1.11	-1.25	-1.41	-2.22
PO Honey Bunches of Oats Reg 16oz.	-0.75	-0.79	-0.90	-1.14	-1.29	-1.46	-2.30
QK Life Regular 21oz.	-0.73	-0.77	-0.87	-1.11	-1.25	-1.41	-2.23
SB Raisin Bran 20oz.	-0.46	-0.48	-0.55	-0.70	-0.79	-0.89	-1.40
Avg of All 64 Products	-0.89	-0.93	-1.06	-1.35	-1.52	-1.72	-2.71
	Price-Cost Margin						
GM Cheerios 15oz.	113%	108%	95%	75%	66%	59%	37%
KG Frosted Flakes 20oz.	137%	131%	115%	90%	80%	71%	45%
PO Honey Bunches of Oats Reg 16oz.	133%	126%	112%	87%	77%	69%	43%
QK Life Regular 21oz.	137%	130%	115%	90%	80%	71%	45%
SB Raisin Bran 20oz.	217%	207%	183%	143%	127%	112%	71%
Avg of All 64 Products	119%	113%	100%	78%	69%	61%	39%

***, **, * denote significance at 1%, 5%, and 10%, respectively, based on White (1980) heteroskedasticity-robust standard errors.
The F statistic reported tests that the first stage coefficients for the excluded instruments are jointly equal to zero. AIC/Obs is for the reduced form.
[a]Excludes merchandising and local advertising variables.

Figure 2.9: IV Estimates, Product-Level and Market-Level

Dep Var: $\ln(S_{jm})$ - $\ln(S_{0m})$	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Market Price	-0.256***	-0.255***	-0.276***	-0.264***	-0.362***	-0.261***	-0.475***
(Standard Error)	(0.0033)	(0.0037)	(0.0040)	(0.0083)	(0.0129)	(0.0138)	(0.0133)
Included Product-Level Variables							
Product Fixed Effects	No	No	No	No	Yes	Yes	Yes
7 Product Characteristics	No	Yes	Yes	Yes	No	No	No
30 Additional Product Chars	No	No	No	Yes	No	No	No
Included Market-Level Variables	Yes	Yes	Yes	Yes	Yes	Yes	Yes[a]
Product-Level Instruments							
Top 5 Rivals' Characteristics	No	Yes	Yes	Yes	No	No	No
Market-Level Instruments							
Average Price in Other Cities	Yes	Yes	No	Yes	No	Yes	Yes
Lagged Price (quarter t-1)	No	No	Yes	Yes	Yes	No	No
Observations	31,478	31,478	28,854	28,854	28,854	31,478	31,478
F	160,000	19,017	13,316	3,341	6,377	4,541	6,124
AIC/Obs	1.033	0.864	0.861	0.466	0.302	0.321	0.723
				Price Elasticity			
GM Cheerios 15oz.	-0.77	-0.77	-0.83	-0.79	-1.09	-0.78	-1.43
KG Frosted Flakes 20oz.	-0.60	-0.59	-0.64	-0.62	-0.84	-0.61	-1.11
PO Honey Bunches of Oats Reg 16oz.	-0.64	-0.64	-0.69	-0.66	-0.91	-0.65	-1.19
QK Life Regular 21oz.	-0.63	-0.63	-0.68	-0.65	-0.89	-0.64	-1.17
SB Raisin Bran 20oz.	-0.41	-0.41	-0.45	-0.43	-0.58	-0.42	-0.77
Avg of All 64 Products	-0.77	-0.77	-0.83	-0.80	-1.09	-0.79	-1.44
				Price-Cost Margin			
GM Cheerios 15oz.	130%	131%	121%	126%	92%	128%	70%
KG Frosted Flakes 20oz.	167%	168%	155%	162%	118%	164%	90%
PO Honey Bunches of Oats Reg 16oz.	156%	156%	145%	151%	110%	153%	84%
QK Life Regular 21oz.	158%	159%	147%	153%	112%	155%	85%
SB Raisin Bran 20oz.	242%	243%	225%	235%	171%	238%	131%
Avg of All 64 Products	136%	137%	126%	132%	96%	134%	73%

***, **, * denote significance at 1%, 5%, and 10%, respectively, based on White (1980) heteroskedasticity-robust standard errors.
The F statistic reported tests that the first stage coefficients for the excluded instruments are jointly equal to zero. AIC/Obs is for the reduced form.
[a]Excludes merchandising and local advertising variables.

Figure 2.10: IV Estimates, Product-Level and Market-Level, Aggregated

	Based on CX[a]			Based on TSLS[b]		
	SPF	MPF	JOM	SPF	MPF	JOM
GM Cheerios 15oz.	48%	56%	84%	59%	68%	102%
KG Frosted Flakes 20oz.	59%	70%	103%	71%	85%	124%
PO Honey Bunches of Oats Reg 16oz.	57%	59%	99%	69%	72%	120%
QK Life Regular 21oz.	58%	59%	103%	71%	72%	124%
SB Raisin Bran 20oz.	93%	95%	163%	112%	115%	197%
Avg of General Mills Products	44%	51%	77%	53%	62%	93%
Avg of Kellogg Products	51%	61%	89%	61%	74%	108%
Avg of Post Products	53%	56%	94%	65%	68%	114%
Avg of Quaker Products	53%	54%	93%	64%	65%	113%
Avg of Store Brand Products	84%	86%	148%	102%	104%	180%
Avg of All 64 Products	51%	58%	89%	61%	70%	108%

SPF = Single-Product Firm, MPF = Multi-Product Firm, JOM = Joint Ownership Monopoly.
[a] CX margins based on estimate from Table VII, column 6 (or columns 4 & 5).
[b] TSLS margins based on estimate from Table VIII, column 6.

Figure 2.11: Price-Cost Margins

	Based on CX[a]			Based on TSLS[b]		
	SPF	MPF	JOM	SPF	MPF	JOM
GM Cheerios 15oz.	34%	40%	60%	37%	43%	65%
KG Frosted Flakes 20oz.	41%	50%	73%	45%	54%	79%
PO Honey Bunches of Oats Reg 16oz.	40%	42%	70%	43%	45%	76%
QK Life Regular 21oz.	41%	42%	73%	45%	45%	79%
SB Raisin Bran 20oz.	66%	67%	115%	71%	73%	125%
Avg of General Mills Products	31%	36%	54%	33%	39%	59%
Avg of Kellogg Products	36%	43%	63%	39%	47%	68%
Avg of Post Products	38%	40%	67%	41%	43%	72%
Avg of Quaker Products	37%	38%	66%	41%	41%	72%
Avg of Store Brand Products	60%	61%	105%	65%	66%	114%
Avg of All 64 Products	36%	41%	63%	39%	44%	68%

SPF = Single-Product Firm, MPF = Multi-Product Firm, JOM = Joint Ownership Monopoly.
[a] CX margins based on estimate from Table VII, column 7 (excludes merchandising and local advertising).
[b] TSLS margins based on estimate from Table VIII, column 7 (excludes merchandising and local advertising).

Figure 2.12: Price-Cost Margins, Excluding Local Promotions

Dep Var: ln(S_{jm}) - ln(S_{0m})	Before		After		
Market Price	-0.723***		-0.687***		
(Standard Error)	(0.0074)		(0.0273)		
	Mean	StDev	Mean	StDev	
Market Price x Product Dummy Vars.	0.073	0.179	0.070	0.189	
Product Fixed Effects	Yes		Yes		
Market-Level Conditioning Instruments	Yes		Yes		
Observations	918,134		182,430		
R-Squared	0.787		0.793		
RMSE	0.386		0.379		
	Average Price-Cost Margin				Change
General Mills	66.6%		70.2%		3.6%
Kellogg	73.0%		70.2%		-2.8%
Post	60.2%		64.4%		4.2%
Quaker	60.8%		62.7%		1.9%
Store Brand	56.7%		56.8%		0.0%
All 64 Products	67.3%		68.2%		0.9%

Before/After cutoff is 7/15/2007, the median date marking when General Mills products changed box size.
Margins are calculated based on multi-product firm (MPF) pricing.

Figure 2.13: Price-Cost Margins Before vs. After Downsizing

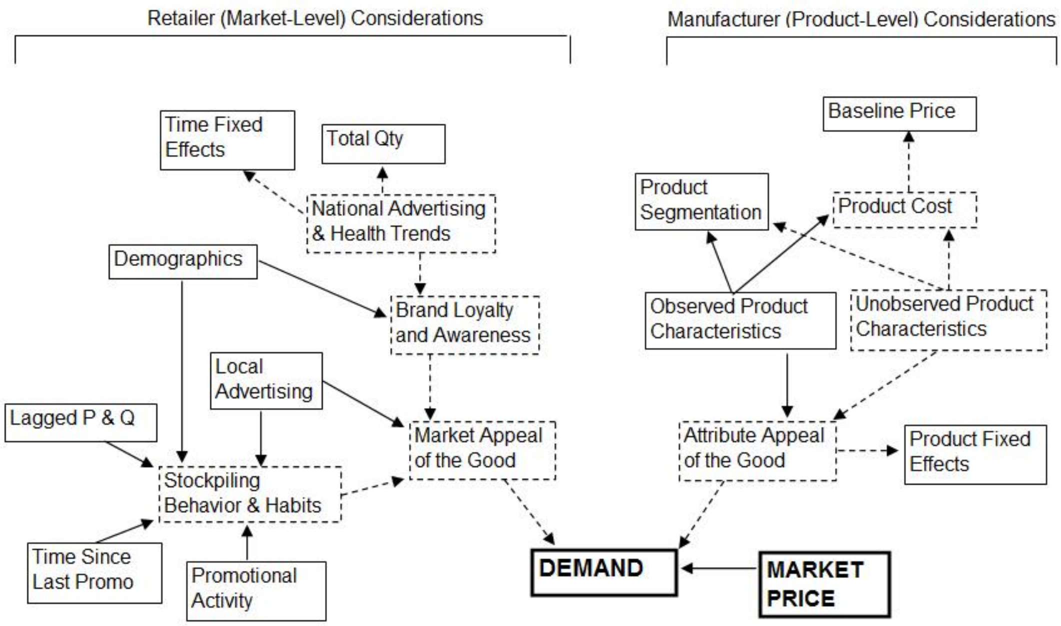


Figure 2.14: Path Diagram - Demand

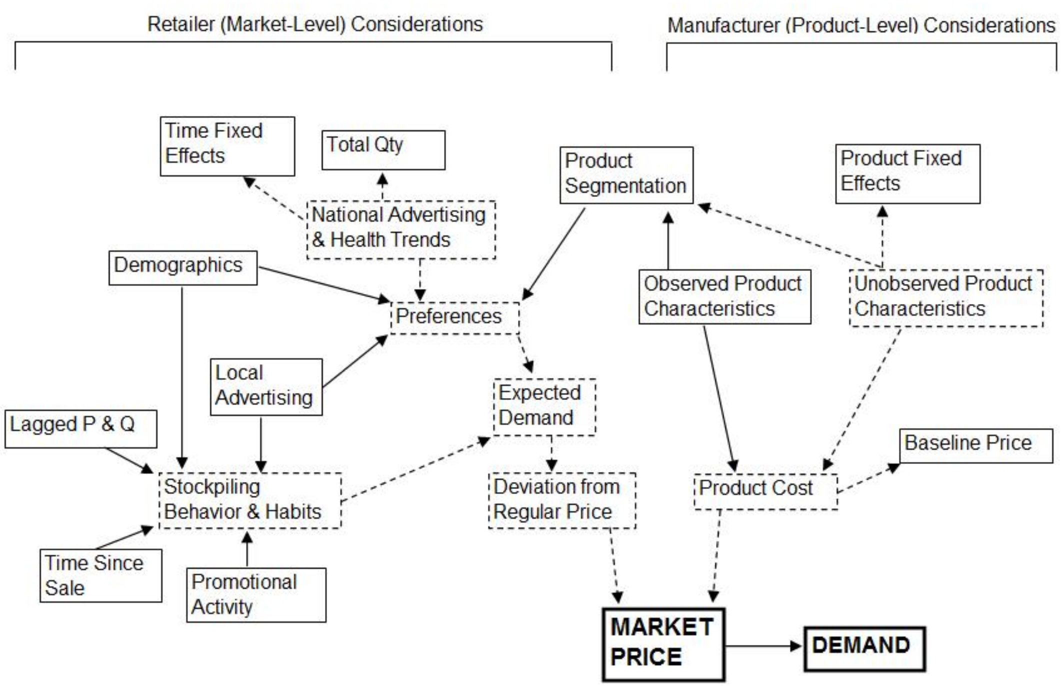


Figure 2.15: Path Diagram - Price

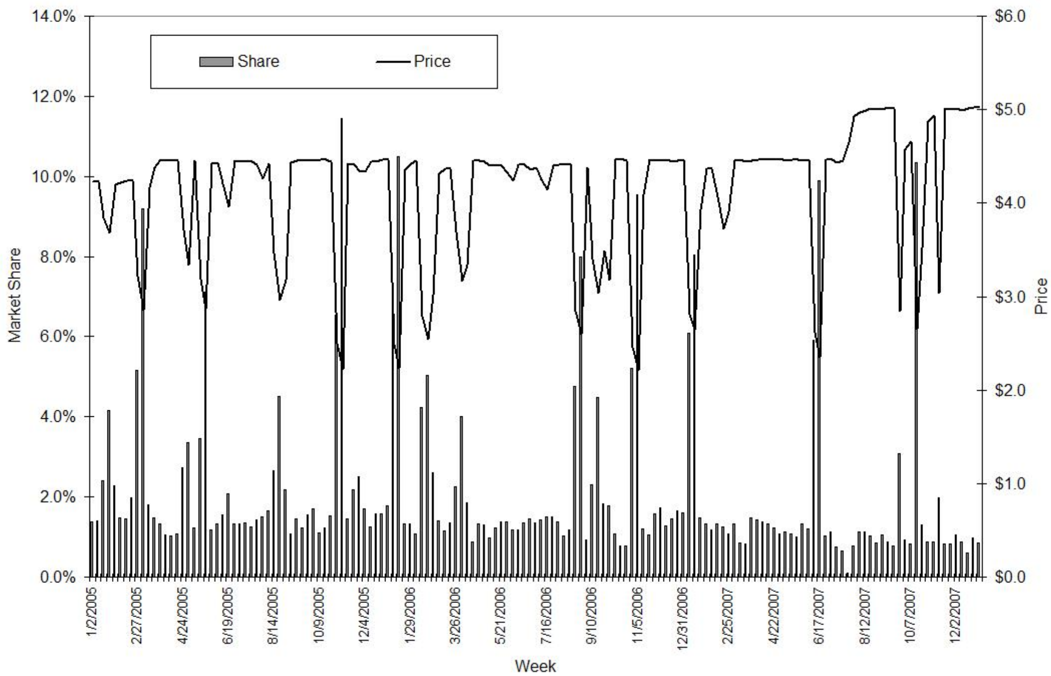


Figure 2.16: Share and Price - GM Cheerios

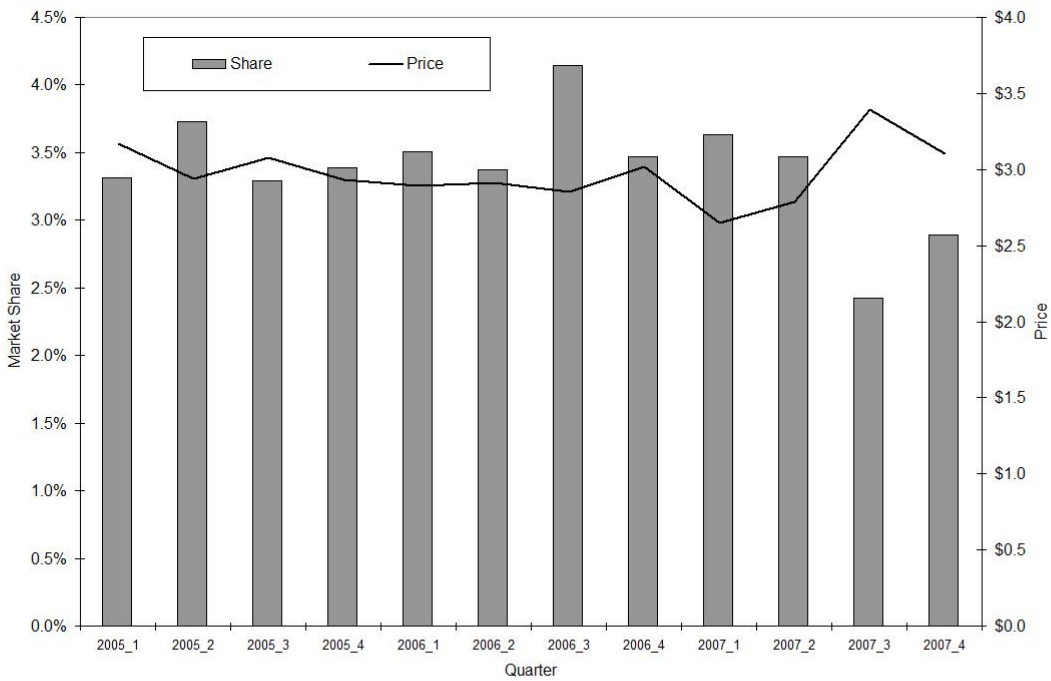


Figure 2.17: Share and Price - GM Cheerios, Aggregated

Chapter 3

Identification and Estimation of Price Effects with a Nonseparable Logit Demand Model

3.1 Introduction

We examine nonparametric estimation of the average marginal effect of price on demand allowing for price heterogeneity. That is, we consider the possibility that price effects, defined as the impact of a change in price on quantity demanded, depend on price, rather than assuming the price effect is a single coefficient estimate. The prototypical indirect utility function in nearly all logit models of demand takes a linear-separable parametric specification in which utility is a function of price, observable and unobservable product characteristics, and a type-I extreme value distributed error term (e.g., Berry, 1994, and Nevo, 2001). We instead allow for nonseparability of price and unobservables by specifying a more flexible indirect utility function that may be nonlinear, and for which observable and unobservable variables may not be separable.

The specification and estimation of demand for differentiated products is particularly important for research in industrial organization (IO) and quantitative marketing. The logit class of discrete choice demand models are widely used for estimating price effects when market-level data are available. Notable studies include Berry, Levinsohn, and Pakes (1995), Nevo (2001), Villas-Boas (2007), Besanko, Gupta, and Jain (1998), and Villas-Boas and Winer (1999), among others. These studies often rely on identification assumptions involving instrument exogeneity. Because demand functions may be nonseparable in the absence of strong restrictions, standard instrumental variables approaches do not help with identification. We instead employ a conditional form of exogeneity for identification: price is independent of unobservable demand drivers conditional on a set of selected conditioning instruments. Conditional exogeneity (conditional independence) has been at the center of recent theoretical research in econometrics, including Altonji and Matzkin (2005), White and Chalak (2006), and Hoderlein and Mammen (2007).

In addition to relaxing the restrictive linear-separable functional forms and identifying assumptions of previous studies, we also employ a rich set of supermarket scanner data on ready-to-eat boxed cereals. The scanner data, a subset of the data employed in Chapters 1 and 2, is relatively disaggregated and contains detailed information on price, quantity, and local promotions for a large set of cereal products. For each of the goods in the sample, the variables are available at the city-supermarket chain aggregation level, with a weekly frequency for the three years ending 2007.

3.2 Demand Model

We employ the conditional logit demand model that has gained popularity since McFadden (1974), especially in IO. The conditional logit model suffers from the well-known independence of irrelevant alternatives problem, which can be addressed with computationally burdensome variants of logit when market-level data are available.[1] However, because of the nonparametric estimation procedures employed here to address nonlinearity and nonseparability of the indirect utility function, the computational burdens are already very high. Thus, we focus on the conditional logit model, noting that potentially unreasonable cross-price effects

[1] Such as the random coefficients logit model of Berry, Levinsohn, and Pakes (1995).

(e.g., cross-price elasticities) arising from this model do not affect our analysis related to price heterogeneity of the own-price effect. That is, we assume that the presence or absence of price heterogeneity does not depend on the underlying type of data generating process in the logit class of demand models.

We consider consumer $i \in \mathcal{I}$ $(\equiv \{1,2,..,I\})$ who has a random indirect utility u_{ikm} from choosing product $k \in \{0\} \cup \mathcal{K} (\equiv \{1,2,..,K\})$ in market $m \in \mathcal{M}$ $(\equiv \{1,2,..,M\})$.[2] We impose the following assumption about $u_{ikm.}$

Assumption A.1 For all $i \in \mathcal{I}$, $k \in \mathcal{K}$, and $m \in \mathcal{M}$, indirect utility is given by

$$\begin{aligned} u_{ikm} &= d\left(p_{km}, \xi_{km}\right) + \varepsilon_{ikm} \\ &\equiv V\left(lp_{km}, \xi_{km}\right) + \varepsilon_{ikm}. \end{aligned} \tag{3.1}$$

where p_{km} is the price of good k in market m, and $lp_{km} \equiv \ln(p_{km})$; ξ_{km} and ε_{ikm} represent market-level/product-level factors and individual level factors that drive preferences, respectively.

Here, we assume that the functional forms of d and V do not depend on i, k or m. The market-level and product-level heterogeneity is captured by ξ_{km}, while individual heterogeneity is captured by ε_{ikm}. ε_{ikm} can also be interpreted as a disturbance term representing an individual-level shock. We allow ξ_{km} (or components of ξ_{km}) and ε_{ikm} to be unobservable. In A.1, we impose the additivity of individual heterogeneity to the indirect utility. This additivity assumption simplifies the derivation of market share, since we only have market-level data. Nevertheless, we do allow arbitrary interactions between price and market-level/product-level factors by using a general nonseparable function V. Furthermore, here the dimension of ξ_{km} can be finite or countably infinite. Since here we impose additivity of ε_{ikm}, and ε_{ikm} can be a function of unobservables, it is without loss of generality to assume ε_{ikm} is a scalar.

Note that the following specification

$$u_{ikm} = V\left(lp_{km}\right) + \xi_{km} + \varepsilon_{ikm} \tag{3.2}$$

[2] A market is defined as a city-supermarket-week combination in the scanner data.

is a special case of our A.1. Also note that in most logit demand models, indirect utility is described by

$$u_{ikm} = \alpha \cdot lp_{km} + \xi_{km} + \varepsilon_{ikm}, \tag{3.3}$$

which is a special case of equation (3.2). This assumes that the effect of price on demand is a single estimate, α, that does not vary with price.[3]

We impose the following assumption about the distribution of ε_{ikm}.

Assumption A.2: (i) $(p_{km}, \xi_{km}) \perp \varepsilon_{ikm}$ for all $i \in \mathcal{I}$, $k \in \{0\} \cup \mathcal{K}$, and $m \in \mathcal{M}$. (ii) ε_{ikm} is a mean-zero type-I extreme value distributed random variable for all $i \in \mathcal{I}$, $k \in \{0\} \cup \mathcal{K}$, and $m \in \mathcal{M}$.

Assumptions A.1 and A.2 yield the demand function of interest in this paper. Because consumer-level data are not available here, we aggregate equation (3.1) up to the market-level representative consumer. We consider a representative consumer i_0, in a representative market m_0. She will choose good k_0 over good k for all $k \neq k_0$ if $U_{k_0 m_0} > U_{k m_0}$ $\forall k \neq k_0$. We denote $lp_{m_0} \equiv (lp_{m_0 0}, lp_{m_0 1}, lp_{m_0 2}, .., lp_{m_0 K})$ and $\xi_{m_0} \equiv (\xi_{m_0 0}, \xi_{m_0 1}, \xi_{m_0 2}, .., \xi_{m_0 K})$. We calculate the probability (s_{k_0}) for the consumer to choose product k_0 in this market given the price level $lp_{m_0} = \overline{lp} \left(\equiv \left(\overline{lp}_0, \overline{lp}_1, \overline{lp}_1, .., \overline{lp}_K\right)\right)$ and the product and market level factor $\xi_{m_0} = \overline{\xi} \left(\equiv \left(\overline{\xi}_0, \overline{\xi}_1, \overline{\xi}_2, .., \overline{\xi}_K\right)\right)$, i.e.,

$$s_{k_0} = \Pr(u_{ik_0m_0} > u_{ikm_0}, \forall k \neq k_0 |\, lp_{m_0} = \overline{lp}, \xi_{m_0} = \overline{\xi}).$$

Under A.1 and A.2, it is easy to show

$$s_{k_0} = \frac{\exp\left(V(\overline{lp}_{k_0}, \overline{\xi}_{k_0})\right)}{\sum_{k \in \{0\} \cup \mathcal{K}} \exp\left(V(\overline{lp}_k, \overline{\xi}_k)\right)}. \tag{3.4}$$

Following Berry (1994), $V\left(\overline{lp}_0, \overline{\xi}_0\right)$ is normalized to take an assumed value of zero, then the share of the "outside good," which is the option of a consumer not purchasing one of the $k = 1, 2, ..., K$ inside goods, is given by

$$s_0 = \frac{1}{\sum_{k \in \{0\} \cup \mathcal{K}} \exp\left(V(\overline{lp}_k, \overline{\xi}_k)\right)}. \tag{3.5}$$

[3]See, for example, Berry (1994) and Chapter 1.

Dividing equation (3.4) by (3.5), and taking natural logs gives

$$\ln(s_{k_0}) - \ln(s_0) = V(\overline{lp}_{k_0}, \overline{\xi}_{k_0}).$$

We denote

$$ls_{k_0} \equiv \ln(s_{k_0}) - \ln(s_0),$$

where ls_{k_0} is the natural logarithm of the market share of product k_0 relative to the outside good. We thus obtain a tractable demand function:

$$ls_{k_0} = V(\overline{lp}_{k_0}, \overline{\xi}_{k_0}). \tag{3.6}$$

The demand equation given by (3.6) is the structural equation of interest here. The function $V(.)$ is general, nonparametric, and nonseparable, as it allows for the possibility that the effect of price on demand depends on price, and it allows for the possibility that price interacts with unobservables. Note that in the linear specification (equation (3.3)), we have

$$ls_{k_0} = \alpha \overline{lp}_{k_0} + \overline{\xi}_{k_0}.$$

We can also derive the own price elasticity of demand e_{k_0}:

$$e_{k_0} = \frac{\partial V(\overline{lp}_{k_0}, \overline{\xi}_{k_0})}{\partial \overline{lp}_{k_0}} \cdot [1 - \bar{s}_{k_0}] = \eta\left(\overline{lp}_{k_0}, \overline{\xi}_{k_0}\right) \cdot [1 - \bar{s}_{k_0}],$$

where

$$\eta\left(\overline{lp}_{k_0}, \overline{\xi}_{k_0}\right) \equiv \frac{\partial V\left(\overline{lp}_{k_0}, \overline{\xi}_{k_0}\right)}{\partial \overline{lp}_{k_0}},$$

and $\bar{s}_{k_0}$ is the market share of product k_0.

Ideally, we would be interested in how the relative market share ls_{k_0} changes in response to the price changes, i.e., $\eta\left(\overline{lp}_{k_0}, \overline{\xi}_{k_0}\right)$ (or own price elasticity, e_{k_0}). It turns out that it is difficult to identify $\eta(.)$ without strong assumptions (for example, monotonicity) that are difficult to justify. For a detailed discussion of monotonicity and testing procedures, see Su, Hoderlein, and White (2010).[4]

[4]Su, Hoderlein, and White (2010) use the same data as we do to implement their test. They assume monotonicity, and reject exogeneity of the covariates in the cereal demand DGP. Had they instead assumed exogeneity of the of the covariates, they would reject monotonicity of the unobservables in the cereal demand DGP. We have implemented their test on our specifications here. Assuming exogeneity, we reject monotonicity of unobservables.

Furthermore, $\bar{\xi}_{k_0}$ is potentially a large dimension vector, which means that estimators of $\eta(.)$ will suffer from a severe "curse of dimensionality" because the convergence rate is determined by the dimension of (lp_{k_0}, ξ_{k_0}). Therefore, we estimate an average version of $\eta\left(\overline{lp}_{k_0}, \bar{\xi}_{k_0}\right)$ by averaging over unobservable market/product heterogeneity ξ_{m_0,k_0}:

$$\theta\left(\overline{lp}_{k_0}\right) \equiv \int \eta\left(\overline{lp}_{k_0}, \bar{\xi}_{k_0}\right) \cdot dF_{\xi}\left(\bar{\xi}_{k_0}\right),$$

where $F_{\xi}(.)$ is the marginal cumulative distribution function (CDF) of ξ_{m_0,k_0}. This is the average marginal effect defined in Blundell and Powell (2003). Note that in the linear specification (equation (3.3)), $\theta(.) = \alpha$, which is examined in detail in Chapter 2.

3.3 Identification and Estimation

In this section, we focus on the identification and estimation of $\theta(.)$. A concern in estimating demand using market data is price endogeneity, defined as the dependence of price with unobservable drivers of demand: $p_{jm} \not\perp \xi_{jm}$. In an ideal world, firms randomly choose products to change price by random increments. This would be tantamount to a randomized experiment in which the researcher could recover the causal treatment effect of interest due to the exogeneity of prices. Of course, prices and unobservable drivers of demand can be correlated for many reasons. We follow Chapter 2 by positing that there are underlying common causes that drive both demand and price at the product and market levels, which are the confounding factors that render market prices endogenous.

By conditioning on variables that act as proxies for these unobserved confounding effects, we claim that prices are conditionally exogenous. We refer the reader to White and Chalak (2006) and Chalak and White (2010) for a precise definition of conditioning instruments, w_{km}, which here are proxies for the confounding variables that drive both price and demand. In Section 4.2, below, we discuss the conditioning instruments that we use to justify conditionally exogenous market prices.

Virtually all studies examining demand estimation for differentiated products with scanner data rely on standard instrument exogeneity conditions for identification. See, for example, Berry, Levinsohn, and Pakes (1995), Hausman (1997), Nevo (2001), Villas-Boas (2007), and Villas-Boas and Winer (1999), just to name a few. Standard instrument exogeneity assumptions require that excluded instruments be independent of the unobserved drivers of demand, yet it is not difficult in most cases to find reasons for why assumed exogenous instruments, such as the proxies for cost shifters used by Hausman (1997) and Nevo (2001), fail to be uncorrelated with unobservables.[5]

Moreover, nonseparable specifications, such as demand equation (3.6), cannot benefit from standard instrument exogeneity assumptions for identification, even if exogenous instruments were readily available. For example, Hahn and Ridder (2008) argue that in a general nonseparable system, the conditional moment type of restriction does not identify any structural object.

3.4 Identification

We rely on the following conditional independence assumption necessary for identification of $\theta(.)$.

Assumption A.3: (i) For $k \in \mathcal{K}$, and $m \in \mathcal{M}$, $lp_{km} \perp \xi_{km}|w_{km}$, where w_{km} is a d_w dimension random vector. (ii) We observe data $\{ls_{km}, lp_{km}, w_{km}\}_{k\in\mathcal{K},\, m\in\mathcal{M}}$ and $\{ls_{kw}, lp_{kw}, w_{kw}\}_{k\in\mathcal{K},\, m\in\mathcal{M}}$ are i.i.d.

Assumption A.3(i) states that, conditional on a vector of conditioning instruments, price is independent of unobservable demand drivers. In the aforementioned ideal (yet fantasy) world where prices are exogenously determined, the vector of conditioning instruments w_{km} is simply the empty set. In the real world, where many of the factors that drive manufacturers' and retailers' pricing decisions also depend on factors that drive consumers' demand, a variety of variables may

[5] See Chapters 1 and 2 for an extensive examination of the standard instruments used in demand estimation.

be needed to populate w_{km} in order to justify that market prices are conditionally exogenous. Given Assumption 3, we can identify $\theta(.)$ as shown in the next lemma.

Lemma 1: *Under A.1, A.2 and A.3,*

$$\theta(lp) = \int m(x, w) \cdot dF_w(w),$$

where

$$m(lp, w) = \frac{\partial E[ls_{km}| lp_{km} = lp, w_{km} = w]}{\partial lp},$$

and $F_w(.)$ *is the CDF of* w_{km}.

3.4.1 Nonparametric Estimation

The estimator of $\theta(lp)$ is $\hat{\theta}(lp)$, given by

$$\hat{\theta}(lp) = \frac{1}{MK}\sum_{m=1}^{M}\sum_{k=1}^{K}\hat{m}(lp, w_{km}),$$

where $\hat{m}(.)$ is an estimator of $m(.)$. We can rewrite

$$\begin{aligned} m(lp, w) &= \frac{\partial}{\partial lp}\left[\frac{g(lp, w)}{f_{lp,w}(lp, w)}\right] \\ &= \left[\frac{g_1(lp, w) \cdot f_{lp,w}(lp, w) - g(lp, w) \cdot f_{lp,w,1}(lp, w)}{[f_{lp,w}(lp, w)]^2}\right], \end{aligned}$$

where

$$\begin{aligned} g(lp, w) &= \int ls \cdot f_{ls,lp,w}(ls, lp, w) \cdot d(ls) \\ g_1(lp, w) &= \frac{\partial g(lp, w)}{\partial lp}, \\ f_{lp,w,1}(lp, w) &= \frac{\partial f(lp, w)}{\partial lp}, \end{aligned}$$

and $f_{lp,w}(.)$ is the density function of (lp_{km}, w_{km}) and $f_{ls,lp,w}(.)$ is the density of $(ls_{km}, lp_{km}, w_{km})$

The corresponding estimator of $m(.)$ is $\hat{m}(.)$:

$$\hat{m}(lp,w) = \left[\frac{\hat{g}_1(lp,w)\cdot\hat{f}_{lp,w}(lp,w) - \hat{g}(lp,w)\cdot\hat{f}_{lp,w,1}(lp,w)}{\left[\hat{f}_{lp,w}(lp,w)\right]^2}\right],$$

where

$$\begin{aligned}
\hat{f}_{lp,w} &= \frac{1}{MK\cdot h^{1+d_w}}\sum_{m=1}^{M}\sum_{k=1}^{K}\left[\mathbf{K}\left(\frac{lp_{km}-lp}{h}\right)\cdot\prod_{q=1}^{d_w}\mathbf{K}\left(\frac{w_{km}^{(q)}-w^{(q)}}{h}\right)\right],\\
\hat{g} &= \frac{1}{MK\cdot h^{1+d_w}}\sum_{m=1}^{M}\sum_{k=1}^{K}\left[\mathbf{K}\left(\frac{lp_{km}-lp}{h}\right)\cdot\prod_{q=1}^{d_w}\mathbf{K}\left(\frac{w_{km}^{(q)}-w^{(q)}}{h}\right)\cdot ls_{km}\right],\\
\hat{f}_{lp,w,1} &= \frac{-1}{MK\cdot h^{2+d_w}}\sum_{m=1}^{M}\sum_{k=1}^{K}\left[\mathbf{K}'\left(\frac{lp_{km}-lp}{h}\right)\cdot\prod_{q=1}^{d_w}\mathbf{K}\left(\frac{w_{km}^{(q)}-w^{(q)}}{h}\right)\right],\\
\hat{g}_1 &= \frac{-1}{MK\cdot h^{2+d_w}}\sum_{m=1}^{M}\sum_{k=1}^{K}\left[\mathbf{K}'\left(\frac{lp_{km}-lp}{h}\right)\cdot\prod_{q=1}^{d_w}\mathbf{K}\left(\frac{w_{km}^{(q)}-w^{(q)}}{h}\right)\cdot ls_{km}\right],
\end{aligned}$$

where $\mathbf{K}$ is a kernel function, $\mathbf{K}'$ is the derivative of $\mathbf{K}$, and h is the bandwidth. We use a six-order Gaussian kernel. As is well known, non-parametric kernel estimation is sensitive to the bandwidth choice. Thus, we try a range of bandwidths near the bandwidth chosen by leave-one-out cross-validation.

3.4.2 Conditioning Instruments

Here, we examine the variables designated as conditioning instruments, w_{km}, that make prices conditionally exogenous in Assumption A.3(i). Recall that ξ_{km} captures drivers of consumer demand, some of which are directly observable, while others are unobservable and must be proxied for with observable variables. We follow Chapter 2 here by proposing that ξ_{km} consists of product-level and market-level factors that also drive price, and are thus sources of confounding that render market prices endogenous. By conditioning on these factors (or proxies for them), we are able to identify the causal effect of market price on demand by relying on the conditional independence assumption $p_{km} \perp \xi_{km}|w_{km}$.

Product-Level Confounding

We consider two sources or "levels" at which confounding may occur rendering market prices endogenous. The first level is at the product, indexed by k. The concern here is that factors that drive the demand of good k also drive the price of good k. If these factors are direct drivers and are observable, then they may be included in w_{km}.

In the ready-to-eat cereal product segment, for example, product quality is a determinant of consumer demand, and certain product attributes, such as the vitamin/mineral content or type of ingredients composing the cereal, drive the marginal cost of production, which in turn determines the price of the good. Simply put, higher quality cereals cost more to produce, thus have higher price, but at the same time the quality of the cereal affects consumers' preferences, which is the source of confounding that prevents causal identification of price effects.

One possibility is that all product characteristics affect both a good's demand and price, in which case, for example, sugar content, fiber, density, ingredients, vitamin/mineral content, shape, among numerous other attributes ought to comprise the conditioning instruments, w_{km}. We encounter two practical problems here. First, the researcher may not have collected a sufficiently detailed set of product characteristics. Second, even if available, the number of product characteristics needed may be too large, rendering the nonparametric estimation techniques proposed earlier infeasible.[6] Both of these setbacks may be well-remedied by relying on a proxy variable that captures the information contained in the comprehensive set of product characteristics needed for identification.

Chapter 2 proposes using the "baseline price" of a good as a predictive proxy for over 30 cereal product characteristics. The baseline price of a good is a measure of central tendency for the price of good j, $p_k = M^{-1} \sum_m p_{km}$. Consider decomposing the price of good k in market m into two parts: $p_{km} = p_k + \Delta p_{km}$, where p_k does not vary across markets but Δp_{km} does. Averaging over markets

[6] This also relates to using product fixed effects in place of product characteristics.

yields
$$M^{-1}\sum_m p_{km} = M^{-1}\sum_m p_k + M^{-1}\sum_m \Delta p_{km},$$
and rearranging gives $p_k = M^{-1}\sum_m p_{km}$. The baseline price captures inter-product variation in product characteristics, both observed and unobserved. Because product attributes, such as a cereal's sweetness, texture, healthfulness, and quality, is driven by product characteristics, and because product characteristics drive manufacturers' marginal cost, the baseline price ought to be a reasonable conditioning instrument as it captures approximately the information contained in the complete set of product characteristics and/or product-level fixed effects.

Moreover, both observed product characteristics and the baseline price can together comprise the conditioning instruments addressing product-level price endogeneity. To the extent that a small set of observable product characteristics do not adequately remedy product-level endogeneity, the baseline price may contain additional information contributing to identification. Chapter 2 finds that when the baseline price and a few key observed product characteristics are used together as the conditioning instrument set, the demand estimates of a linear-separable conditional logit model are approximately the same as when product-specific dummy variables are used to control for product-level unobservables. We examine this point in the empirical section of this paper by using cereal density, calories, fiber, sugar, and baseline price as our product-level conditioning instruments.

Market-Level Confounding

The second source of price endogeneity to consider is at the market level. The concern here is that market-level unobservable factors that determine consumers' demand for good k also influence the price of good k in that market. A market m is defined to be a city-supermarket-week combination. Factors that affect market demand for a good, such as brand awareness, consumer preferences, and stockpiling behavior, may also drive market prices, and are thus market-level confounding effects that prevent identification in the demand system.

Consumers' preferences and brand awareness in a particular market depend on the demographics of the geographic location, local advertising, and national advertising. These factors may also drive market prices in particular cities over time. The extent to which these variables are observable or not by the researcher is dictated by the available data set. In the context of supermarket scanner data, local advertising variables may be available, but data related to health trends and national advertising are typically not observed. We follow Chapter 2 and exploit the panel structure of the scanner data to obtain a response variable that captures information contained in an aggregate demand shock. For example, if national advertising, health trends, or any other aggregate demand shock affect consumers' preferences, then the quantity purchased of the product segment (or individual good) ought to be affected in all geographic locations. Thus the quantity of a good contemporaneously purchased in all other cities is a predictive proxy for an aggregate demand shock in a particular city.

Next, consider dynamics in pricing and consumer demand. For example, if the retailer temporarily reduces the price of a good periodically to capture welfare from low-valuation consumers, and if consumers tend to be the type to stockpile the good in low-price time periods, then confounding may exist. A few variables that drive stockpiling (or non-stockpiling) behavior of consumers and retailers' pricing decisions include local advertising, and lagged prices and quantities. We thus include these variables, along with income[7] and quantity sold in other cities, in our conditioning instrument set, w_{km}, in order to justify Assumption 3.

3.5 Data and Results

We estimate demand for a large sample of ready-to-eat breakfast cereals. Currently, the four major nationally branded cereal manufacturers operating in the U.S. are General Mills, Kellogg, Post, and Quaker. A fifth important source of cereal sales are store brand products, which are lower-priced versions of the nationally branded cereals.

[7]In order to keep the number of conditioning instruments small, we include only income, as it is usually correlated with other demographic variables.

The breakfast cereal industry has been the subject of extensive investigation by IO researchers. Two notable studies that examine demand estimation in the cereal industry are Hausman (1997) and Nevo (2001), which rely on standard instrument exogeneity assumptions for identification.[8] In contrast, we rely on the conditional exogeneity assumption $p_{km} \perp \xi_{km}|w_{km}$ for identification.

3.5.1 Data

The primary data set utilized here is supermarket scanner data from Information Resources, Inc. (IRI). The data set consists of variables measuring price, quantity, and local promotional variables for 150 of the top-selling cereal stock-keeping units (SKUs) for January 2005 through December 2007. We consider $K = 64$ products as the inside goods, while the remaining 86 products are set as the outside good.[9] The scanner data set has a panel structure, where the time dimension consists of a weekly frequency (2005 through 2007) and the individual dimension consists of a supermarket retail chain located in a geographic market. This data is a subset of that used in Chapter 2. Because of the computational considerations here, we focus on four city-supermarkets that cover four major U.S. regional areas.

Confidentiality agreements prevent us from revealing the supermarket names. We identify the four supermarket retail chains as R1, R2, R3, and R4. The sample thus consists of Chicago-R1, Houston-R2, Philadelphia-R3, and Sacramento-R4. Each of the geographic areas defined by IRI are similar to the Census Bureau's definition of a metropolitan statistical area or combined MSA. Note that unlike Hausman (1997) and Nevo (2001), this market-level scanner data is not heavily aggregated, and thus does not suffer from serious aggregation bias. In Chapter 2, demand is estimated using both disaggregated (city-supermarket-week) data

[8]Specifically, they rely on what have come to be known as "Hausman instruments." The Hausman instrument for the price of brand A in city B is the price of brand A in cities not B. See Chapters 1 and 2 for a detailed examination of this identification scheme.

[9]The set of 86 cereal products comprising the outside good are products that either had data quality issues, were missing price information, or had very high price correlations with one or more of the 64 inside products.

and aggregated (city-quarter) data, and finds that there is considerable aggregation bias that causes the aggregated estimates to be much less elastic than the estimates based on disaggregated scanner data.

Finally, for each cereal, we use product characteristics data on density, calories, fiber, and sugar. Because using product fixed effects would be far too computationally burdensome in the nonparametric estimation framework presented here, we rely on a good's baseline price, defined earlier, to act as a predictive proxy for product characteristics that ought to be included in addition to the four aforementioned, but are not included due to preserving a parsimonious specification. Consistent with Section 4.2, we employ the following conditioning instruments: local promotional intensity, lagged price, lagged market share, quantity in all other cities, density, calories, fiber, sugar, baseline price, and market wage.

Descriptive Statistics

Figure 3.1 presents aggregate market share and pricing information[10] on the 150 top-selling products (or stock keeping units, SKUs) for the four city-supermarket chains in our sample between 2005 and 2007. Kellogg and General Mills dominate the four markets, which is also true of the full sample found in Chapter 2. The results in Figure 2.1 are similar to the full sample of 121 city-supermarket chains in Chapter 2, except that the market share gap between Kellogg and General Mills is slightly wider here. All four of the major manufacturers price well above the store brand products, selling at a 27 to 63 percent premium, while enjoying higher market shares.

Figure 3.4 presents the aggregated monthly market shares for the four city-supermarket chains in the sample. The average market share of the store brand products during the 36 months ending December 2007 is 6.2 percent with a standard deviation of only 0.50 percent. Post and Quaker's market shares vary more than store brands, but are still relatively less volatile than the two biggest manufacturers, Kellogg and General Mills. The average monthly market share for Post and Quaker are 14 percent and 8 percent, with standard deviations of 2 percent

[10] Market share is based on pounds of cereal sold, and price is dollars per pound.

and 1.5 percent, respectively. In contrast, Kellogg and General Mills have market shares that vary more over time than the other manufacturers. Kellogg has an average market share of 42 percent with a standard deviation of 3 percent, and General Mills has an average market share of 30 percent with a standard deviation of 3.0 percent. Finally, Figure 3.4 indicates that not only do the market shares of Kellogg and General Mills vary considerably over time, but that they are nearly mirror images. The correlation of the monthly market share for Kellogg and General Mills is -0.65, indicating that Kellogg's gain in market share is typically General Mills' loss, and vice-versa.

These results are all approximately in line with the market share and pricing trends based on the full data set in Chapter 2. Thus, the four city-supermarket chains that we focus on here appear to be a good representation of the full sample of 121 city-supermarkets. Finally, Figure 3.5 presents the weekly price per pound and market share for General Mills Cheerios 15oz., the best-selling cereal product during 2005-2007, selling in one of the supermarket chains in Chicago. This figure fully demonstrates the degree of disaggregation of the scanner data we use for estimating $\theta(lp)$. Clearly, price and market share are negatively correlated, indicating an expected downward sloping demand for the Cheerios product.

Figures 3.2 and 3.3 present a variety of summary statistics of the data we use for estimating demand. Again, market share is defined as the ratio of pounds of cereal sold for good j in market m to the total amount of cereal sold of all 150 goods in market m. Price is in dollars per pound. Promotion is a variable that is given by $0.25(a_1 + a_2) + 0.50a_3$, where a_1 is the promotional intensity of local advertising, a_2 is the promotional intensity of in-store displays, and a_3 is the promotional intensity of both local advertising and in-store displays for a good.[11]

For each cereal, density is measured as ounces per cup, calories are per cup, fiber and sugar are grams per cup, and the baseline price is the weighted average price of good j across all 121 city-supermarkets in Chapter 2. Wage is the average weekly wage of a given geographic region from the Bureau of Labor

[11]Note that a_1, a_2, and a_3 are mutually exclusive. If good j has $a_2 = 0.45$, this means that good j has in-store displays in a set of supermarkets that represent 45% of all commodity volume in the geographic region under consideration.

Statistics' Quarterly Census of Employment and Wages database. To note, the source of variation for market share, price, and promotion is from products and markets; the source of variation for density, calories, fiber, sugar, and baseline price is product only; the source of variation for wage is geographic and temporal (quarterly).

3.5.2 Estimation Results

In this section, we present results of the nonparametric estimates of $\theta(lp)$. Because market shares are typically small in our sample, examining $\theta(lp)$ is approximately equivalent to examining the own price elasticity. We consider four specifications: (i) a linear-separable model without conditioning instruments; (ii) the linear-separable model with conditioning instruments; (iii) a nonparametric, nonseparable model without conditioning instruments; and (iv) the nonparametric, nonseparable model with conditioning instruments.

The conditioning instruments we include are local promotional intensity, lagged natural log price, lagged natural log market share, natural log of quantity in all other cities, cereal density, fiber, sugar, baseline price, and wage. Following Chapter 2, we define the market share of the outside good ($k = 0$) to be the quantity-based market share of the 86 cereal products that are not included as one of the 64 inside goods.

Chicago-R1

The linear-parametric estimate of $\theta(lp)$ is -1.70 without conditioning instruments. I.e., the average own price elasticity of demand for the 64 cereal products in the sample for Chicago-R1 ("retail supermarket chain one") indicates that a 1% increase in price leads to a 1.7% decrease in purchases of cereal. This result is elastic, but becomes even more elastic when we include the conditioning instruments that remedy the price endogeneity that arises from product-level sources, such as unobserved product quality, and market-level sources, such as stockpiling behavior by consumers. The estimate of $\theta(lp)$ is -2.51 when we include the conditioning instruments, which suggests severe misspecification for the previous case where conditioning instruments are omitted.

Next we examine the nonparametric estimates of $\theta(lp)$. We employ the Gaussian kernel density and bandwidth $h = 1.25$. Figure 3.6 presents the nonparametric estimates of the price effect $\theta(lp)$ without conditioning instruments. Figure 3.7 presents the nonparametric estimate of $\theta(lp)$ including the conditioning instruments. The price effect $\widehat{\theta}$ is reported on the Y-axis while the natural log of price (lp) is reported on the X-axis.

First, it is clear that as log price increases, the price effect (approximately the own price elasticity of demand) becomes less elastic. The linear-parametric estimate of θ with conditioning instruments is -2.51, but the more flexible nonparametric estimates of θ found in Figure 3.7 shows that $\widehat{\theta}$ ranges from -4 to -0.25. This suggests price heterogeneity is present, which means that linear-parametric estimates of demand can be misleading. That is, if the price effect is truly -2.51, we should expect to see a horizontal line at $\widehat{\theta} =$ -2.51 in Figure 3.7, which is clearly not the case.

The increasing nature of the series, whereby price sensitivity becomes less elastic as (log) price increases, potentially addresses a well-known problem in standard linear-parametric specifications of logit demand. High-quality goods that typically enjoy relatively higher prices but lower market shares will necessarily have relatively large own price elasticities compared to, say, mid-quality goods with relatively lower prices but higher market shares. This result directly stems from the price elasticity equation for standard linear-parametric logit models of demand.

However, this result is counterintuitive, as we should expect that relatively high-quality, high-priced goods enjoy higher price-cost margins, not lower ones. For example, a high-quality organic cereal priced at \$6.00 per pound with 0.25% market share, which ought to have a relatively low price elasticity, reasonably commands a higher price-cost margin (e.g., through a Lerner index type of argument) compared to a low or mid-quality cereal priced at \$3.00 per pound with 4% market share, which ought to have a relatively high price elasticity. This is not the case with a linear-parametric logit model of demand.

The nonparametric estimates found in Figures 3.6 and 3.7 addresses this pitfall. As prices increase, the price effect becomes less elastic, thus providing higher price-cost margins according to standard oligopoly models. By relaxing the potentially stringent structure of a linear-parametric logit demand function, we permit the price effect to vary in price, which yields lower price elasticities at higher prices.

Consistent with the linear-parametric estimates, the nonparametric estimates are less elastic when conditioning instruments are left out. We note that a problem that arises in both figures is that the price effect quickly becomes inelastic for a broad range of prices. In Figure 3.7, for example, $\widehat{\theta} \in [-4, -0.25]$, but it would be more realistic for $\widehat{\theta} \in [-4, -1)$ to be consistent with implied price-cost margins for the cereal industry. The result may be due to misspecification of the conditioning instruments. That is, it's possible that the set of conditioning instruments we have included do not sufficiently render price conditionally exogenous. We shall examine this matter more thoroughly in future revisions of this research.

3.6 Conclusions and Further Work

We have examined nonparametric estimation of the average marginal effect of price on demand allowing for price heterogeneity. We find that price heterogeneity exists, meaning that consumers' purchasing sensitivities depends on price, rather than being a fixed coefficient as in standard linear-parametric logit demand models. We employ a detailed, disaggregated set of supermarket scanner data on ready-to-eat cereals for 2005-2007.

Our identification strategy relies on a conditional form of exogeneity: Price is independent of unobservable demand drivers conditional on a set of selected conditioning instruments. Conditional exogeneity (conditional independence) has been at the center of recent theoretical research on identification in econometrics, making our application a timely pursuit. We rely on Megerdichian's (2009) research on identification of price effects in discrete choice models of demand in selecting our conditioning instruments.

Our nonparametric estimates of the effect of price on demand yields some useful results. First, we find that as price increases for a product (or as we go from cheaper cereal products to more expensive ones), the effect of price on quantity demanded becomes less elastic. This addresses a well-known criticism of standard linear-parametric logit demand models that render high-priced (low market share) products to have high price elasticities, which are not consistent with profit margin expectations. Second, both our nonparametric and parametric estimates yield price effects that are more elastic when we include conditioning instruments. This suggests that our instruments and identification strategy are properly addressing price endogeneity issues, such as the dependence of price on unobserved product quality.

A glaring problem with our nonparametric analysis is that the price effect estimates quickly become too inelastic for mid and high price ranges. Inelastic estimates are not consistent with price-cost margins derived from standard oligopoly models of profit maximization. We return to this problem in future work.

Finally, we note that further work here includes the following: (i) employing cross-validation methods for bandwidth selection; (ii) constructing confidence intervals for the nonparametric estimates; (iii) examining the effects of data aggregation on nonparametric estimates of price effects; (iv) implementing the nonparametric estimation for all four of the city-supermarkets in our sample; and (v) linking the research presented here with Berry and Haile's (2010) research on nonparametric identification of demand.

3.7 Acknowledgment

Chapter 3 is coauthored with Xun Lu.

3.8 Tables and Figures

	Market Share	Number of SKUs	Wtd. Avg. Price/Lb.	Abv./Blw Avg. Price	Abv./Blw StrBds. Price
Kellogg	42.0%	55	$2.82	-5.6%	29.7%
General Mills	29.9%	45	$3.54	18.4%	62.7%
Post	14.3%	25	$2.75	-7.8%	26.7%
Quaker	7.6%	13	$2.85	-4.5%	31.2%
Store Brands	6.2%	12	$2.17	-27.2%	
Total	100.0%	150	$2.99		

Source: IRI scanner data. Based on four city-supermarkets from 2005-2007. Kellogg Includes six Kashi products. Market share is based on quantity (pounds of cereal) sold.

Figure 3.1: Market Share and Price

							Percentiles				
	Min	Max	Mean	StDev	Skew	Kurt.	2.5	25	50	75	97.5
Market Share	0.000	0.156	0.009	0.012	4.750	31.775	0.002	0.004	0.006	0.010	0.042
Price	0.621	6.909	3.533	1.044	0.171	-0.438	1.674	2.758	3.466	4.281	5.626
Promotion	0.000	0.500	0.059	0.114	1.800	2.102	0.000	0.000	0.000	0.030	0.383
Density	0.753	4.092	1.508	0.528	1.983	6.822	0.931	1.111	1.388	1.870	2.540
Calories	86.7	400.0	155.4	50.0	2.1	7.4	100.0	120.0	146.7	181.7	280.0
Fiber	0.0	28.0	3.8	4.5	3.3	13.1	0.0	1.3	2.7	5.0	20.0
Sugar	0.0	20.0	10.9	5.9	-0.3	-1.0	0.0	6.7	12.0	14.7	20.0
Baseline Price	1.627	4.367	3.057	0.658	0.060	-0.440	1.672	2.603	3.096	3.400	4.328
Wage	786.0	1096.0	923.6	68.8	0.4	-0.3	807.0	874.5	912.0	973.5	1067.0

Based on 64 inside goods for four city-supermarkets, 2005-2007. Number of observations is 39,936 (64 goods x 4 city-supermarkets x 156 weeks).

Figure 3.2: Summary Statistics

	Mkt. Shr.	Price	Promo.	Density	Calories	Fiber	Sugar	Base Price
All Goods	0.009	3.53	0.059	1.51	155.4	3.8	10.9	3.1
	(0.0001)	(0.005)	(0.001)	(0.066)	(6.26)	(0.56)	(0.73)	(0.08)
General Mills	0.010	3.97	0.048	1.39	144.5	3.4	10.7	3.4
	(0.0001)	(0.008)	(0.001)	(0.070)	(6.49)	(1.12)	(1.23)	(0.09)
Kellogg	0.010	3.44	0.081	1.49	151.2	3.8	11.4	3.1
	(0.0001)	(0.008)	(0.001)	(0.092)	(8.57)	(0.85)	(1.20)	(0.14)
Post	0.008	3.34	0.049	1.87	193.7	5.2	10.4	2.8
	(0.0001)	(0.013)	(0.001)	(0.305)	(29.03)	(1.23)	(1.92)	(0.18)
Quaker	0.009	3.40	0.034	1.58	172.2	3.0	11.3	2.8
	(0.0002)	(0.019)	(0.002)	(0.169)	(15.74)	(0.88)	(1.96)	(0.23)
Store Brands	0.007	2.16	0.029	1.44	145.8	4.3	9.9	1.9
	(0.0001)	(0.011)	(0.002)	(0.200)	(15.92)	(1.29)	(3.16)	(0.16)

Based on data for 64 inside goods for four city-supermarkets, 2005-2007. Standard errors in parentheses.

Figure 3.3: Averages by Manufacturer

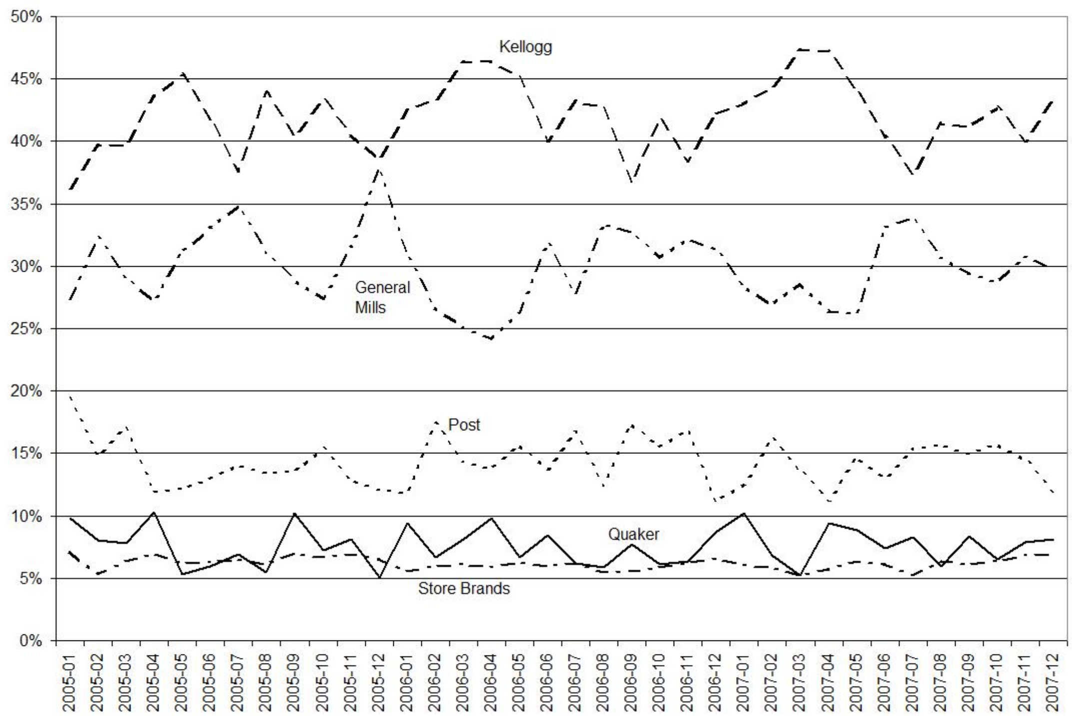


Figure 3.4: Market Share

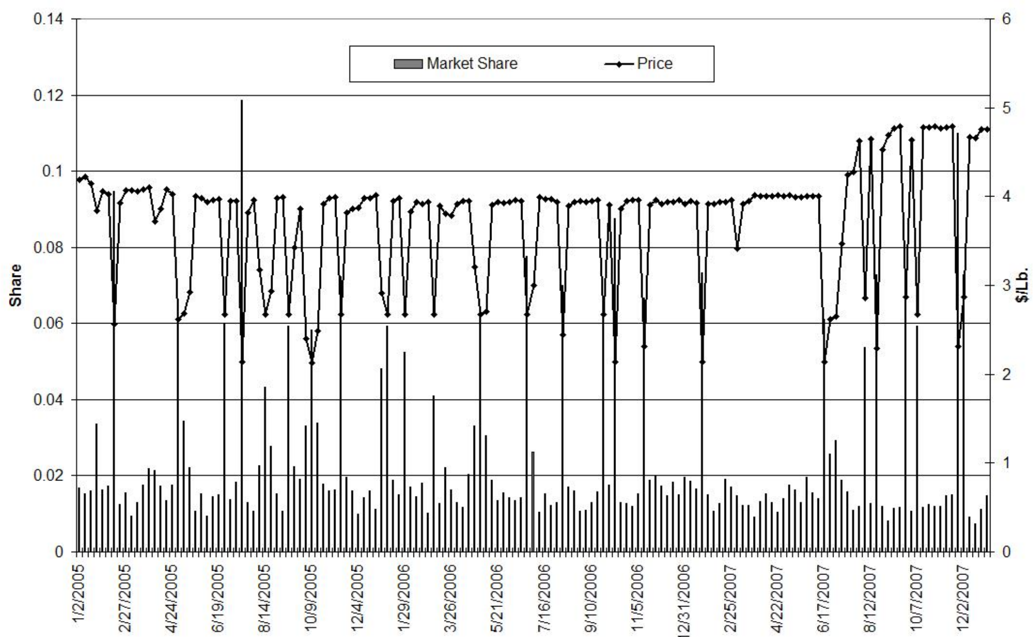


Figure 3.5: General Mills Cheerios Market Share and Price

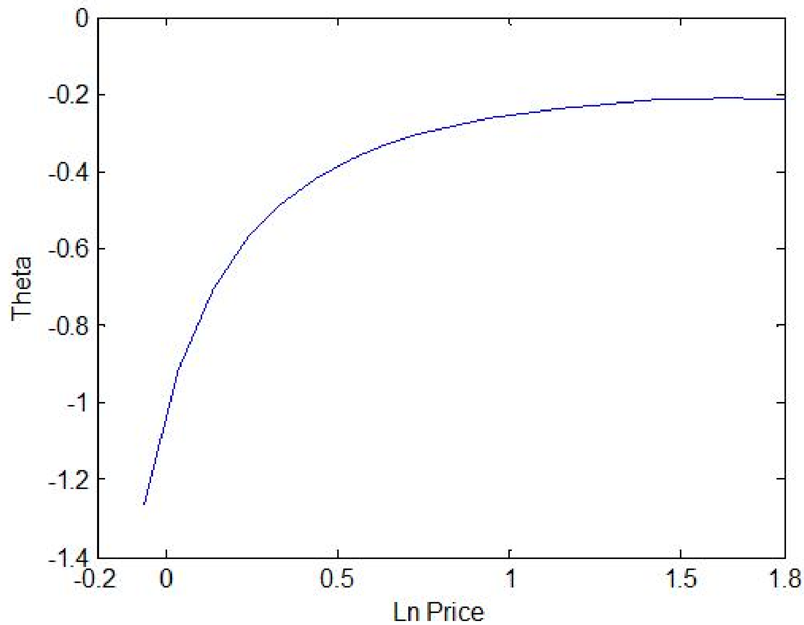


Figure 3.6: Nonparametric Estimate without Conditioning

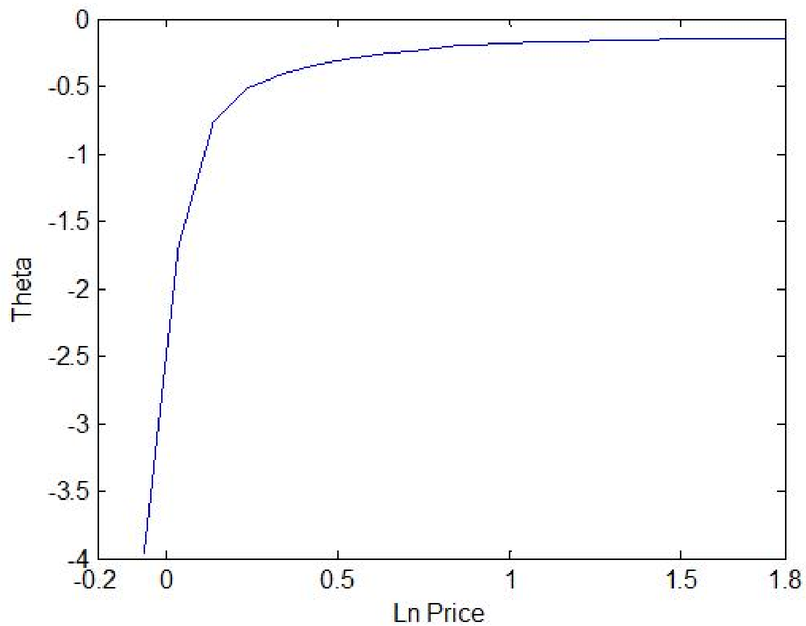


Figure 3.7: Nonparametric Estimate with Conditioning

Chapter 4

Product Downsizing and Hidden Price Changes in the Ready-to-Eat Cereal Market

4.1 Introduction

Firms have at least three ways to change the prices of their products. The first way, defined here as an overt price change, is to change the dollar amount charged per package without changing the amount of the good contained in the package. The second way to change price, defined here as a hidden price change, is to change the package size without changing the dollar amount charged for the package. Finally, a third possible way to change price is with a combination of an overt and hidden price change, defined here as a hybrid price change.

There are very few prior studies on hidden price changes. The only two relevant studies we are aware of are Adams, di Benedetto, and Chandran (1991) and Gourville and Koehler (2004), neither of which directly examines the impact on a firm's revenues or profits arising from a hidden price change. We therefore contribute to the existing literature by studying the extent to which a firm is successful in increasing its products' prices with a hidden price change. If a sufficient number of consumers do not respond to the hidden price increase, then we expect the firm to benefit from the price increase. In contrast, if a sufficient number of consumers do notice the hidden price increase and thus substitute to other goods, then we expect the firm to suffer from the price increase.

Furthermore, we examine hidden price changes in the ready-to-eat breakfast cereal industry. During 2006 and 2007, the prices of commodities that are primary inputs for cereal production, such as wheat, rice, and corn, increased dramatically. This increase in input costs prompted General Mills, the second largest firm in the industry, to increase prices for nearly all of its products with a hidden (or hybrid) price change in July 2007. For example, General Mills changed its flagship line of Cheerios products from 10oz., 15oz., and 20oz. down to 8.9oz., 14oz., and 18oz., respectively. We employ an extremely rich scanner data set that allows us to determine which products had a price change and how the price change was implemented. The detail of the data also allows us to undertake a variety of empirical analyses, including estimating a full demand system to examine the success of a hidden price change, which no other study has done to date.

We conduct four empirical investigations to determine the impact of hidden price changes on General Mills' cereal business. First, we estimate a basic linear demand model as in Gourville and Koehler (2004), and we also improve their specification by including prices of all cereals and promotional variables. Second, we estimate demand with the almost ideal demand system (AIDS) to predict expenditure share for the after-downsizing sample period, and compare it to the actual expenditure share. We also estimate a difference-in-difference specification for expenditure share to determine how General Mills' portfolio of products that were downsized fared as a whole. Finally, we further examine Chapter 2's analysis of product downsizing's effect on General Mills' profitability.

We find that out of the 20 General Mills products in our sample that downsized cereal content, about half the products increased expenditure share by more than what the AIDS predicts, indicating a sufficient number of consumers didn't notice the product downsizing for these goods. A key finding is that the products that did benefit from the downsizing tend to be large size boxes (by weight), while those that did not benefit from the downsizing tend to be small size boxes. One explanation is that consumers are more likely to notice hidden price changes when already small boxes become even smaller, leading them to substitute to larger boxes. An alternative explanation is the possibility that consumers don't notice hidden price changes, but because they prefer boxes to not be too big nor too small, they substitute away from small boxes that are downsized to large boxes that are downsized.

With a difference-in-difference estimation scheme, we find that the downsizing had a negligible effect on the expenditure share for the 20 General Mills products as a whole. While this result may appear to suggest that consumers noticed the hidden price change, it is perhaps more plausible the result just confirms the AIDS prediction exercise in that some of General Mills' products benefitted because enough consumer didn't notice, while other products did not because consumers did notice, thus leading to a negligible net effect.

This chapter is organized as follows. Section 2 provides an overview of the hidden price change phenomenon, including a review of prior relevant studies on the topic. Section 3 describes the supermarket scanner data on cereals that we employ here. In Section 4, we present a variety of empirical work outlined above. Section 5 concludes.

4.2 Hidden Price Changes

Firms have more than one way to change price. The first way, defined here as an overt price change, is to change the dollar amount charged per package without changing the amount of content in the package. For example, a cereal manufacturer may increase price from \$4.00 per 16oz. box to \$4.50 per 16oz. box, a 12.5% price increase. The second way to change price, defined here as a hidden price change, is to change the amount of content in a package without changing the dollar amount charged for the package. For example, a cereal manufacturer may decrease the size of a 16oz. box to 14.22oz. per box while keeping the price fixed at \$4.00 per box, also constituting a 12.5% price increase. Consistent with previous literature on this topic, we refer to the act of decreasing the amount of content in a package as "downsizing." Finally, a firm may opt for some combination of an overt and hidden price change, defined here as a hybrid price change. For example, a firm can decrease the size and price of a $4.00 per 16oz. box to $3.75 per 12.69oz. box, which is again a 12.5 % increase in price measured in dollars per ounce.

There are important aspects of a hidden price change worth noting. First, a hidden price change is tantamount to an overt price change, but in a way that some consumers may not notice (hence, hidden). Second, changing the amount of content in a package may require redesigning the package to accommodate the new size, or at the very least require the firm to change the weight information on the package. Thus, a hidden price change may involve production considerations, whereas an overt price change does not. Finally, it may be relatively difficult to undo a hidden price change by going back to the pre-downsized package; thus, undoing a hidden price change may require a counteracting overt price change. We examine these ideas in subsequent sections.

Gourville and Koehler (2004) refer to several product segments that have undergone downsizing, including coffee, yogurt, potato chips, baby diapers, bottled water, breakfast cereal, and ice cream. Adams, di Benedetto, and Chandran (1991) similarly point to these industries, and also note several other product segments, including candy bars and paper towels. Notably, the examples have the commonality of being packaged goods that sell primarily in supermarkets and, for the most part, appear to be differentiated products in an oligopoly-type market. That the primary outlet for these products are supermarkets is an important point, as temporary price promotions in supermarkets account for a significant part of a good's price variation (see Pesendorfer, 2002, and Hosken and Reiffen, 2004).[1] Therefore, downsizing may be more likely to go unnoticed when sticker prices vary considerably week-to-week.

[1]See also Figure 4.15 that conveys the volatility of retail prices for General Mills Cheerios.

4.2.1 Cost as a Driver of Downsizing

A reason manufacturers may change the price of a product is to remain profit maximizing when faced with a change in marginal cost. Consider a profit maximizing monopolist that sets price according to the well-known Lerner price-cost margin

$$\frac{p-c}{p} = -\frac{1}{\eta},$$

where η is the own price elasticity of demand, p is price, and c is marginal cost. Rearranging gives the familiar optimal pricing rule

$$p = \left(\frac{\eta}{1+\eta}\right) c.$$

According to this pricing equation, a firm may find it necessary to change price in response to a change in its marginal cost. During the last several years leading up to the start of the 2008 U.S. recession, prices of input commodities, including petroleum, grains, and rice increased markedly.

Figure 4.11 presents monthly price indices for several commodities that are inputs for the production of breakfast cereals. Prices of corn, rice, wheat, and petroleum[2] were on the rise between 2003 and 2007, and increased considerably starting in 2006. The price of sugar, on the other hand, was relatively stable over the five-year time period. Figure 4.1 presents yearly data for the commodities. Corn and wheat had large price movements in 2006 and 2007, while rice and petroleum had their largest price movements in 2004 and 2005. According to Chapter 2, 60% of the cereal products in his sample list corn or wheat as the primary ingredient. It is clear that cereal manufacturers faced rising input prices during the last decade, particularly during 2006 and 2007, which drove some of them to react with a hidden price increase; specifically, General Mills downsized its boxes for 20 out of its 23 products in our sample in July 2007.

Moreover, as one industry commentator reported in December 2007, "to counter rising commodity expenses, General Mills has reduced merchandise price discounts and promotions... General Mills even shrank its cereal-box size and

[2]Petroleum can be thought of as a proxy for general energy costs incurred to produce and transport cereal.

increased the average price of cereal per ounce to help counter rising costs and better compete against Kellogg Co., its larger rival in cereals... These actions, on top of price increases for select items and new product introductions, are expected to offset this cost pressure..."[3]

If the variable input cost and thus marginal cost of production for a firm increases, the firm will find it necessary to increase the price of its products to remain profit maximizing. The question of interest is should the manufacturer overtly increase the price of the product or should it change price through a hidden increase? The answer to this question depends in large part on how consumers react to overt versus hidden price changes.

Consider the example given earlier of changing the price of a 16oz. box of cereal from \$4.00 to \$4.50 (a 12.5% increase) versus attaining an equivalent price increase by keeping the price at \$4.00 but downsizing the box from 16oz. to 14.22oz. Assume that the intermediary retailer simply passes on price changes to the consumer. If consumers consider the two different pricing schemes to be completely equivalent – i.e., they are rational – then it does not matter which way firms increase price since consumers will perceive both identical. If enough consumers do not perceive them to be equivalent, potentially because consumers simply don't notice that the contents of the box changed from 16oz. to 14.22oz., then the manufacturer should engage in a hidden price increase.

Of course, a manufacturer can't keep decreasing the contents of its package, because at some point it becomes obvious to the consumer that they aren't getting much for their money. In fact, our empirical estimates show that, on average, larger-sized cereal products gained share relative to smaller-sized products after the downsizing.

4.2.2 Previous Research on Downsizing

To date, the most comprehensive academic study examining product downsizing and the hidden price change phenomenon is Gourville and Koehler (2004).

[3]"General Mills Backs Year View as Profit Rises," by Matt Andrejczak and Angela Moore, www.marketwatch.com, December 19, 2007.

Although they refer to several product segments that engage in hidden price changes, an indication that it is a fairly common occurrence, there are very few academic studies that rigorously examine product downsizing. The few studies we found on the topic are tangentially relevant, as none of them adequately measure the quantitative impact of hidden price changes.

Granger and Billson (1972) conduct an experiment with consumers to determine the extent to which they are aware of the price per unit of measurement (e.g., dollar per ounce) of soft drinks and laundry detergents. They find that consumers are generally not aware of the price per unit. When the price per unit of measurement is explicitly revealed to the consumers, they tend to substitute towards larger-sized items that typically have a lower price per unit. Note, however, that this study took place prior to retailers adopting unit pricing standards. As of 2008, 19 states in the U.S. have adopted unit pricing regulations that requires retailers to present price per unit of measurement information;[4] the majority of retailers, particularly supermarkets, voluntarily offer such information.[5] From this study we conclude that consumers tend to be more sensitive to overt price changes than hidden price changes.

Adams, di Benedetto, and Chandran (1991) offer several reasons for why firms find the need to downsize packaging, including maintaining a price point, increasing margin and profitability, increasing frequency of purchase, and offsetting raw material cost increases. They identify 25 product segments that committed downsizing primarily during the 1980s, including candy bars, coffee, cereal, ketchup, paper towels, and soap. In many instances, they document that the same packaging is used to hold the downsized contents.

Their study, although informative, provides very little empirical evidence assessing whether firms are successful in implementing hidden price increases. For example, they examine three brands of cereal that were downsized during the early 1980s, and make inferences about the impact on expenditures after the downsize by simply examining expenditures after the downsize. However, many other con-

[4]National Inst. of Standards and Technology, Uniform Laws and Regulations, NIST Handbook 130, 2009.

[5]Food Marketing Institute, Item and Unit Pricing, October 2007 (electronic article).

founding changes may take place during the relevant time period after the downsize that would lead to incorrect inferences. For example, if some of the many other cereals in the market that are substitutes for the downsized product decreased their price, or if promotions for the downsized product increases after the hidden price change, this may lead to an increase in expenditures for the downsized product for reasons potentially unrelated to the hidden price increase. We address these concerns in our empirical study in Section 4.

Finally, Gupta et al. (2006) examine the legal and ethical issues related to product downsizing and hidden price increases. They maintain that although hidden price changes through product downsizing may adversely affect consumers that do not notice the price increase in the same way as they would notice if it were an overt price increase, manufacturers are nevertheless in compliance with U.S. Food and Drug Administration regulations. However, they point out that there may be room for the Federal Trade Commission to pursue firms if the hidden price change is viewed as deceptive. To date, we are not aware of any legal or regulatory consequences resulting from hidden price changes. The reason is likely due to the fact that most retailers, regardless of which state they are located, provide consumers with a dollar per unit of measurement price.[6]

Gourville and Koehler (2004)

We pay special attention to Gourville and Koehler's (2004) Harvard Business School working paper (hereafter, "GK"), as it is the most complete study that directly examines product downsizing and hidden price changes. GK examine multiple product categories, using both surveys and market data, to determine that consumers react to overt price increases much more strongly than to hidden price increases. GK conduct four studies to assess hidden price increases; we describe and assess three of the relevant ones here.

First, GK examine the package size and shelf price of ready-to-eat cereals

[6]For example, most grocery stores present a shelf sticker that states a 16oz. box of cereal is \$3.99; the same sticker will state \$0.25/oz. below the \$3.99 figure, although typically in much smaller font.

from a single supermarket at one point in time. They suggest that if consumers are more sensitive to overt price changes relative to hidden price changes, and if manufacturers are aware of this, then we should expect to see more variation in package sizes than in prices. For 157 cereal products, they examine the ratio of the standard deviation of package sizes to the mean of package sizes, and compare this to a similar ratio for price. For example, they find that the ratio for package size is 0.244 for General Mills products compared to a ratio for price of 0.122 for the same General Mills products. They conclude that because there is more variation in package sizes relative to prices, manufacturers are exploiting the fact that consumers don't notice hidden price changes relative to overt price changes.

We don't necessarily dispute GK's conclusion that consumers are less sensitive to hidden price changes, but the way they implemented their empirical test is not appropriate. Assuming that examining variance is applicable here, then what makes sense is to examine the variance of shelf price changes for a good over time versus the variance of its package size changes over time. In other words, their idea to examine variance may be valid, but it should be done as a time-series exercise for each good separately, rather than a cross-sectional exercise for all goods at one point in time.

For example, let s_t denote a good's package size in time t, $d_t \equiv 1[\text{downsize}]$, $\delta \equiv s_b - s_a$, where s_b and s_a are the sizes of the good before and after the downsizing, respectively. Consider a relevant time period $t = 1, 2, ..., T$, and that the product downsizes once during that period. The mean and the standard deviation of s_t is given by $\overline{s} = T^{-1}\sum_t s_t = (1-\pi)s_b + \pi s_a = s_b - \pi\delta$ and $\sigma = \delta\sqrt{\pi(1-\pi)}$, respectively, where $\pi \equiv T^{-1}\sum_t d_t$. If the product downsizes once, $\sqrt{\pi(1-\pi)} \in (0, 0.50]$. Following GK's idea to use $\sigma/\overline{s}$ to evaluate the impact of product downsizing, we have that $\sigma/\overline{s} = \delta/(\delta + 2s_a)$ for the extreme case where $\sqrt{\pi(1-\pi)} = 0.50$. It does not take much for $\sigma/\overline{s}$ to be very small. Consider a good that is downsized from 20 ounces to 18 ounces exactly at the halfway point of a relevant time period that lasts $T = 260$ weeks. In this case, $\delta = 2$, $\overline{s} = 20$, $\sigma = 1$, and GK's ratio is $\sigma/\overline{s} = 0.05$, a very small number when examining downsizing correctly as a time series exercise, which would in this case indicate that downsizing is not undertaken very often by the firm.

The fact that GK find variation in package sizes for a cross section of products at a particular point in time is not surprising for a differentiated products market. It is well known that at any point in time cereals come in a wide variety of shapes, sizes, and densities, among other attributes, to satisfy consumers' preferences, which includes preferences for different package sizes. GK's analysis is simply picking up the fact that, for example, Cheerios comes in three different packages sizes.

In their second study, GK survey 60 adults about how they would price coffee if they were in charge of pricing at a gourmet coffee shop. Their survey asks subjects whether it is better to price coffee at $6 per half pound or $12 per pound; the subjects are also asked to indicate which of the pricing schemes would promote sales better for the store.[7] Nearly 87 percent of respondents chose the $6 per half pound pricing option, and the mean score is 2.85 for the second question, indicating that respondents believe the $6 per half pound pricing scheme to be more lucrative for the business even though the two pricing schemes are the same price in terms of dollars per pound. Respondents justify their answers by noting that consumers won't notice the "per half pound" part of the price, or that the unit of measurement (the denominator of price) is not as important to consumers than the dollars (the numerator of price). GK find that this sentiment favors a hidden price increase strategy for firms.

In their fourth study, GK return to the cereal industry, but this time use data covering 145 weeks during the late 1990s for an undisclosed number of supermarkets that they apparently have aggregated into a single representative store. Their data is for four cereal products from a single manufacturer that engaged in product downsizing during the relevant time period, and contains information on unit sales, price, and package size. They specify a standard linear regression equation with quantity purchases of cereal (in boxes) as a function of a trend variable, own price (presumably, in dollars per box), and size (ounces per box).

[7] The latter question is based on a one to seven scale. A score of one is described to be that $6 per half pound is "much more effective," while a score of seven is described to be that $12 per pound is "much more effective," and four is described to be that both are "equally effective."

GK explain that if consumers do not notice hidden price changes, then the price variable will have more impact on quantity than the size variable in their regressions. For their sample of four cereals, they obtain statistically significant negative coefficients on price that are relatively large in magnitude compared to their estimates of the size coefficient. They conclude that this is evidence that consumers pay more attention to overt price changes than hidden price changes from downsizing. We note here that although it's not fair to compare the coefficient estimates on price and size directly, because price and size have different units of measurement, comparing the respective T-statistics of price and size, which are unit free, still shows price dominating size in their regressions.

Except for one of the four cereals, the coefficient on size is not statistically significant, and GK note that the negative signs on the size variable are "counter-intuitive." However, there is an intuitive explanation for the negative coefficients. Letting q denote quantity purchases and s size, their estimate for three of the four products is $\partial q/\partial s < 0$, meaning that a decrease in the amount of cereal in a package leads to an increase in the number of boxes of cereal purchased. The reason this is actually intuitive is that we may expect that, *ceteris paribus*, consumers who don't notice the downsizing and who consume a fixed amount of cereal during a relevant time period to purchase more boxes of cereal to maintain their consumption if there is less content in each box.

There are a few empirical issues with their analysis. First, as GK themselves correctly point out, they have left out marketing-mix variables, such as local advertising and in-store displays, that typically play an important role in driving purchases in supermarkets. Second, they are essentially estimating a poorly-specified linear demand equation for each of the four cereals in their sample by leaving out the prices of the many other cereals in the market.[8] Third, it appears the price variable they use is the regular shelf price for the cereal, not the effective transaction price that takes into account temporary price promotions. The effective transaction price for cereals during a promotional week can sometimes be half

[8] They do not intentionally leave out prices of other cereals (potentially substitutes) and marketing-mix variables; it appears they simply don't have the data.

of the shelf price,[9] which may impact their regression results. Lastly, it appears they have taken data from multiple supermarkets and aggregated it into a single representative supermarket that they use for their regressions. This may effect the results, particularly the size coefficient, due to aggregation bias.

We do not disagree with GK's overall conclusion that consumers are less sensitive to hidden price changes versus overt price changes. Nevertheless, none of their studies adequately addresses whether the hidden price change was successful from the firm's point of view. They only address consumers' perceptions about hidden price changes, and then draw conclusions about how those perceptions may affect firms that engage in product downsizing. In contrast, we examine the impact of hidden price changes on a firm's revenue and profitability, and we fix most of the problems with their demand analysis.

4.3 Data

Our application here focuses on the ready-to-eat cereal industry. We refer the reader to numerous studies by industrial organization and marketing researchers for a background on the cereal market.[10] Currently, the four major nationally-branded cereal manufacturers operating in the U.S. are General Mills, Kellogg, Post, and Quaker. A fifth important source of cereal sales are store brand products, which are typically inexpensive versions of nationally-branded cereal products. Figure 4.2 presents aggregate market share and pricing data. General Mills and Kellogg are the two largest players in this market, followed by Post, Quaker, and store brands, respectively.

We employ a rich set of supermarket scanner data from Information Resources, Inc. (IRI). The data set consists of variables measuring price, quantity, and promotional variables for 150 of the top-selling cereal products sold in the U.S. for three years between January 2005 and December 2007. The data set has

[9] See Figure 4.15.

[10] See, for example, Schmalensee (1978), Ippolito and Mathios (1990), Hausman (1997), Rubinfeld (2000), Nevo (2001), Nevo and Wolfram (2002), Shum (2004), and Chapters 1 and 2.

a panel structure, where the time dimension has a weekly frequency and the individual dimension is a city-supermarket chain.[11] Except for (sometimes serious) cases of missing observations and other irregularities, there are 157 weeks of data for each of the 121 city-supermarket chains covering 41 cities[12] across the United States. We refer the reader to Chapters 1 and 2 for further details.

4.3.1 Market Share and Pricing

Figure 4.12 displays the monthly market share based on pounds of cereal sold of the four cereal manufacturers as well as store brand products. The store brand market share does not deviate much from five percent on a month-to-month basis, while the branded manufacturers have considerably more volatility over time.[13] Figure 4.12 indicates that not only do the market shares of Kellogg and General Mills vary over time, but that they are nearly mirror images. The correlation of the monthly market share for Kellogg and General Mills is -0.78, indicating that Kellogg's gain in market share is typically General Mills' loss, and vice-versa.

Figure 4.13 presents the weighted average monthly retail price per pound for the manufacturers between 2005 and 2007. Figure 4.14 presents the weighted average price per box. General Mills products tend have higher prices than the other manufacturers, while the store brands products are sold at a deep discount compared to the nationally-branded products. What is notable here is that about July 2007, the point during which General Mills downsized 20 of the 23 products in the sample, the price per pound (Figure 4.13) of General Mills products increases, while their price per box (Figure 4.14) during the same time period is relatively flat. This is indicative of their hidden price increase campaign. Moreover, the timing of the increase in the price per pound for General Mills is roughly in line with the sky-rocketing commodity prices presented in Figure 4.11.

[11]The data are not for individual stores; rather, they report information for a supermarket chain, which may be comprised of many stores in a given city.

[12]A city is approximately equivalent to the Census Bureau's metropolitan statistical area (MSA) or combined metropolitan statistical area (CMSA).

[13]For more details, see Chapter 2.

The variation of the manufacturers' prices over time is due to temporary price promotions that are more pronounced in the disaggregated weekly data. Figure 4.15 presents the weekly price per pound and price per box of General Mills Cheerios 15oz. (14oz. after the downsizing), the best selling product in the sample over the three year period. The series is for a single supermarket chain located in a major U.S. city.[14] The numerous price decreases are the result of temporary promotions, which are an important source of price variation in this data. The richness of this data set permits us to determine when downsizing took place for General Mills products in 2007. We explore this next.

Identifying Downsizing

Importantly, the disaggregated nature of the data allows us to identify when General Mills (GM) downsized their products.[15] General Mills downsized a substantial portion of its cereal products in 2007, and downsized the remainder in 2008. Kellogg started downsizing its products in 2008. To note, 2008 and 2009 are beyond our sample.

Dividing a good's price in dollars per box by its corresponding price in dollars per pound yields the ratio pounds per box, or ounces per box after multiplying by 16. This ratio is constant over time for products that did not downsize, but decreases if the product underwent a downsize. Figure 4.15 presents the ounces per box ratio for GM Cheerios 15oz. using the price series in the same figure. The product downsizing for this good took place at this particular supermarket chain in early July 2007, as GM downsized its product from 15 ounces to 14 ounces. This can be seen in two ways. First, the gap between the two price series increases during July 2007. Second, the ounces per box, calculated as the ratio of the two price series, drops from 15oz. to 14oz. during July 2007.

Figure 4.3 presents the 20 GM products that underwent downsizing during

[14]Confidentiality agreements entered into for this UCSD PhD dissertation do not permit revealing the identities of the supermarkets.

[15]General Mills downsized a substantial portion of its cereal products in 2007, and downsized the remainder in 2008. Kellogg started downsizing its products in 2008. To note, 2008 and 2009 are beyond our sample.

July 2007.[16] Because different supermarkets changed over to the new, smaller boxes at different times, the downsizing date reported in Figure 4.3 is an average across all city-supermarket chains in the sample. The variable ounces per box reports the amount of cereal in the box before and after the downsizing. For example, GM changed the 10oz. box of Cheerios to a 8.9oz. box near mid-July 2007. This is an 11% decrease in the amount of cereal in the box, which translates into an implied hidden price increase of 12.4%.

Figure 4.3 also provides average prices for two measures: price per pound and price per box. Consider again the Cheerios product that changed from 10 ounces to 8.9 ounces. The average price per box across all the city-supermarket-weeks is \$2.95 before the downsizing, and is \$2.96 after the downsizing. Yet the average price per pound jumps 11.9%, from \$4.73 per pound to \$5.29 per pound. This is a clear-cut example of a pure hidden price change, which can also be seen by comparing the implied price change with the change in the price per pound, which is about the same for the 10oz. Cheerios product.

Figure 4.3 shows that some products, on average, underwent a hybrid price change, whereby the product had both downsizing and an overt price change. Nine of the 20 products lowered the price per box by more than five percent after the downsizing. For example, Chex Rice 15.6oz. changed to 12.8oz., implying a 21.9% hidden price increase, but the average price per pound for that product only increased by 7.8% after the downsizing. This discrepancy is due to the fact that the average price per box for this product declined by 10.2% in the after-downsizing period.

Finally, Figure 4.3 provides average expenditure share information for each of the products. The expenditure share is calculated as the proportion of a good's dollar sales to the dollar sales of all 150 goods in the sample at a city-supermarket-week. Some of the products increased share in the after downsize period (e.g., Cheerios 20oz.), while others saw a decline in expenditure share (e.g., Cheerios 15oz.). For some products, the large swings may be attributable to the decrease

[16]There may have been more, however we only focus on a subsample of 65 out of the 150 cereal products in our sample. Out of 65 cereal products, 20 out of 23 General Mills products were downsized.

in price per box. For example, Chex Rice 15.6oz. increased expenditure share by 15.6%, which may be due to the 10.2% drop in price per box that outweighed the hefty downsizing from 15.6 ounces to 12.8 ounces. On the other hand, Cheerios 20oz. witnessed an increase in expenditure share of 22.7% after its downsizing, but its price per box decreased by only 2.8%.

Thus, there appears to be other factors that are influencing the expenditure share of the products in the before to after downsizing periods. This may include not just the good's own price, but the good's promotional intensity, the prices of the numerous other cereals in the market, seasonality, and trend. By estimating a full demand system, we take these factors into consideration.

4.4 Empirical Results

In this section, we conduct four empirical investigations to determine the impact of hidden price changes on GM's cereal business. First, we estimate a basic linear demand model, such as the one in GK, and we also extend GK's specification by including prices of all cereals and promotional variables. Second, we estimate demand with the almost ideal demand system (AIDS) to predict expenditure share for the after-downsizing sample period in order to compare it to the actual expenditure share. We also estimate a difference-in-difference specification for expenditure share to determine how GM's portfolio of products that were downsized fared as a whole. Finally, we discuss Chapter 2's analysis of product downsizing's effect on GM's profitability.

4.4.1 Linear Demand

Figure 4.4 presents basic linear demand estimates tantamount to GK's "individual regressions" presented in their Figure 4.4 and discussed above in Section 2. Here, we first focus on GK's specification as it is a reasonable starting point given that it's the only empirical study of hidden price changes to date. For each of the $j = 1, 2, ..., J = 20$ GM products that underwent downsizing presented in

Figure 4.3, quantity demanded is described by the data generation process (DGP)

$$q_{jm} = \lambda_j + \alpha_j s_{jm} + \gamma_j p_{jm} + \omega_j T + \lambda_{jt} + \lambda_{jr} + \varepsilon_{jm},$$

where q is quantity purchases of boxes of good j in market m, s is the box size in ounces, p is price measured in dollars per box, T is a trend variable, λ_{jt} represents time fixed effects,[17] λ_{jr} represents city-retailer fixed effects, and ε_{jm} captures unobservable variables that drive q_{jm}. A market m is defined as a city-retailer-week combination; i.e., the index m subsumes r (city-retailer) and t (week). The OLS estimates presented in Figure 4.4 are based on 157 weeks for each of the 121 city-supermarkets.[18]

As in GK's analysis, Figure 4.4's results show that good j's own price dominates the box size variable in explaining quantity demanded. Of course, price and box size are two different units, but even the T-statistic, which is unit free, shows that price has a much bigger impact on quantity than box size does. While price is always statistically significant, the box size variable rarely is. Moreover, estimates of α_j are negative except for Lucky Charms 20oz. and Reese's 14.25oz. Recall that GK noted that their finding of negative estimates of α_j were "counter-intuitive"; yet we maintain that $\widehat{\alpha}_j \equiv \partial q_j / \partial s_j < 0$ is in line with consumers not noticing hidden price changes, since when s declines due to downsizing we expect q to increase in a relevant time period because consumers will purchase more boxes since there is less cereal in each box.

Figure 4.5 presents estimates of the own price effect, γ_{jk} for $j = k$, and the box size effect, α_j, for a more detailed specification of a linear demand DGP given by

$$q_{jm} = \lambda_j + \alpha_j s_{jm} + \sum_k \gamma_{jk} p_{km} + \mathbf{x}_{jm}\theta_j + \mathbf{z}_m \pi_j + \omega_j T + \lambda_{jt} + \lambda_{jr} + \varepsilon_{jm}.$$

Prices are given by p_k for $k = 1, 2, ..., K = 65$ cereal products,[19] $\mathbf{x}$ is a vector of good

[17]Based on months, not weeks, although the frequency of the data is weekly.

[18]The number of observations would be $121 \times 157 = 18,997$ but for missing price data.

[19]Including 23 General Mills products (20 of which are the downsized products in Figure 4.3), 26 Kellogg products, 9 Post products, 3 Quaker products, and 4 store brand products. See Chapter 1 for a complete list of the 65 products.

j's marketing-mix or promotional variables,[20] and $\mathbf{z}$ is a vector income and demographic variables.[21] In comparison to GK's specification in Figure 4.4, estimates of the own price effect have decreased in magnitude with the full specification, which is due to the omitted variable bias plaguing the simple GK specification. Moreover, many of the negative estimates of α in Figure 4.4 are now positive in Figure 4.5, although still statistically insignificant. A positive estimate of the size effect, $\widehat{\alpha}_j > 0$, suggests that consumers notice the downsizing and resulting hidden price change and are buying fewer boxes when the size declines.

Consider the estimate of α_j for GM Cheerios 15oz., $\widehat{\alpha}_j = -830$. GM Cheerios downsized from 15 ounces to 14 ounces (see Figure 4.3), meaning that sales of Cheerios 15/14 oz. are expected to increase by 830 boxes as a result of the downsizing. On the other hand, Cheerios Honey Nut downsized from 27 ounces to 25.25 ounces, resulting in a loss in sales of 250 boxes after the downsizing of that product.

Two primary conclusions emerge. First, it appears that the box size variable doesn't explain as much of the variation in quantity demanded as price. This is in line with GK's conclusion that overt price changes matter more to consumers than hidden price changes. Second, some products increased sales (in boxes) after the downsizing, while others did not, and most of the GM products had a size effect, $\widehat{\alpha}_j$, that were not statistically significant.

This approach, although an informative starting point, suffers a few drawbacks. First, even our "better-specified" version of GK's simple demand analysis is problematic because it is an *ad hoc* linear demand equation that may not be the true DGP for quantity sales. Moreover, properly making causal inferences about the own price effects, γ_{jk} for $j = k$, and the box size effect, α_j, relies on the assumption that we have identified those causal effects, which may be unrea-

[20]Includes ad only, display only, ad and display, and distribution. See Little (1998) for a detailed description of these variables. We've also included the number of weeks that have elapsed since the last promotion took place within a city-supermarket. See Megerdichian Chapters 1 and 2, and Pesendorfer (2002).

[21]Includes percentage of population under age 19, percentage of population over age 55, percentage of households with children, percentage of population that is white, income (weekly average wage).

sonable due to misspecification or endogeneity. We refer the reader to Hausman (1997), Nevo (2001), and Chapters 1 and 2 for an extensive review of addressing endogeneity in demand estimation. In the next section, we attempt to circumvent these issues with a predictive demand model of expenditure shares to determine whether downsizing was a successful strategy for General Mills.

4.4.2 AIDS Predictions

Deaton and Muellbauer's (1980) almost ideal demand system (AIDS) is rooted in consumer theory, and is considered to have a high level of econometric flexibility, in that even if the true demand function that describes the data is not the AIDS, the AIDS model will provide an approximate estimate that is near the true DGP. Moreover, the AIDS specification has expenditure share as the dependent variable, and is thus a convenient framework to determine how successful, from a revenue share perspective, GM's product downsizing strategy was in 2007. Because expenditure share is the proportion a good's sales relative to sales of all the cereals, it is ideal because it addresses changes in sales of a product due to changes in demand for all cereal. Finally, because we utilize this demand model for predictive purposes only, it is not necessary to obtain causal estimates of the coefficients.

For each of the $j = 1, 2, ..., J = 20$ General Mills products that underwent package downsizing, expenditure share is described by the AIDS specification

$$w_{jm} = \lambda_j + \sum_k \gamma_{jk} \ln p_{km} + \beta_j \ln\left(\frac{X_m}{P_m}\right) + \mathbf{x}_{jm}\theta_j + \mathbf{z}_m\pi_j + \omega_j T + \lambda_{jt} + \lambda_{jr} + \varepsilon_{jm},$$

where w_{jm} is the expenditure share based on dollar sales of good j in market m, X_m is the total expenditure of all $k = 1, 2, ..., K = 65$ cereal products in market m, and P_m is the Stone price index given by $\ln P_m = \sum_k w_{km} \ln p_{km}$. The remaining variables and indexing are defined the same as for the linear model, except that p here is price measured as dollars per pound. For further details on the AIDS, see Deaton and Muellbauer (1980), Chalfant (1987), Green and Alston (1990), Hausman (1997), Hausman and Leonard (2002), and Chapter 1.

We use the AIDS specification to predict the expenditure share for the 20 GM goods in question after the downsizing took place. In the first stage the analysis, the AIDS equation for the expenditure share of good j is estimated using only the before-downsizing sample period. In the second stage, the predicted expenditure share, $\widehat{w}_j$, is obtained by evaluating the estimated AIDS equations using the after-downsizing sample period. If a sufficient number of consumers didn't notice the product downsizing and resulting hidden (or hybrid) price change for good j, then we expect the actual expenditure share (w_j) to be higher than the predicted AIDS expenditure share ($\widehat{w}_j$) during the after-downsizing sample period.

Figure 4.6 presents the mean actual expenditure share and the mean predicted expenditure share for each of the 20 GM products for the after-downsizing sample period. Figure 4.6 also presents the ratio of actual to predicted expenditure share, the difference, and the corresponding P-value for the difference from a two-sided T-test of no difference between actual and predicted expenditure share. A ratio greater than one means that the actual expenditure share in the after-downsizing sample period is greater than the predicted expenditure share, indicating that the downsizing was successful for that GM product in terms of gaining revenue relative to the other cereal products in the demand system. Conversely, a ratio less than one means that the actual expenditure share of the good is less than the predicted expenditure share, indicating that the downsizing was not successful for that GM product. A few regression diagnostics from the AIDS estimation are also presented.

Figure 4.7 presents the same information as Figure 4.6, but sorts in descending order based on the expenditure share ratio. The first eleven products, GM Chex Rice 15.6oz. through GM Cookie Crisp 12.25oz., all have ratios greater than one that are statistically significant. It appears that for these eleven products, GM's campaign to downsize during 2007 was successful from a revenue gain perspective. The next four of the GM products, Reese's 14.25oz. through Cheerios 15oz., have actual-predicted mean shares that are either not statistically different from zero or have ratios close enough to one to call into question whether they gained or lost share relative to what is predicted by the AIDS model. Finally, the last five goods, Golden Grahams 13oz. through Cheerios Honey Nut 14oz., have actual-to-predicted ratios that are far enough below one to conclude that these products did not benefit from the downsizing and resulting hidden price change. That is, enough consumers noticed the hidden price change on these five goods, causing them to substitute to other products, so that GM attained a lower expenditure share relative to what the AIDS equations predicts for the after-downsizing period.

An interesting pattern emerges from the results in Figure 4.7 regarding the three groupings of GM products based on the ratio of actual-to-predicted expenditure share. The first eleven products with ratios greater than one have an average before-downsizing package size of 18.28 ounces (median 20oz.); the middle four products have an average size of 15.13 ounces (median 14.63oz.); and the last five products with ratios sufficiently less than one have an average size of 12.6 ounces (median 13oz.). This suggests that the larger-sized products fared better than smaller-sized products from the downsizing, which makes sense as consumers are more likely to notice downsizing of already small boxes, causing them to substitute to other cereals.

To make this point concrete, Figure 4.8 presents 10 of the 20 GM products in our sample that have multiple sizes within the same brand-line, including Cheerios (10oz, 15oz., 20oz.), Cheerios Honey Nut (14oz., 20oz., 27oz.), Cinnamon Toast Crunch (14oz., 20.25oz.), and Lucky Charms (14oz., 20oz.). Within each of the four brand groupings, the products are sorted in descending order based on the actual-to-predicted expenditure share ratio. What is of interest here is that the ratio within each product grouping is nearly monotonic in box size, with only a slight exception for 20oz. and 27oz. Cheerios Honey Nut. For example, consider the first grouping of the three Cheerios products. As we go from 20oz. to 15oz. to 10oz. down the list, the expenditure share ratio declines from 1.22 to 0.94 to 0.59, respectively.

This pattern generally holds for each of the four brand groupings, as the larger boxes within the brand enjoy higher ratios, indicating they tended to succeed in gaining expenditure share relative to the smaller boxes of the same brand based on the AIDS predictions. This evidence suggests that the downsizing caused consumers to shift from the smaller boxes of a brand to the larger boxes of the same brand. Moreover, the results may suggest that consumers initially believed that the large boxes were too large and the small boxes were too small, so that the downsizing benefitted the large boxes and harmed the small boxes.

4.4.3 A Diff-in-Diff Approach

The AIDS analysis presented above resulted in 11 products gaining share relative to what the AIDS model predicted, while 9 of the products were either even or lost share. Here, we take a different approach to determine how GM's 20 downsized products performed after the downsizing.

We consider a difference-in-difference (diff-in-diff) strategy, which can be particularly useful for natural experiments. The downsizing that took place by GM in July 2007 serves as a natural experiment because the impetus for the downsizings and resulting hidden price increases was the dramatic increase in input prices faced by GM (see Figure 4.1 and Figure 4.11). That GM's marginal costs exogenously increased, causing them to raise price by downsizing, is good reason to believe that the downsizing event serves as a treatment in a natural experiment. As a concomitant, we are hopeful that this approach partly allays Angrist and Pischke's (2010) concerns regarding causal inference in empirical IO studies.

We consider the following diff-in-diff specification:[22]

$$\begin{aligned} w_{krt} &= \lambda + \beta_1 G_k + \beta_2 D_t + \beta_3 GD_{kt} + \gamma p_{krt} + \mathbf{x}_{krt}\theta \\ &\quad + \mathbf{z}_{rt}\pi + \omega T + \lambda_k + \lambda_r + \lambda_t + \varepsilon_{krt}, \end{aligned}$$

where k indexes the 65 goods in the sample, r indexes city-retailers, and t indexes week. G_k is an indicator variable that takes on a value of one if good k is one of the 20 GM products that downsized, D_t is an indicator variable that takes on a value of one if week t is during the after-downsize period, and GD_{kt} is the product

[22] For an excellent example of this type of identification strategy in IO, see Shepard (1991).

of G_k and D_t. Finally, the variables p, $\mathbf{x}$, $\mathbf{z}$, and T are defined as in the earlier AIDS model, λ_k denotes product fixed effects, λ_r denotes city-retailer fixed effects, and λ_t denotes month fixed effects. Here, w_{krt} is the expenditure share of good k calculated as the proportion of dollar sales of good k in r during t to the dollar sales of all 150 products in the full sample (not just the $K = 65$ products in the AIDS analysis) in r during t.[23]

The diff-in-diff quantity of interest is $\Delta \equiv (w_g^a - w_g^b) - (w_{\widetilde{g}}^a - w_{\widetilde{g}}^b)$, where g denotes the 20 GM products that underwent downsizing (the "treatment" or "intervention"), $\widetilde{g}$ denotes the 45 other products in the sample that are not one of the 20 downsized GM products (the "control"), a denotes the after-downsize sample period, and b denotes the before-downsize sample period. The coefficient of interest is β_3, as this captures the diff-in-diff value of Δ :

$$\begin{aligned}
\Delta &\equiv \left[w_g^a - w_g^b\right] - \left[w_{\widetilde{g}}^a - w_{\widetilde{g}}^b\right] \\
&= \left[E(w \mid G = 1, D = 1, GD = 1) - E(w \mid G = 1, D = 0, GD = 0)\right] \\
&\quad - \left[E(w \mid G = 0, D = 1, GD = 0) - E(w \mid G = 0, D = 0, GD = 0)\right] \\
&= \left[(\beta_1 + \beta_2 + \beta_3 + ...) - (\beta_1 + ...)\right] - \left[(\beta_2 + ...) - (0 + ...)\right] \\
&= \beta_3.
\end{aligned}$$

Thus, the estimate of β_3 captures the aggregate net effect of the downsizing for the 20 GM products under investigation during the after-downsizing period, conditional on the additional covariates described above.

Column (6) of Figure 4.9 reports the OLS results. For comparison purposes, columns (1) through (5) exclude certain variables. The estimate of β_3 is sensitive to specification, as it varies from -0.546 to 0.025 depending on the included covariates and fixed effects. The R-squared and AIC improve substantially with the full specification in column (6). The estimate of β_3 changes from negative to positive when product fixed effects are included in column (6). The estimate of β_3 from column (6) is 0.025, which means that GM gained about 0.025 percentage points (e.g., from 23% to 23.025%) on its 20 downsized products during the

[23]Here, expenditure share is multiplied by 100 for scale purposes in reporting the regression results. E.g., if good k has 1.8 percent share, this is $w_{krt} = 1.8$ in the data, not $w_{krt} = 0.018$.

after-downsizing period. In other words, the effect of the downsizing that remains after controlling for confounding is an anemic increase at best, suggesting that, overall, not enough consumers fell for the hidden price change to increase GM's total expenditure share by a meaningful amount.

Therefore, this analysis, in conjunction with the individual product-by-product AIDS prediction analysis conducted earlier, suggests that some products performed well after the downsizing, while others did not (due perhaps to cannibalization), which ultimately leads to a negligible net aggregate effect for expenditure share. In the next section, we examine the impact of the downsizing on GM's profitability.

4.4.4 Logit Demand and Profitability

Finally, here we extensively examine Chapter 2's analysis of GM's change in profitability due to the downsizing of their 20 products in our sample. He first estimates a logit model of demand to obtain a matrix of demand elasticity estimates before and after the hidden price change (approximately July 2007).[24] Consistent with Nevo (2001), he assumes a multiproduct differentiated Nash-Bertrand oligopoly model for the cereal industry, and then proceeds to estimate implied price-cost margins (PCM) for each good j, $(p_j - c_j)/p_j$, for the before-downsizing and after-downsizing sample periods based on the demand estimates. Figure 4.10 present's the PCM results for the before and after downsizing periods found in Figure 2.13 of Chapter 2. General Mills increased its profitability 3.6%; Kellogg declined by 2.8%; Post increased by 4.2%; Quaker increased by 1.9%; store brands was flat.

We note that the estimates of GM's increase in profitability after the downsizing is roughly consistent with external company reports. GM reported that its net sales for "Big G" cereals in their U.S. retail segment increased by 2.7% for the quarter ending November 25, 2007, and that their operating profits for their

[24]Extensive attention is paid to identification; in particular, he evaluates standard instrumental variables approaches as well as applying conditional independence assumptions and related instruments.

entire U.S. retail segment[25] grew 1.4% for the 6-month period ending November 25, 2007.[26] The change in the profitability of GM found in Figure 4.10 (+3.6%) is roughly in line with the accounting information. The discrepancy is not unexpected for three reasons: (i) The accounting information reported is for all retail channels, not just supermarkets; (2) the time periods are not consistent; and (iii) the profitability estimates for GM in Figure 4.10 is for only 23 of GM's cereal products, not their entire portfolio of products as in the accounting reports. Nevertheless, the 3.6% estimated increase for GM supports the idea that after raising price by downsizing, GM increased their profits to get back to profit maximization, which had been eroded due to the rising input costs in 2006 and early 2007.

Another result from Figure 4.10 that is in line with what has transpired in the cereal industry is that Kellogg's profitability decreased after GM downsized, while all the other firms either increased profitability or did not change. Beginning in 2008, which is beyond our sample, Kellogg also increased the prices of its cereals through downsizing, which none of the other cereal manufacturers have done to date. This action by Kellogg, and the corresponding non-action by the other firms, comports with the margin calculations in Figure 4.10. Kellogg's profitability declined during that period because they too faced rising input costs in 2007, yet they didn't commit to a hidden price increase until 2008. General Mills moved first in 2007 with the downsizing, and now Kellogg – the only firm with an estimated drop in profitability – is the only firm to have followed suit and downsized in 2008. Moreover, the margins for the other three manufacturers have either increased or not changed, potentially explaining why they have not followed General Mills and Kellogg with hidden price increases.

4.5 Conclusions

We investigated the hidden price change phenomenon that results from product downsizing in the ready-to-eat cereal industry. General Mills, the second

[25]Includes all their U.S. product segments, such as yogurt, snacks, baking products, cereals, and more.

[26]Source: General Mills SEC Form 10-Q for quarter ended November 25, 2007.

largest manufacturer of cereal in the U.S., increased the price of most of its cereals by decreasing the content of its boxes in July 2007. This change in price was driven by an increase in several key commodity prices that are inputs for producing cereal.

We examined the few studies on hidden price changes and product downsizing. Gourville and Koehler (2004), while limited and flawed in many ways, is the most comprehensive study on product downsizing to date. Yet it does not adequately address how successful downsizing is from a revenue or profit standpoint. We thus contribute to the literature by undertaking several empirical studies on a rich set of scanner data.

We conducted several empirical investigations to determine the impact of hidden price changes on General Mills' cereal business. Out of the 20 General Mills products in our sample that downsized, about half the products increased expenditure share in the category by more than what the demand model predicts, indicating a sufficient number of consumers didn't notice the product downsizing for these goods, or that these products cannibalized sales from the other downsized products that became too small for consumers' tastes.

A key finding is that the products that did benefit from the downsizing tended to be large-sized boxes (by weight), while those that did not benefit from the downsizing tended to be small-sized boxes. One explanation is that consumers are more likely to notice when already small boxes get smaller, and substitute to the larger boxes. Moreover, when examining different size products within the same brand-line (e.g., Cheerios 10oz., 15oz., and 20oz.), the larger-sized box is the one that benefitted the most from the downsizing relative to its smaller-sized sibling. In fact, this pattern held nearly perfectly within each of the four brand groupings. A potential explanation here is that consumers don't notice hidden price changes, but because they prefer boxes not to be too big nor too small, they substitute away from small boxes that have been downsized to large boxes that have been downsized.

Because the downsizing by General Mills in 2007 was due to an exogenous increase in its input costs, it is arguably a natural experiment that lends itself to diff-in-diff estimation. After controlling for a number of confounding factors, we find that the downsizing had a negligible effect on the expenditure share for the 20 General Mills products as a whole. While this result may appear to suggest that consumers noticed the hidden price change, it is more consistent with the notion that some of General Mills products benefitted because enough consumer didn't notice, while other products did not benefit because consumers did notice, thus leading to a negligible net effect. This explanation is in line with the results from the expenditure share predictions from the demand model that finds consumers may have substituted from smaller size boxes up to the larger size boxes.

Finally, we examined Chapter 2's investigation of General Mills' change in profitability due to the downsizing. Two of his main findings are consistent with accounting profits and what has transpired in the cereal industry since the end of 2007. His profitability estimates for General Mills before and after the downsizing are approximately consistent with accounting profits reported by the firm. His finding that General Mills' profits increased by nearly 4% after the downsizing is expected if increasing prices was the appropriate profit maximizing action to undertake when they incurred higher input costs in 2006 and 2007. Lastly, the only firm whose profitability declined after General Mills' downsizing was Kellogg, which is the only firm that also committed downsizing starting in 2008, most likely in an attempt to re-attain profit maximization as General Mills did.

4.6 Tables and Figures

	Petroleum		Corn		Rice		Sugar		Wheat	
2003	46		102		70		92		87	
2004	61	31%	109	6%	86	23%	88	-4%	94	7%
2005	86	41%	96	-12%	101	17%	90	2%	91	-3%
2006	103	20%	118	24%	107	5%	95	5%	115	26%
2007	114	11%	159	34%	117	10%	89	-6%	153	33%

Source: International Monetary Fund, Energy and Commodities Surveillance Unit.
The first column is a price index that is constructed from monthly data by taking the average for each year. January 2006 is indexed to 100.
The second column is the year-over-year percentage change.

Figure 4.1: Commodity Prices

	Market Share - DS	Market Share - PS	Number of SKUs	Wtd. Avg. Price Per Lb.	% Above/Below Total Avg. Price
Kellogg	37.9%	37.9%	55	$2.95	-0.1%
General Mills	34.9%	30.3%	45	$3.40	15.2%
Post	14.7%	16.0%	25	$2.71	-8.5%
Quaker	7.6%	8.4%	13	$2.66	-10.2%
Store Brands	5.0%	7.3%	12	$2.01	-31.9%
Total	100.0%	100.0%	150	$2.96	

Source: IRI scanner data. Based on 121 city-supermarket chains, 2005-2007. SKU is "stock keeping unit."
Market Share - DS is based on dollar sales. Market Share - PS is based on sales in pounds of cereal.

Figure 4.2: Market Share and Price

Product	Date of Downsizing	Ounces Per Box Bef	Aft	Chg	Implied Price Chg	Price Per Pound Bef	Aft	Chg	Price Per Box Bef	Aft	Chg	Expenditure Share Bef	Aft	Chg
CHEERIOS BOX 10OZ	07/19/07	10	8.9	-11.0%	12.4%	4.73	5.29	11.9%	2.95	2.96	0.2%	0.019	0.019	0.1%
CHEERIOS BOX 15OZ	07/19/07	15	14	-6.7%	7.1%	3.46	3.80	9.8%	3.25	3.34	2.7%	0.027	0.022	-17.2%
CHEERIOS BOX 20OZ	07/11/07	20	18	-10.0%	11.1%	3.53	3.80	7.6%	4.41	4.29	-2.8%	0.016	0.020	22.7%
CHEERIOS FROSTED BOX 20.25OZ	07/22/07	20.25	17.2	-15.1%	17.7%	3.02	3.64	20.5%	3.82	3.97	3.9%	0.005	0.005	-1.4%
CHEERIOS HONEY NUT BOX 14OZ	07/24/07	14	12.25	-12.5%	14.3%	3.72	4.25	14.0%	3.26	3.27	0.4%	0.025	0.026	1.4%
CHEERIOS HONEY NUT BOX 20OZ	07/13/07	20	17	-15.0%	17.6%	3.56	4.17	17.1%	4.45	4.47	0.4%	0.017	0.017	-1.5%
CHEERIOS HONEY NUT BOX 27OZ	07/12/07	27	25.25	-6.5%	6.9%	3.13	3.40	8.7%	5.28	5.38	1.8%	0.014	0.015	8.9%
CHEX RICE BOX 15.6OZ	07/20/07	15.6	12.8	-17.9%	21.9%	3.63	3.91	7.8%	3.54	3.18	-10.2%	0.007	0.008	15.6%
CINNAMON TOAST CRUNCH BOX 14OZ	08/04/07	14	12.8	-8.6%	9.4%	3.60	3.73	3.8%	3.20	3.00	-6.0%	0.015	0.014	-2.1%
CINNAMON TOAST CRUNCH BOX 20.25OZ	07/20/07	20.25	17	-16.0%	19.1%	3.51	3.91	11.4%	4.44	4.21	-5.1%	0.010	0.010	2.4%
COCOA PUFFS BOX 13.75OZ	07/25/07	13.75	11.8	-14.2%	16.5%	3.68	3.89	5.4%	3.21	2.91	-9.5%	0.009	0.009	4.0%
COOKIE CRISP BOX 12.25OZ	08/04/07	12.25	11.25	-8.2%	8.9%	5.01	5.18	3.4%	3.84	3.67	-4.5%	0.010	0.008	-23.4%
GOLDEN GRAHAMS BOX 13OZ	08/13/07	13	12	-7.7%	8.3%	4.10	3.88	-5.3%	3.34	2.93	-12.2%	0.006	0.007	14.5%
LUCKY CHARMS BOX 14OZ	07/22/07	14	11.5	-17.9%	21.7%	3.84	4.52	18.0%	3.39	3.30	-2.7%	0.014	0.014	-0.3%
LUCKY CHARMS BOX 20OZ	07/19/07	20	16	-20.0%	25.0%	3.80	4.42	16.4%	4.75	4.50	-5.3%	0.009	0.010	8.6%
OATMEAL CRISP RAISIN BOX 19.25OZ	08/08/07	19.25	18	-6.5%	6.9%	3.35	3.52	5.0%	4.03	3.98	-1.2%	0.004	0.004	-10.6%
REESES PNT BTR PUFFS BOX 14.25OZ	07/22/07	14.25	13	-8.8%	9.6%	4.08	4.13	1.3%	3.64	3.37	-7.4%	0.008	0.009	2.8%
TOTAL WHOLE GRAIN BOX 12OZ	08/05/07	12	10.6	-11.7%	13.2%	4.83	5.11	5.8%	3.62	3.43	-5.2%	0.006	0.005	-15.7%
TRIX BOX 12OZ	08/06/07	12	10.7	-10.8%	12.1%	4.13	4.23	2.5%	3.15	2.87	-8.7%	0.008	0.008	-7.7%
WHEATIES BOX 18OZ	07/23/07	18	15.6	-13.3%	15.4%	3.53	4.11	16.4%	3.97	4.06	2.1%	0.007	0.007	-3.6%

Source: IRI scanner data. Based on 121 city-supermarket chains, 2005-2007.

Figure 4.3: General Mills Downsizing

Dependent Var: Qty Boxes Purchased	Price Per Box Coeff	Tstat	Pval	Box Size Coeff	Tstat	Pval	Regression Stats Obs	R-Sqd	RMSE
CHEERIOS BOX 10OZ	-3657.0	-27.53	0.00	-236.9	-1.14	0.25	18760	0.47	2655.4
CHEERIOS BOX 15OZ	-3401.2	-40.58	0.00	-806.7	-3.45	0.00	18987	0.52	3162.4
CHEERIOS BOX 20OZ	-1607.6	-22.03	0.00	-71.4	-1.83	0.07	18995	0.47	1480.4
CHEERIOS FROSTED BOX 20.25OZ	-568.9	-27.08	0.00	-6.2	-0.80	0.43	17654	0.48	535.7
CHEERIOS HONEY NUT BOX 14OZ	-4191.5	-40.93	0.00	-68.4	-0.46	0.65	18997	0.50	3844.6
CHEERIOS HONEY NUT BOX 20OZ	-2042.4	-21.33	0.00	-63.8	-1.92	0.06	18983	0.46	1827.5
CHEERIOS HONEY NUT BOX 27OZ	-938.5	-17.16	0.00	-72.1	-1.74	0.08	17508	0.43	1150.9
CHEX RICE BOX 15.6OZ	-499.7	-34.97	0.00	-6.0	-0.71	0.48	18024	0.62	450.6
CINNAMON TOAST CRUNCH BOX 14OZ	-2635.7	-40.59	0.00	-8.1	-0.20	0.84	18996	0.41	2790.3
CINNAMON TOAST CRUNCH BOX 20.25OZ	-1007.1	-28.21	0.00	-15.4	-0.64	0.52	18992	0.44	1088.0
COCOA PUFFS BOX 13.75OZ	-1490.6	-30.52	0.00	-19.6	-0.53	0.60	18636	0.40	1583.0
COOKIE CRISP BOX 12.25OZ	-1269.5	-29.11	0.00	-144.0	-3.53	0.00	18891	0.43	1470.1
GOLDEN GRAHAMS BOX 13OZ	-1135.1	-31.86	0.00	-20.6	-0.58	0.57	17602	0.39	1144.5
LUCKY CHARMS BOX 14OZ	-1983.5	-43.36	0.00	-21.5	-0.49	0.63	18997	0.44	2271.3
LUCKY CHARMS BOX 20OZ	-902.4	-26.73	0.00	11.8	0.91	0.36	18843	0.44	951.8
OATMEAL CRISP RAISIN BOX 19.25OZ	-404.7	-20.11	0.00	-17.0	-0.88	0.38	17751	0.51	384.8
REESES PNT BTR PUFFS BOX 14.25OZ	-1104.6	-38.61	0.00	3.6	0.09	0.93	18865	0.43	1223.5
TOTAL WHOLE GRAIN BOX 12OZ	-913.4	-16.52	0.00	-143.3	-2.00	0.05	18455	0.38	1005.0
TRIX BOX 12OZ	-1435.9	-30.79	0.00	-6.0	-0.16	0.88	18728	0.37	1509.3
WHEATIES BOX 18OZ	-635.3	-20.02	0.00	-18.5	-1.79	0.07	18418	0.53	541.5

T-stats and P-values are based on White (1980) heteroskedasticity-robust standard errors.

Figure 4.4: Demand - GK's Specification

	Price Per Box			Box Size			Regression Stats		
Dependent Var: Qty Boxes Purchased	Coeff	Tstat	Pval	Coeff	Tstat	Pval	Obs	R-Sqd	RMSE
CHEERIOS BOX 10OZ	-2831.7	-17.46	0.00	0.5	0.00	1.00	8251	0.59	2501.6
CHEERIOS BOX 15OZ	-3077.2	-19.24	0.00	-830.3	-2.69	0.01	8251	0.61	3047.0
CHEERIOS BOX 20OZ	-1251.8	-14.33	0.00	43.6	0.91	0.37	8251	0.62	1286.1
CHEERIOS FROSTED BOX 20.25OZ	-480.4	-14.87	0.00	-32.5	-2.52	0.01	8251	0.54	556.8
CHEERIOS HONEY NUT BOX 14OZ	-3018.2	-20.46	0.00	62.9	0.41	0.68	8251	0.59	3380.3
CHEERIOS HONEY NUT BOX 20OZ	-1245.6	-13.50	0.00	39.9	1.26	0.21	8251	0.63	1430.5
CHEERIOS HONEY NUT BOX 27OZ	-897.5	-13.98	0.00	142.7	2.49	0.01	8251	0.58	1083.1
CHEX RICE BOX 15.6OZ	-452.7	-19.10	0.00	-7.7	-0.62	0.54	8251	0.69	437.1
CINNAMON TOAST CRUNCH BOX 14OZ	-2026.5	-20.63	0.00	34.3	0.61	0.54	8251	0.54	2167.9
CINNAMON TOAST CRUNCH BOX 20.25OZ	-518.3	-14.33	0.00	34.1	1.50	0.14	8251	0.63	663.7
COCOA PUFFS BOX 13.75OZ	-1151.6	-18.47	0.00	-109.2	-2.78	0.01	8251	0.52	1331.5
COOKIE CRISP BOX 12.25OZ	-964.0	-16.62	0.00	-66.9	-1.31	0.19	8251	0.57	1156.7
GOLDEN GRAHAMS BOX 13OZ	-906.9	-17.68	0.00	62.1	2.15	0.03	8251	0.50	1091.4
LUCKY CHARMS BOX 14OZ	-1481.3	-21.01	0.00	32.3	0.54	0.59	8251	0.57	1701.7
LUCKY CHARMS BOX 20OZ	-565.9	-14.42	0.00	-15.6	-1.29	0.20	8251	0.64	579.4
OATMEAL CRISP RAISIN BOX 19.25OZ	-314.1	-12.52	0.00	10.5	0.57	0.57	8251	0.61	346.8
REESES PNT BTR PUFFS BOX 14.25OZ	-879.6	-16.56	0.00	43.9	0.70	0.48	8251	0.47	1239.4
TOTAL WHOLE GRAIN BOX 12OZ	-736.8	-8.53	0.00	17.6	0.25	0.80	8251	0.46	847.0
TRIX BOX 12OZ	-1206.4	-21.04	0.00	-21.9	-0.79	0.43	8251	0.54	1179.3
WHEATIES BOX 18OZ	-549.5	-12.40	0.00	-42.0	-2.82	0.01	8251	0.59	551.6

T-stats and P-values are based on White (1980) heteroskedasticity-robust standard errors.

Figure 4.5: Demand - Full Specification

Product	Mean Expend. Share (After) Actual	Predict	Ratio	Diff	P-Val	Regression Stats Obs.	R-Sqd	RMSE
CHEERIOS BOX 10OZ	2.95%	5.02%	0.59	-2.07%	0.000	7152	0.696	0.011
CHEERIOS BOX 15OZ	3.73%	3.98%	0.94	-0.26%	0.009	7145	0.759	0.013
CHEERIOS BOX 20OZ	3.28%	2.69%	1.22	0.59%	0.000	7114	0.749	0.007
CHEERIOS FROSTED BOX 20.25OZ	0.89%	0.76%	1.17	0.13%	0.000	7160	0.708	0.003
CHEERIOS HONEY NUT BOX 14OZ	4.15%	7.23%	0.57	-3.07%	0.000	7224	0.777	0.014
CHEERIOS HONEY NUT BOX 20OZ	2.61%	1.89%	1.39	0.73%	0.000	7125	0.785	0.008
CHEERIOS HONEY NUT BOX 27OZ	2.69%	2.06%	1.30	0.62%	0.000	7100	0.757	0.007
CHEX RICE BOX 15.6OZ	1.23%	0.06%	21.79	1.18%	0.000	7143	0.769	0.004
CINNAMON TOAST CRUNCH BOX 14OZ	2.33%	3.87%	0.60	-1.53%	0.000	7292	0.768	0.010
CINNAMON TOAST CRUNCH BOX 20.25OZ	1.61%	1.29%	1.25	0.32%	0.000	7157	0.756	0.005
COCOA PUFFS BOX 13.75OZ	1.37%	0.82%	1.68	0.55%	0.000	7178	0.726	0.005
COOKIE CRISP BOX 12.25OZ	1.21%	1.06%	1.14	0.15%	0.002	7231	0.791	0.005
GOLDEN GRAHAMS BOX 13OZ	1.07%	1.33%	0.81	-0.25%	0.000	7245	0.704	0.005
LUCKY CHARMS BOX 14OZ	2.09%	1.74%	1.20	0.35%	0.000	7143	0.737	0.008
LUCKY CHARMS BOX 20OZ	1.51%	0.62%	2.43	0.89%	0.000	7137	0.728	0.005
OATMEAL CRISP RAISIN BOX 19.25OZ	0.68%	0.73%	0.94	-0.05%	0.001	7240	0.716	0.002
REESES PNT BTR PUFFS BOX 14.25OZ	1.40%	1.38%	1.01	0.02%	0.755	7159	0.619	0.007
TOTAL WHOLE GRAIN BOX 12OZ	0.84%	1.32%	0.63	-0.48%	0.000	7241	0.682	0.003
TRIX BOX 12OZ	1.30%	1.38%	0.94	-0.08%	0.076	7231	0.759	0.005
WHEATIES BOX 18OZ	1.07%	0.78%	1.38	0.29%	0.000	7165	0.709	0.003

Figure 4.6: AIDS Predictions

Product	Mean Expend. Share (After) Actual	Predict	Ratio	Diff	P-Val
CHEX RICE BOX 15.6OZ	1.23%	0.06%	21.79	1.18%	0.000
LUCKY CHARMS BOX 20OZ	1.51%	0.62%	2.43	0.89%	0.000
COCOA PUFFS BOX 13.75OZ	1.37%	0.82%	1.68	0.55%	0.000
CHEERIOS HONEY NUT BOX 20OZ	2.61%	1.89%	1.39	0.73%	0.000
WHEATIES BOX 18OZ	1.07%	0.78%	1.38	0.29%	0.000
CHEERIOS HONEY NUT BOX 27OZ	2.69%	2.06%	1.30	0.62%	0.000
CINNAMON TOAST CRUNCH BOX 20.25OZ	1.61%	1.29%	1.25	0.32%	0.000
CHEERIOS BOX 20OZ	3.28%	2.69%	1.22	0.59%	0.000
LUCKY CHARMS BOX 14OZ	2.09%	1.74%	1.20	0.35%	0.000
CHEERIOS FROSTED BOX 20.25OZ	0.89%	0.76%	1.17	0.13%	0.000
COOKIE CRISP BOX 12.25OZ	1.21%	1.06%	1.14	0.15%	0.002
REESES PNT BTR PUFFS BOX 14.25OZ	1.40%	1.38%	1.01	0.02%	0.755
TRIX BOX 12OZ	1.30%	1.38%	0.94	-0.08%	0.076
OATMEAL CRISP RAISIN BOX 19.25OZ	0.68%	0.73%	0.94	-0.05%	0.001
CHEERIOS BOX 15OZ	3.73%	3.98%	0.94	-0.26%	0.009
GOLDEN GRAHAMS BOX 13OZ	1.07%	1.33%	0.81	-0.25%	0.000
TOTAL WHOLE GRAIN BOX 12OZ	0.84%	1.32%	0.63	-0.48%	0.000
CINNAMON TOAST CRUNCH BOX 14OZ	2.33%	3.87%	0.60	-1.53%	0.000
CHEERIOS BOX 10OZ	2.95%	5.02%	0.59	-2.07%	0.000
CHEERIOS HONEY NUT BOX 14OZ	4.15%	7.23%	0.57	-3.07%	0.000

See Table VI.

Figure 4.7: AIDS Predictions, Sorted by Ratio

Product	Actual	Mean Expend. Share (After) Predict	Ratio	Diff	P-Val
CHEERIOS BOX 20OZ	3.28%	2.69%	1.22	0.59%	0.000
CHEERIOS BOX 15OZ	3.73%	3.98%	0.94	-0.26%	0.009
CHEERIOS BOX 10OZ	2.95%	5.02%	0.59	-2.07%	0.000
CHEERIOS HONEY NUT BOX 20OZ	2.61%	1.89%	1.39	0.73%	0.000
CHEERIOS HONEY NUT BOX 27OZ	2.69%	2.06%	1.30	0.62%	0.000
CHEERIOS HONEY NUT BOX 14OZ	4.15%	7.23%	0.57	-3.07%	0.000
CINNAMON TOAST CRUNCH BOX 20.25OZ	1.61%	1.29%	1.25	0.32%	0.000
CINNAMON TOAST CRUNCH BOX 14OZ	2.33%	3.87%	0.60	-1.53%	0.000
LUCKY CHARMS BOX 20OZ	1.51%	0.62%	2.43	0.89%	0.000
LUCKY CHARMS BOX 14OZ	2.09%	1.74%	1.20	0.35%	0.000

See Table VII.

Figure 4.8: AIDS Predictions, Multiple Sizes

	(1)		(2)		(3)		(4)		(5)		(6)	
Dep. Var.: Expenditure Share x 100	Coeff.	P-Val.	Coeff.	P-Val.	Coeff.	P-Val.	Coeff.	P-Val.	Coeff.	P-Val.	Coeff.	P-Val.
1[Downsize]	0.547	0.00	0.537	0.00	0.378	0.00	0.408	0.00	0.394	0.00	0.083	0.00
1[GM 20 Products]	0.338	0.00	0.468	0.00	0.355	0.00	0.349	0.00	0.350	0.00	1.399	0.00
1[Downsize] x 1[GM 20 Products]	-0.546	0.00	-0.447	0.00	-0.351	0.00	-0.344	0.00	-0.326	0.00	0.025	0.00
Price			-0.272	0.00	-0.115	0.00	-0.115	0.00	-0.123	0.00	-0.306	0.00
Ad Only					0.162	0.00	0.165	0.00	0.191	0.00	0.114	0.00
Display Only					1.706	0.00	1.710	0.00	1.655	0.00	1.390	0.00
Ad and Display					3.132	0.00	3.136	0.00	3.122	0.00	2.692	0.00
Distribution					1.798	0.00	1.798	0.00	2.272	0.00	1.411	0.00
Trend							-0.001	0.00	-0.001	0.00	-0.001	0.00
Month Fixed Effects	n		n		n		y		y		y	
City-Retailer Fixed Effects	n		n		n		n		y		y	
Product Fixed Effects	n		n		n		n		n		y	
Other Covariates	n		n		n		n		n		y	
N	1,203,794		1,203,794		1,203,794		1,203,794		1,203,794		1,203,794	
R-Sqd	0.031		0.121		0.426		0.427		0.445		0.595	
RMSE	0.909		0.866		0.700		0.699		0.689		0.588	
AIC	3,187,593		3,070,339		2,558,204		2,555,122		2,517,900		2,137,608	
AIC/N	2.648		2.551		2.125		2.123		2.092		1.776	

P-values are based on White (1980) heteroskedasticity-robust standard errors. "Other Covariates" includes weeks elapsed since last promotion, percentage of populatio under age 19, percentage of population over age 55, percentage of households with children, percentage of population that is white, income (weekly wage).

Figure 4.9: Diff-in-Diff Results

	Before	After	Change
General Mills	66.6%	70.2%	3.6%
Kellogg	73.0%	70.2%	-2.8%
Post	60.2%	64.4%	4.2%
Quaker	60.8%	62.7%	1.9%
Store Brand	56.7%	56.8%	0.0%
All 64 Products	67.3%	68.2%	0.9%

Source: Chapter 2, Figure 2.13.

Figure 4.10: Price-Cost Margins

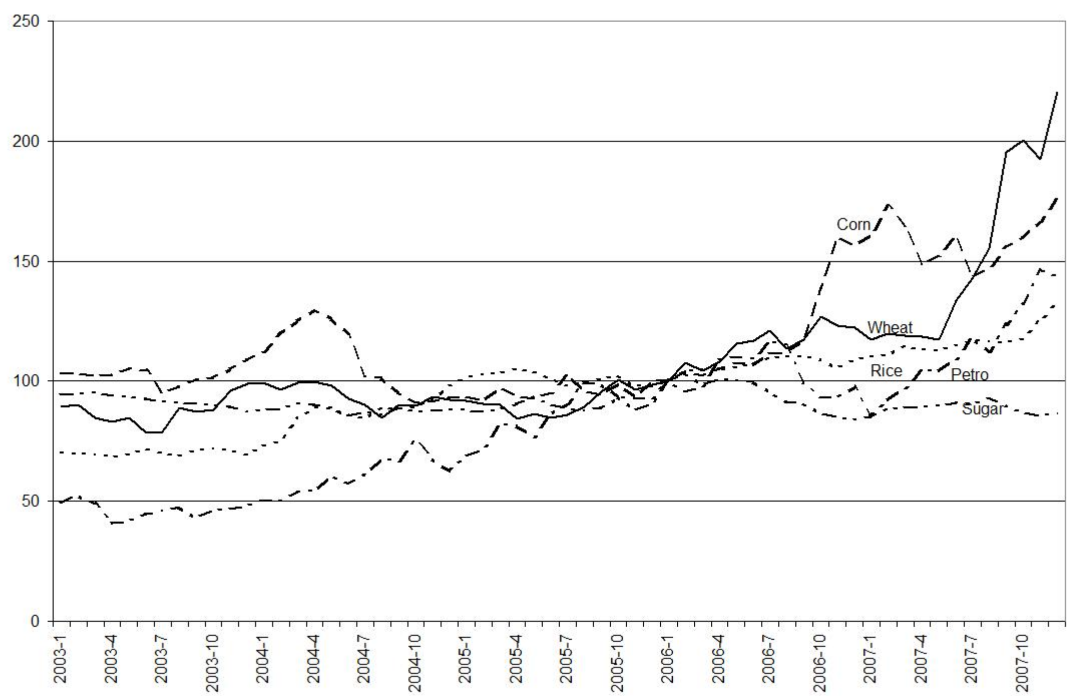


Figure 4.11: Input Commodity Costs

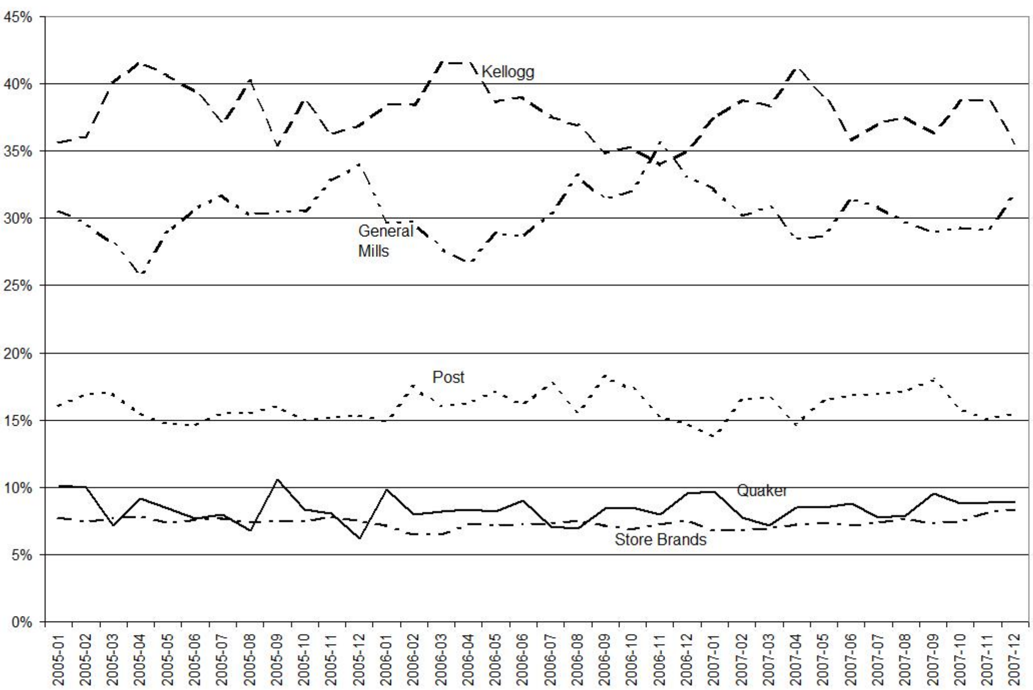


Figure 4.12: Market Share

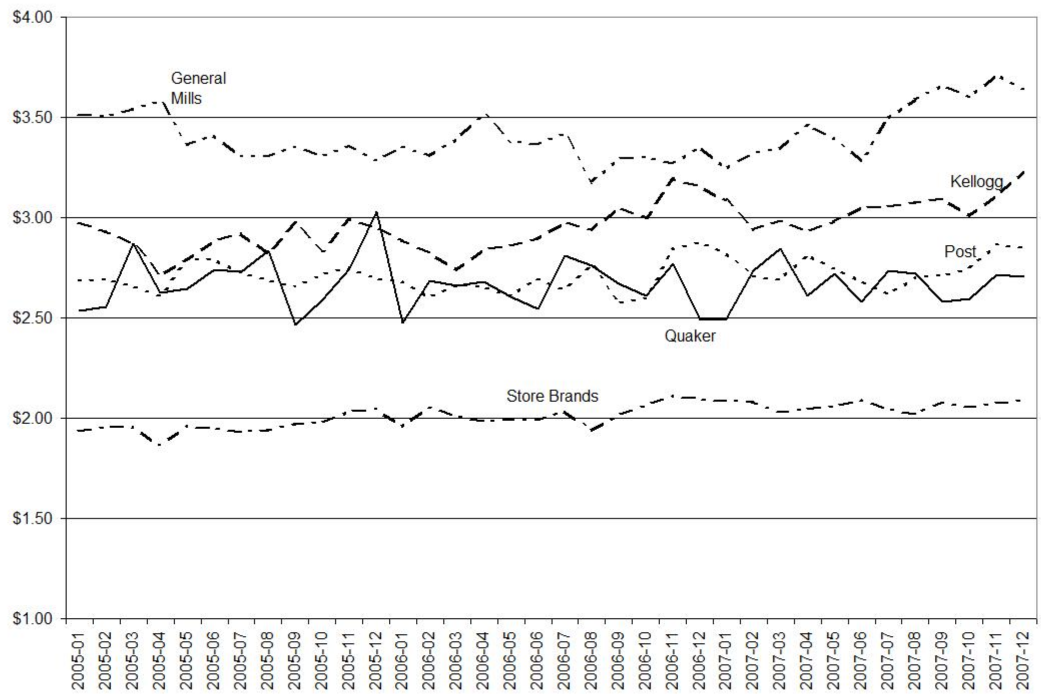


Figure 4.13: Price Per Pound

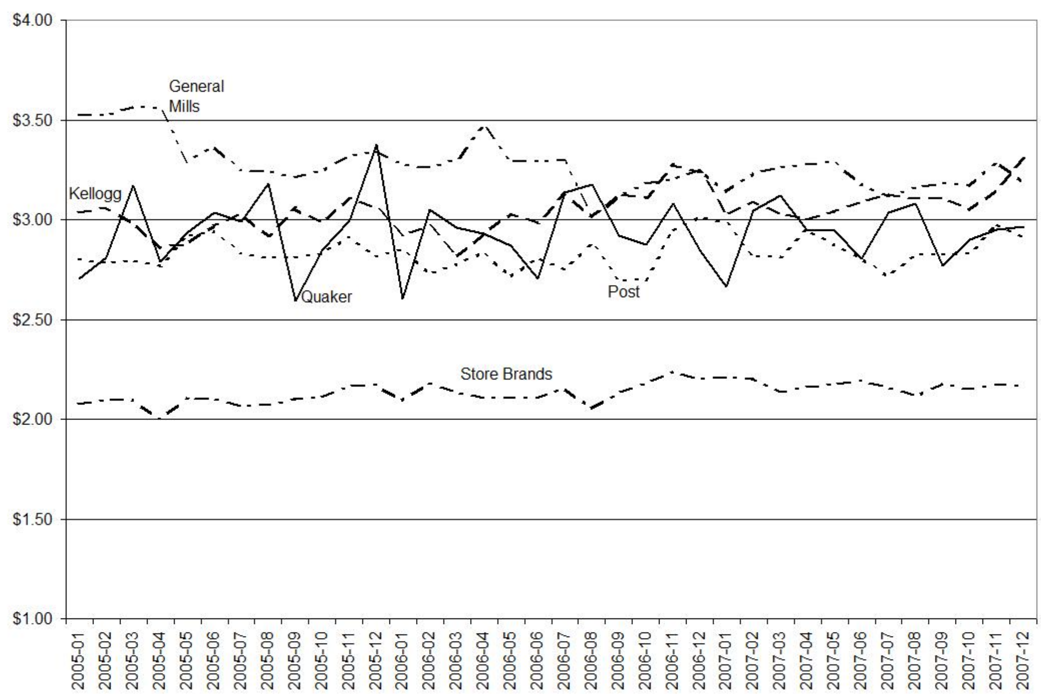


Figure 4.14: Price per Box

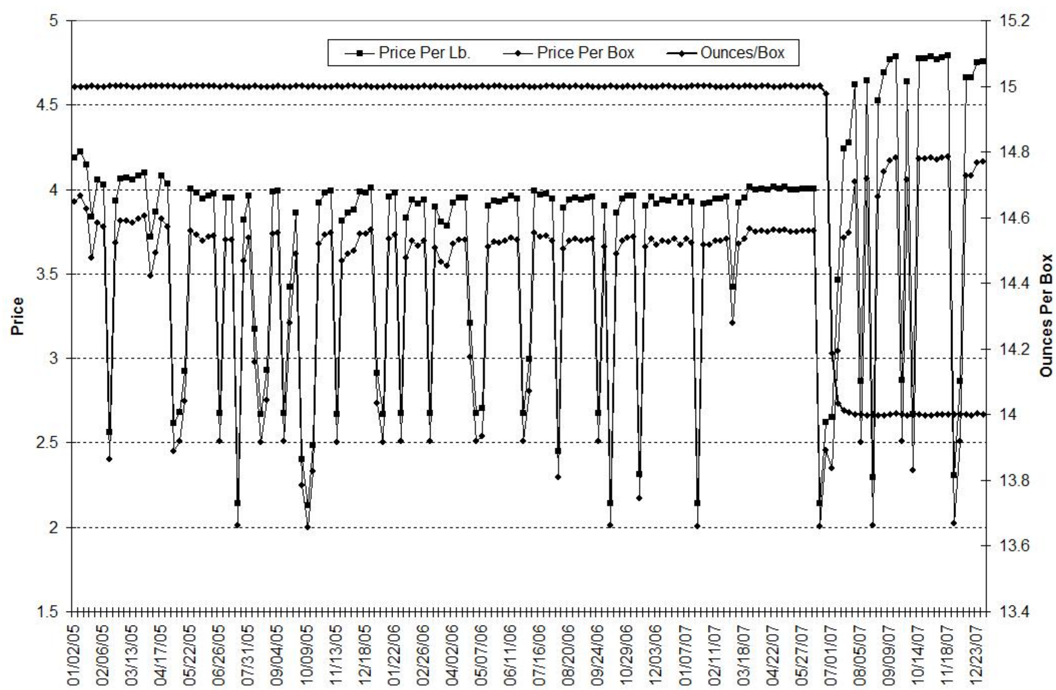


Figure 4.15: Cheerios 15oz. to 14oz. Downsize

Bibliography

[1] Ackerberg, Daniel A., and Gregory S. Crawford. (2008) "Estimating Price Elasticities in Differentiated Product Demand Models with Endogenous Characteristics." Working Paper.

[2] Adams, Anthony C., Anthony di Benedetto, and Rajan Chandran. (1991) "Can You Reduce Your Package Size Without Damaging Sales?" *Long Range Planning.* 24(4), 86-96.

[3] Altonji, J. and R. Matzkin. (2005) "Cross Section and Panel Data Estimators for Nonseparable Models with Endogenous Regressors." *Econometrica.* 73, 1053-1102.

[4] Angrist, Joshua, and Jorn-Steffen Pischke. (2010) "The Credibility Revolution in Empirical Economics: How Better Research Design is Taking the Con Out of Econometrics." NBER Working Paper 15794.

[5] Ashenfelter, Orley, and Daniel Hosken. (2008) "The Effect of Mergers on Consumer Prices: Evidence from Five Selected Case Studies." NBER Working Paper 13859.

[6] Baker, Jonathan B., and Timothy F. Bresnahan. (1985) "The Gains from Merger or Collusion in Product-Differentiated Industries." *The Journal of Industrial Economics.* 33(4), 427-444.

[7] Berry, Steven T. (1994) "Estimating Discrete Choice Models of Product Differentiation." *Rand Journal of Economics.* 25(2), 242-262.

[8] Berry, Steven T., and Philip A. Haile. (2010) "Identification in Differentiated Products Markets Using Market Level Data." NBER Working Paper 15641.

[9] Berry, Steven T., James Levinsohn, and Ariel Pakes. (1995) "Automobile Prices in Equilibrium." *Econometrica.* 63(4), 841-890.

[10] Besanko, David, Jean-Pierre Dube, and Sachin Gupta. (2005) "Own-Brand and Cross-Brand Retail Pass-Through." *Marketing Science.* 24(1), 123-137.

[11] Besanko, David, Sachin Gupta, and Dipak Jain. (1998) "Logit Demand Estimation Under Competitive Pricing Behavior: An Equilibrium Framework." *Management Science*. 44(11), 1533-1547.

[12] Blundell, Richard, and James Powell. (2003) "Endogeneity in Nonparametric and Semiparametric Regression Models", in M. Dewatripont, L. Hansen and S.J. Turnsovsky (eds), *Advances in Economics and Econometrics*, Cambridge: Cambridge University Press.

[13] Bound, John, David A. Jaeger, and Regina M. Baker. (1995) "Problems with Instrumental Variables Regression when the Correlation Between the Instruments and the Endogenous Explanatory Variable is Weak." *Journal of the American Statistical Society*. 90(430), 443-450.

[14] Bresnahan, Timothy. (1997) "Comment" in Bresnahan, Timothy and Robert J. Gordon, eds. *The Economics of New Goods: NBER Studies in Income and Wealth.* Volume 58, 237-47.

[15] Chalak, Karim, and Halbert White. (2010) "An Extended Class of Instrumental Variables for the Estimation of Causal Effects." UCSD Working Paper.

[16] Chalfant, James A. (1987) "A Globally Flexible Almost Ideal Demand System." *Journal of Business and Economic Statistics*. 5(2), 233-242.

[17] Chevalier, Judith A., Anil K. Kashyap, and Peter E. Rossi. (2003) "Why Don't Prices Rise During Periods of Peak Demand? Evidence from Scanner Data." *American Economic Review*. 93(1), 15-37.

[18] Chintagunta, Pradeep K. (2002) "Investigating Category Pricing Behavior at a Retail Chain." *Journal of Marketing Research*. Vol 39, 141-154.

[19] Dawid, A. P. (1979) "Conditional Independence in Statistical Theory." *Journal of the Royal Statistical Society*. 41(1), 1-31.

[20] Deaton, Angus, and John Muellbauer. (1980) "An Almost Ideal Demand System." *American Economic Review*. 70(3), 312-326.

[21] Eichenbaum, Martin, Nir Jaimovich, and Sergio Rebelo. (2008) "Reference Prices and Nominal Rigidities." Working Paper.

[22] Gourville, John T., and Jonathan J. Koehler. (2004) "Downsizing Price Increases: A Greater Sensitivity to Price than Quantity in Consumer Markets." Harvard Business School Marketing Research Paper No. 04-01.

[23] Granger, C.W., and A. Billson. (1972) "Attitudes Toward Package Size and Price." *Journal of Marketing Research*. 9(3), 239-248.

[24] Green, Richard, and Julian M. Alston. (1990) "Elasticities in AIDS Models." *American Journal of Agricultural Economics.* 72(2), 442-445.

[25] Gupta, Omprakash K., Sudhir Tandon, Sukumar Debnath, and Anna S. Rominger. (2007) "Package Downsizing: Is it Ethical?" *AI and Society.* 21(3), 239-250.

[26] Hahn, Jinyoung, and Geert Ridder. (2008) "Conditional Moment Restrictions and Triangular Simultaneous Equations." University of Southern California, IEPR Working Paper.

[27] Hamilton, James D. (1994) *Time Series Analysis.* Princeton University Press.

[28] Hayashi, Fumio. (2000) *Econometrics.* Princeton University Press.

[29] Hausman, Jerry A. (1978) "Specification Tests in Econometrics." *Econometrica.* 46(6), 1251-71.

[30] Hausman, Jerry A. (1997) "Valuation of New Goods Under Perfect and Imperfect Competition." in Bresnahan, Timothy and Robert J. Gordon, eds. *The Economics of New Goods: NBER Studies in Income and Wealth.* Volume 58, 209-37.

[31] Hausman, Jerry A., and William E. Taylor. (1981) "Panel Data and Unobservable Individual Effects." *Econometrica.* 49(6), 1377-98.

[32] Hausman, Jerry A., Leonard, Gregory K., and J.D. Zona. (1994) "Competitive Analysis with Differentiated Products." *Annales d'Economie et de Statistique.* 34, 159-80.

[33] Hausman, Jerry A., and Gregory K. Leonard. (2002) "The Competitive Effects of a New Product Introduction: A Case Study." *Journal of Industrial Economics.* Volume L, No. 3, 237-263.

[34] Hausman, Jerry A., and Gregory K. Leonard. (2005) "Competitive Analysis Using a Flexible Demand Specification." *Journal of Competition Law and Economics.* 1(2), 279-301.

[35] Hausman, Jerry A., and Daniel McFadden. (1984) "Specification Tests for the Multinomial Logit Model." *Econometrica.* 52(5), 1219-1240.

[36] Hausman, Jerry A., and William E. Taylor. (1981) "Panel Data and Unobservable Individual Effects." *Econometrica.* 49(6), 1377-1398.

[37] Heckman, J., H. Ichimura, and P. Todd. (1997) "Matching as an Econometric Evaluation Estimator: Evidence from Evaluating a Job Training Program." *Review of Economic Studies.* 64, 605-54.

[38] Heckman J. and E. Vytlacil. (2005) "Structural Equations, Treatment Effects, and Econometric Policy Evaluation." *Econometrica.* 73, 669-738.

[39] Hendel, Igal, and Aviv Nevo (2003). "The Post-Promotion Dip Puzzle: What Do the Data Have to Say?" *Quantitative Marketing and Economics.* (1), 409-424.

[40] Hendel, Igal, and Aviv Nevo. (2006) "Sales and Consumer Inventory." *Rand Journal of Economics.* 37(3), 543-61.

[41] Hoch, Stephen J., Byung-Do Kim, Alan L. Montgomery, and Peter E. Rossi. (1995) "Determinants of Store-Level Price Elasticity." *Journal of Marketing Research*. Vol XXXII, 17-29.

[42] Hoderlein, S. and E. Mammen. (2007) "Identification of Marginal Effects in Nonseparable Models without Monotonicity." *Econometrica*. 75, 1513-1518.

[43] Hosken, Daniel, Daniel O'Brien, David Scheffman, and Michael Vita. (2002) "Demand System Estimation and its Application to Horizontal Merger Analysis." Federal Trade Commission, Working Paper 246.

[44] Hosken, Daniel, and David Reiffen. (2001) "Multiproduct Retailers and the Sale Phenomenon." *Agribusiness*. 17(1), 115-37.

[45] Hosken, Daniel, and David Reiffen. (2004a) "How Retailers Determine Which Products Should Go on Sale: Evidence from Store-Level Data." *Journal of Consumer Policy*. 141-77.

[46] Hosken, Daniel, and David Reiffen. (2004b) "Patterns of Retail Price Variation." *Rand Journal of Economics*. 35(1), 128-46.

[47] Ippolito, Pauline M., and Alan D. Mathios. (1990) "Information, Advertising and Health Choices: A Study of the Cereal Market." *Rand Journal of Economics*. 21(3), 459-480.

[48] Kennedy, Peter. (1992) *A Guide to Econometrics*, Third Edition. MIT Press.

[49] Lal, Rajiv, and Carmen Matutes. (1994) "Retail Pricing and Advertising Strategies." *Journal of Business*. 67(3), 345-70.

[50] Little, John D.C. (1998) "Integrated Measures of Sales, Merchandising, and Distribution." *International Journal of Research in Marketing.* 15, 473-85.

[51] McFadden, Daniel. (1974) "Conditional Logit Analysis of Qualitative Choice Behavior." in P. Zarembka, ed. *Frontiers in Econometrics.* New York: Academic Press, 105-142.

[52] Nevo, Aviv. (2000a) "Mergers and Differentiated Products: The Case of the Ready-To-Eat Cereal Industry." *Rand Journal of Economics.* 31 (3), 395-421.

[53] Nevo, Aviv. (2000b) "A Practitioner's Guide to Estimation of Random Coefficients Logit Models of Demand." *Journal of Economics and Management Strategy.*" 9(4), 513-48.

[54] Nevo, Aviv. (2001) "Measuring Market Power in the Ready to Eat Cereal Industry." *Econometrica.* 69(2), 307-42.

[55] Nevo, Aviv, and Catherine D. Wolfram. (2002) "Why do Manufacturers Issue Coupons? An Empirical Analysis of Breakfast Cereals." *Rand Journal of Economics.* 33 (2), 319-339.

[56] Pearl, Judea. (2000) *Causality.* New York: Cambridge University Press.

[57] Pesendorfer, Martin. (2002) "Retail Sales: A Study of Pricing Behavior in Supermarkets." *Journal of Business.* 75(1), 33-66.

[58] Peters, Craig. (2006) "Evaluating the Performance of Merger Simulation: Evidence from the U.S. Airline Industry." *Journal of Law and Economics.* XLIX, 627-649.

[59] Rosenbaum, P. and D. Rubin. (1983) "The Central Role of the Propensity Score in Observational Studies for Causal Effects." *Biometrika.* 70, 41-55.

[60] Rubin, D. (1974) Estimating Causal Effects of Treatments in Randomized and Nonrandomized Studies." *Journal of Educational Psychology.* 66, 688-701.

[61] Rubinfeld, Daniel L. (2000) "Market Definition with Differentiated Products: The Post-Nabisco Cereal Merger." *Antitrust Law Journal.* 68(1), 163-185.

[62] Scheffman, David T., and Pablo T. Spiller. (1996) "Econometric Market Delineation." *Managerial and Decision Economics.* 17, 165-78.

[63] Schmalensee, Richard. (1978) "Entry Deterrence in the Ready-to-Eat Breakfast Cereal Industry." *Bell Journal of Economics.* 9(2), 305-327.

[64] Shankar, Venkatesh, and Ruth N. Bolton. (2004) "An Empirical Analysis of Determinants of Retailer Pricing Strategy." *Marketing Science.* 23(1), 28-49.

[65] Shepard, Andrea. (1991) "Price Discrimination and Retail Configuration." *Journal of Political Economy.* 99(1), 30-53.

[66] Shum, Matthew. (2004) "Does Advertising Overcome Brand Loyalty? Evidence from the Breakfast-Cereals Market." *Journal of Economics and Management Strategy.* 13(2), 241-272.

[67] Staiger, Douglas, and James H. Stock. (1997) "Instrumental Variables Regression with Weak Instruments." *Econometrica*. 65(3), 557-586.

[68] Su, Liangjun, Stefan Hoderlein, and Halbert White. (2010) "Testing Monotonicity in Unobservables with Panel Data." Working Paper.

[69] Trajtenberg, Manuel. (1989) "The Welfare Analysis of Product Innovations, with an Application to Computed Tomography Scanners." *Journal of Political Economy*. 97(2), 444-479.

[70] Train, Kenneth E. (2003) *Discrete Choice Methods with Simulation*, Cambridge University Press.

[71] Villas-Boas, Sofia B. (2007) "Vertical Relationships Between Manufacturers and Retailers: Inference with Limited Data." *Review of Economic Studies*. 74, 625-652.

[72] Villas-Boas, Miguel J., and Russel. S. Winer. (1999) "Endogeneity in Brand Choice Models." *Management Science*. 45(10), 1324-1338.

[73] Weinberg, Matthew C., and Daniel Hosken. (2008) "Using Mergers to Test a Model of Oligopoly." Working Paper.

[74] Werden, Gregory J. (1998) "Demand Elasticities in Antitrust Analysis." *Antitrust Law Journal*. 66, 363-.

[75] Werden, Gregory J., and Luke M. Froeb. (1994) "The Effects of Mergers in Differentiated Products Industries: Logit Demand and Merger Policy." *The Journal of Law, Economics, and Organization*. 10(2), 407-425.

[76] White, Halbert. (1980) "A Heteroskedasticity-Consistent Covariance Matrix Estimator and a Direct Test of Heteroskedasticity." *Econometrica*. 48(4), 817-838.

[77] White, Halbert. (2006) "Time Series Estimation of the Effects of Natural Experiments." *Journal of Econometrics*. 135, 527-66.

[78] White, Halbert, and Karim Chalak. (2006) "Parametric and Nonparametric Estimation of Covariate-Conditioned Average Causal Effects." UCSD Working Paper.

[79] White, Halbert, and Karim Chalak. (2009) "Settable Systems: An Extension of Pearl's Causal Model with Optimization, Equilibrium, and Learning." *Journal of Machine Learning Research*. 10, 1759-99.

[80] Wright, Philip G. (1928) *The Tariff on Animal and Vegetable Oils*. New York: Macmillan.