Review of Industrial Organization 8: 717–740, 1993

The Inverse Association Between the Margins of Manufacturers and Retailers

ROBERT L. STEINER
3112 Q St. N.W., Washington, D.C. 20007–3027, U.S.A.

Abstract. The margins of manufacturers and retailers are largely determined by the absolute and relative magnitudes of two cross-elasticities that define the willingness of consumers to switch brands within store and to switch stores within brand. When one of these cross-elasticities is high and the other low, margins of firms at the two stages are inversely associated. This phenomenon is widespread but not universal in industries whose retailing segments are imperfectly competitive, as is typically true. The inverse association is inconsistent with "single stage" models which assume that retailing is perfectly competitive and that the derived demand theorem holds. This article explores the dynamics that produce the negative correlation between margins at the two stages, summarizes the empirical evidence and identifies some important areas in which accepted conclusions should be re-examined in light of this relationship.

Key words. Single and dual stage models, inter and intrabrand cross-elasticities, manufacturers' brand domination, retailer domination, retail gross margin.

I. Background

The inverse association between the margins of consumer goods manufacturers and the firms that distribute their products to household consumers is a prevalent although not ubiquitous phenomenon. This relationship has prevailed since the introduction and rapid spread of branding and of manufacturers' brand advertising in the late nineteenth century. It has been experienced by generations of business people, many of whom have put its lessons to work.

Yet strangely, the negative correlation between margins at the two stages – which for brevity will be referred to as the "inverse association" – and its important implications have gone largely unrecognized in the economics literature. An early exception was Marshall's brief observation in *Industry and Trade* (1920) that while retailers were forced to sell popular advertised brands "at prices that barely covered expenses" (p. 301), the manufacturers were selling them at wholesale for relatively high prices.

In the contemporary marketing literature Steiner (1978a, 1978c) presented an informal "dual stage" model that predicted the inverse association, which relationship he had observed in toys and other industries (1973). Subsequently, Albion (1983) and Albion and Farris (1981a, 1981b) have endorsed and expanded this analysis.

Lynch (1986) developed a formal model based on Steiner's dual stage construct. He demonstrates that with a monopolistically competitive retailing segment the elasticity of demand facing a brand's manufacturer can change inversely with that experienced by the brand's retailers. Lynch points out that this result is inconsistent with the predictions of standard models that posit some combination of pure competition and pure monopoly at the two stages and with any model in which the derived demand theorem holds. All such models predict that changes in elasticities and margins at the two levels will be either positively related or uncorrelated but never inversely related. Lynch also finds it theoretically interesting that with a monopolistically competitive retailing segment the behavior of firms at the two levels is not bounded by their conduct in the polar extremes of pure competition and pure monopoly.

'The present article builds on and extends the dual stage model. It shows that the structure of a consumer goods industry and the relative and absolute margins of manufacturers and retailers are largely determined by the magnitudes of two cross-elasticities that define the willingness of consumers to switch brands within store and to switch stores within brand. When the magnitudes are markedly different, margins at the two stages will be negatively related. The article also presents empirical evidence of the prevalence of this inverse relationship and asserts that it requires some amending of presently accepted analytical techniques and conclusions.

We begin by defining and identifying the terms and concepts that will be used. Section III summarizes the extensive empirical support for the inverse association. Section IV lays out the dynamics that produce this result in a dual stage world. It draws in part on the author's own business experience as a consumer goods manufacturer. The following section defines the scope of the inverse relationship and describes the industry structures in which the values of the two key cross-elasticities are roughly similar, causing margins at the two stages to be positively correlated. The concluding section spotlights some of the major areas that require substantial rethinking in light of the negative relationship between margins at the two stages.

II. Some Concepts and Definitions

Commodity and non-commodity categories. We will investigate margin relationships at the two stages in commodity classes and in various structures found in non-commodity industries. Commodity classes are those in which the goods are physically fungible, or virtually so, and consumers recognize this homogeneity both across brands and across stores. Examples are milk, eggs and sugar. If consumers fail to recognize the physical homogeneity, the category is a non-commodity one. This failure can result from the manufacturer's product differentiation efforts (aspirin, gasoline) or from consumer ignorance, as in many physically homogeneous apparel categories (Dardis and Skow, 1969).

Single and dual stage models. Economics does not lack for models of vertical relationships nor for those that posit some form of imperfect competition. Still, there is a strong tendency to employ "single stage" models to analyze consumer goods industries. In this methodology the wholesale/retail markets that intervene between consumer goods manufacturers and household consumers are ignored by the usually implicit assumption that they are inert and perfectly competitive. The manufacturer's or factory price PM is then a reasonable and unbiased proxy for the price consumers pay PC.

In real consumer goods industries wholesalers and retailers have a degree of market power and face downward sloping demand schedules. Stores are differentiated by location, reputation and product assortment, while large chains often enjoy economies of scale and scope not attainable by independent merchants. Moreover, retailers often have market power as buyers. This upstream leverage arises because the merchant who desires to stock, say, 4 brands in a product category often finds that he can select from among perhaps 20 brands offered by the category's manufacturers.

Therefore, the markets downstream from the manufacturer are best thought of as monopolistically competitive with varying degrees of oligopoly and monopsony, depending on the product class. Consumer goods industries can be appropriately analyzed through a simplified dual stage model in which manufacturers sell to independent "retailers" who resell to household consumers.[1] Retail firms perform all the distributive functions necessary to move goods from factories to households.

Dual stage effects. The dual stage manufacturer's demand schedule is shaped by three parameters that play little or no role when producers sell to consumers directly or through an inert retailing segment. Retail penetration is a measure of a brand's distribution. A brand with an X% retail penetration is distributed in stores that together account for X% of category volume. Dealer support is the store display, local advertising and other promotional efforts that retailers place behind a brand. Retail gross margin (RGM) is the difference between the brand's consumer (retail) price PC and its factory selling price PM (the retailer's invoice cost in a dual stage world) divided by PC. This same ratio is termed gross distribution margin (GDM) when manufacturers also sell to wholesalers.

Note that RGM is the margin over retailers' invoice costs. Their non-invoice marginal costs are excluded. Hence, RGM overstates the true retail margin and understates E_{DR}, although not materially, since the non-invoice portion of retailers' total marginal cost is small (Preston, 1963), under 10% (Farris, 1993).[2]

Derived demand and demand elasticities. In a single stage world with its inert and perfectly competitive retailing sector, a brand's consumer level demand schedule DK specifies the quantities demanded at each retail price by a constant size group of consumers, exposed to some constant, nominal level of dealer support. The manufacturer's demand schedule DM is derived from DK by subtracting the cost

of distribution including a competitive markup at each quantity. Since the elasticity of demand faced by the retailer (E_{DR}) is infinite, a change in E_{DK} is uncorrelated with E_{DR}, and through the derived demand theorem E_{DK} and E_{DM} are positively associated (See Lynch, 1986).

In a dual stage world a brand's true consumer level demand curve DC is more elastic than DK because DC reflects that retail penetration and dealer support vary inversely with PM (Steiner, 1984, pp. 183–185). Moreover, except in the mutual dependence structure (Section 5-A), changes in E_{DR} tend to induce inverse changes in E_{DM}; and when E_{DM} falls (rises), it also decreases (increases) relative to E_{DC}.

Horizontal and vertical components of market power. In a dual stage environment the market power and the margins of an individual manufacturer or retailer are a joint function of its horizontal competitive position against firms at the same level and its vertical bargaining power with firms at the other stage. A firm's market share is a rough surrogate for the former. Some sources of retailer market power have already been noted. Manufacturers' market power is generally attributed to such potential entry barriers as scale economies, high capital to sales requirements, patents and ownership of a popular, trademarked advertised brand.

Vertically, manufacturers and retailers vie with one another to increase their respective share of a brand's retail price and thus to capture a larger portion of the available rents in the vertical system. RGM is a reasonable surrogate for the vertical position of retailers and one minus RGM for that of manufacturers.[3] To illustrate, Steiner (1991a) has shown that holding constant the vigor of competition at the manufacturing stage and consumer utility functions, a monopolist manufacturer's margin will rise when intrabrand competition among the retail resellers of the brand becomes intensified and its RGM falls. For a non-monopolist manufacturer, a below-industry average RGM not only increases the firm's margin but improves its horizontal competitive position by forcing rival, higher RGM brands to set a lower factory price to attain the same retail price (See discussion in Section IV and Albion, 1983; Nelson and Hilke, 1991).

Key cross-elasticity relationships. At the retail level, interbrand competition takes place between stores and on the counters of the same store. It will be more vigorous in the second environment, since consumer search costs are far lower within than among stores. Interbrand competition within store is therefore the more important determinant of manufacturers' margins.

The chief horizontal determinant of retailer margins in a product class is clearly the intensity of competition among retail stores rather than the extent of interbrand competition within store. Among stores, competition takes place on both an inter and an intrabrand basis. The former involves competition among differentiated items and therefore cannot rise to the same level of intensity as competition on the same brand. To simplify, we therefore omit the effects of interbrand competi-

tion among stores and focus on the roles of interbrand competition within stores and intrabrand competition among stores.[4]

We wish to discover whether, when a retailer raises the price of a brand or discontinues stocking it, consumers are more disposed to switch brands within store or to switch stores within brand. These responses can be represented by two cross-elasticities that capture both the horizontal and vertical determinants of market power and margins. Where PC and Q are the consumer prices and quantities sold of Brands X and Y in stores a and b:

$$E_{b\cdot s}, \text{ the interbrand cross-elasticity} = \frac{dQY_a}{QY_a} \div \frac{dPCX_a}{PCX_a} \quad (1)$$

$$E_{s\cdot b}, \text{ the intrabrand cross-elasticity} = \frac{dQX_b}{QX_b} \div \frac{dPCX_a}{PCX_a} \quad (2)$$

For a little-known manufacturer's brand or in product categories dominated by such brands, the interbrand cross-elasticity is high and the intrabrand cross-elasticity low, so manufacturers are predicted to have slim margins and retailers wide ones. For a leading advertised brand or in categories dominated by such brands, the relative cross-elasticity and margin relationships will be reversed. When the magnitudes of $E_{b\cdot s}$ and $E_{s\cdot b}$ are similar, so too should be margins at the two stages. The more consumers are disposed both to switch brands within store and stores within brand, the lower the expected total margin in the vertical system.

III. Evidence of the Inverse Association

A. THE INVERSE ASSOCIATION IN SPECIFIC INDUSTRIES

Only a few studies seem to have compared margins of manufacturers and retailers in the same industry; all found them to be negatively correlated.

Food products. Wills and Mueller (1989) found a positive association between brand rank, market share, LNA media advertising outlays and retail price in 133 physically homogeneous food classes. In 74 of these classes they also obtained brand wholesale prices. At the expenditure level with the maximum impact on brand price, advertising elevated wholesale prices by 30% more than retail prices, implying a strong negative association between margins at the two levels. The magnitude of this effect would probably have been even larger had the sample been confined to non-commodity classes.[5]

Toys. In a series of articles (1973, 1978a, 1978c, 1991a, 1991b) Steiner presented the evidence for an inverse association in the toy business. Children under 7 form the heart of the toy market. They do not read ads. The industry therefore remained very lightly advertised in the pre-television era, and few toy brands enjoyed a

loyal consumer following. In 1958 rates of return for the 1,327 U.S. producers of toys, games and dolls were well below average for U.S. manufacturing industries – yet the mean industry gross distribution margin was around 49%.

Beginning in the late 1950s toys underwent the same kind of transformation that many other consumer goods industries had experienced 50 to 75 years earlier. Between 1958 and 1970 toy advertising in the major media jumped from under $7 million to over $80 million, and manufacturers' profits rose strongly. Manufacturers' variable margins (net sales minus production wages, purchased materials, sales commissions, inventor and character merchandise royalties and freight out divided by manufacturers' net sales) increased steadily from 25% in 1958 to 33% in 1972. Over the same period, the mean industry gross distribution margin plunged from 49% to 33%, led by the best selling televised toys whose GDMs had fallen to around 20% in the U.S. and Canada by 1972.

Prescription drugs. Rates of return in prescription drug manufacturing have consistently ranked among the highest in U.S. manufacturing industries (Comanor, 1986), especially for the "research intensive" companies. While still on patent, the drugs of these companies are known as "single source" products. I compared RGMs in the 18 highest volume single source drug entities with RGMs in entities where interbrand competition was the most vigorous – namely, in multi-source entities where the market share of generics plus secondary brands exceeded 20%. The 58.4% mean RGM in the multi-source entities was more than double the 27.4% mean RGM in the 18 leading single source entities.

A multiple regression that included 62 single and multi-source drug entities was estimated to control for 2 other variables, the drug's invoice cost and its refill rate, that were predicted to affect RGMs. The regression revealed that the $RGM rose 1¢ for every 1% increase in the generic/secondary brand market share. Thus, as competition becomes more vigorous among pharmaceutical manufacturers, it becomes less vigorous among retail pharmacies. (Information in this section on prescription drugs is from Steiner 1991a.)

Apparel. Apparel accounts for about 6% of U.S. consumption expenditures. Characteristically, apparel categories contain a myriad of little known manufacturers' brands and private labels (although manufacturers' brands have recently gained market share in some categories). Bankruptcies are rampant, average firm size is small, advertising intensity has been light and concentration ratios are low. Net income as a percentage of sales and of net worth is well below average for U.S. manufacturing industries.

Yet at the retail level margins are very high. The "Keystone" pricing convention is prevalent, where the store doubles the factory price to set its retail price and thereby obtains a 50% RGM. In multi-product retail establishments, apparel department RGMs are well above the store-wide average, especially in women's

apparel and accessories, the largest industry segment. (Apparel data in this section is from Steiner, 1978c and 1993.)

B. ADVERTISING INTENSITY AND MARGINS

Manufacturing sector. Recent research has established that within industries there is a significant positive correlation between a manufacturer's market share and its price, price/cost margin and profit (Ravenscraft, 1983, Weiss, 1989, Schmalensee and Willig, 1989, Scherer and Ross, 1990, Greer, 1991). In consumer goods industries in which advertising is important, it is the large market share brands that have the large advertising budgets. Comanor and Wilson (1974), Porter (1976) and others have shown that advertising intensity is positively associated with high rates of return for manufacturers. This result has "proved to be quite robust" (Scherer and Ross, 1990, p. 436) for consumer goods manufacturing industries in the U.S. and in other countries.

The distribution sector. As brand advertising swept across the consumer goods economy, a few economists began to comment on its propensity to drive down the spread between factory and consumer price. Perhaps the earliest was Fogg-Meade (1901). She discovered that dealers were forced to resell Pear's Soap, with one of the largest advertising budgets of its day, at its invoice cost of 10¢ and "So cannot make a cent on the sales" (p. 242). Marshall (1920, pp. 301, 302) observed that when a strongly advertised brand "had won its way, the dealers can be forced to handle it at a low rate of profit," because a refusal to do so "would simply drive away customers" (p. 302). Haring described the retail margin depressing effects of manufacturers' brand advertising in almost the same language, pointing to cigarettes as an example (1935, p. 144).

Patent medicines and proprietary drugs were the earliest class of products to be aggressively advertised. These remedies became subject to intensive retail price cutting in the U.S. and England (Grether, 1935, 1937, Palamountain, 1968). Grether reported that "almost invariably, products with extremely low dealer-margins were well-known and highly advertised" (1935, p. 313). Lydia E. Pinkham's Vegetable Compound,with one of the largest advertising budgets of any brand in America around the turn of the century, still had a razor thin 13.3% gross distribution margin in 1939 (Borden and Marshall, 1959). This compares to the traditional 45% GDM in the industry.

Yamey (1952) relates that English consumers had relied on salesclerks in specialized shops for product information in such goods as proprietary medicines, tobacco and tea. Once brand advertising began to take over this function, new and more efficient types of mass retailers entered these categories. With less skilled, lower-paid salesclerks and other cost advantages they could profitably undersell the specialist retailers. As the new-type retailers gained market share, RGMs declined substantially in these goods.

During the 1890s, bicycles became the first class of durable goods to become intensively advertised. By 1898 retail gross margins had plummeted across America, as department stores, then the new and more efficient form of retailing, cut prices of the leading bicycle brands to capture market share from traditional wheel goods dealers (Steiner, 1978b).

Borden (1942) was one of the first to measure RGMs of advertised brands and their private label counterparts in drug and grocery products. He found, as have later investigators in virtually all lines of merchandise, that nationally advertised brands had materially lower RGMs.

Albion and Farris (1981a) analyzed RGMs in 51 product categories for 488 individual brands sold in supermarkets. To capture the carry-over effects, advertising intensity was represented by a 4 year average of LNA 6-Media Advertising outlays. Using a brand gross margin ratio (BGMR) that measured the extent to which a brand's RGM was above or below the category average, permitted pooling across categories. "The results show, on average, the highly advertised brands sell for gross margins that are 22% lower than the unadvertised brands and 12% lower than the less advertised brands. These differences are statistically significant at the 99% level" (p. 11).

Most other studies – e.g. Harris (1979) for breakfast cereals in the U.S. and Reekie (1979) for a number of products in the U.K. – also find an inverse association between manufacturers' advertising and RGM or GDM. For a summary of studies see Albion (1983, Table 3–3, pp. 58–61).

In sharp contrast to product categories that became intensively advertised – such as soap, bicycles, patent medicines and toys – RGMs in categories that remained lightly advertised have remained high. For example, in women's outerwear (dresses, blouses, waists, coats) from 1958 to 1970 the always meager level of brand advertising expenditures fell in constant dollars and the women's outerwear A/S ratio dropped from 0.3% to 0.2%. Over the same period, RGMs increased from 48% to 51% (Steiner, 1978c).

Most segments of the stationery, housewares, luggage, drapery and bedding industries have remained lightly advertised. General merchandise retailers, the principal outlets, enjoy above store-average RGMs in the departments that resell these products. (Departmental RGMs are published in Chain Store Age General Merchandise Edition, Discount Store News, Discount Merchandiser and by the National Retail Merchants Association and the International Mass Retailing Institute.).

To my knowledge, no intraindustry studies and only 2 interindustry studies have failed to show that advertising intensity and RGMs are negatively correlated. Connor and Weimer (1986) found a significant positive relationship between advertising/sales ratios and RGMs in 30 food categories in supermarkets. However, 10 were commodity products and 2 primarily producer goods. These 12 categories, as would be predicted, had low A/S ratios and low RGMs. Their prominence,

constituting 40% of the sample, prevents generalizing the results to non-commodity food categories.

Weiss, Pascoe and Martin (1983) found a non-significant positive relationship between FTC A/S ratios and RGMs in over 80 FTC consumer lines of business from passenger cars to cane sugar. The authors obtained average store-wide RGMs for 12 types of retail outlets (drugstores, auto dealers, etc.) and then assigned each line of business exclusively to one of them. If supermarkets were the leading outlets for cigarettes and lightbulbs, both products were assigned the 21.1% average supermarket retail RGM, although the RGM of cigarettes runs around 11% and of lightbulbs around 55% in supermarkets (Chain Store Age Supermarkets, various years). Moreover, both lightbulbs and cigarettes are sold in large volume through other types of retailers – often at quite different retail gross margins. Hence, nothing can be concluded about the relationship between advertising intensity and RGMs across consumer goods industries from this study.

Moreover, both the above studies use the A/S ratio rather than dollar advertising expenditures to represent advertising intensity. In my judgment, this constitutes a significant methodological problem.[6]

IV. Dynamics of the Inverse Association

A. RETAILER DOMINATION IN UNADVERTISED INDUSTRIES

The classical preconditions for a vigorously competitive market are a host of buyers and sellers, excellent information and product homogeneity. These conditions are not remotely fulfilled in the second stage (retailer/consumer) market in the typical unadvertised product class.

Consumers don't have strong preferences among the myriad of little-known brands. Since individual brands lack a host of consumer buyers,they are each stocked by a limited number of retail sellers. Without a set of dominant brands that are carried by most dealers, consumers must canvass a large number of outlets to discover the prices asked for different brands and for the same brand. Moreover, there will be relatively little retailer advertising. Dealers have learned that promoting unknown brands, even at steep discounts, seldom produces substantial additional sales. Therefore consumer information is limited by the high cost of store search and the small volume of advertising by manufacturers and retailers alike.

In trade parlance, "the merchandise is blind". Consumers are not simply "blind" about the qualities and prices of different brands. Even more important to the determination of retail margins, they are also blind in the intrabrand market. That is, they do not readily recognize that X_a, X_b and X_c are the same manufacturer's brand X on sale in different stores. Hence, in the intrabrand arena, the product homogeneity criterion for a competitive market is also not fulfilled. Retailers can mark up Brand X almost as though no other stores stocked it without fear of

losing sales to rival dealers. Nor will they attract much business from other stores by cutting the price of Brand X.

In sum, consumers are not disposed to switch stores within brand. The low value of $E_{s \cdot b}$ depresses the elasticity of retailer demand curves E_{DR} and leads to large optimal RGMs, on the order of 50% in keeping with the ubiquitous Keystone pricing convention. This implies that E_{DR} is close to 2.

RGMs are also pushed upward from the cost side. Retailers' non-invoice marginal costs tend to decline with volume and are therefore higher on slow moving items. Longer term, entry may erode retailers' profits by raising their costs, but the substitution of costs for profits is unlikely materially to reduce RGM's or consumer prices.

In categories where brands do not enjoy a solid franchise with consumers, shoppers enter retail stores with a generic demand for a particular class of goods. If the factory price of Brand X is raised slightly, the dealer can replace it with Brands Y and Z that have the same general attributes. The store's sales will hardly be affected. The ease with which consumers will switch brands within store empowers retailers to play off one maker against the next in search of a better price. The high value of $E_{b \cdot s}$ therefore leads to high retailer elasticities of substitution. In turn, this forces the industry's manufacturers to face very elastic demand functions and to have low optimal margins. In this manner, the natural tendency for manufacturers' margins to be thin when there are a large number of horizontal competitors is reinforced by the vertical bargaining power of retailers.

Note that when private labels rather than manufacturers' unadvertised brands dominate a category, manufacturers' margins may be even lower. The producer's name is unknown to consumers and all brand goodwill inures to the retailer's benefit. Yet RGM's in private label domination will be even higher. Stores of the same chain do not compete by price on the chain's own-label brands, so there is no intrabrand price competition. Rival chains do not stock each other's private labels, so there is less within-store search, which reduces the vigor of interbrand competition.

B. MANUFACTURERS' BRAND DOMINATION IN INTENSIVELY ADVERTISED INDUSTRIES

Manufacturer/retailer relationships in categories dominated by a handful of highly advertised brands are the reverse of those in retailer domination. For the category's leading brands, although not for its fringe brands, $E_{s \cdot b}$ is high and $E_{b \cdot s}$ is low, so manufacturers have wide margins and retailers narrow ones.

Consumers have far stronger loyalties for individual brands and are not readily disposed to switch brands within (or among) stores. Hence, retailers have relatively low elasticities of substitution for a leading brand. This allows their makers to increase factory prices without suffering a substantial loss of demand through

diminished retail penetration and dealer support, in contrast to the experience of fringe producers.

In the retailing segment, when a brand becomes well-known through advertising or for any reason, the intrabrand product differentiation that characterizes the competition between stores on a "blind" item is swept away and $E_{s \cdot b}$ rises. Consumers quickly recognize that the Tide or the Barbie dolls on sale at Stores a, b and c are identical. Moreover, a high-market share brand in a large-volume category is important to consumers. A retailer will lose market share by charging more than other retailers for what consumers recognize as the same thing. Once intensified intrabrand competition has forced down the RGM and resale price of the leading brand, interbrand competition within store depresses the margins and resale prices of competing goods.

Leading advertised brands in numerous product classes are periodically offered at temporarily reduced factory prices, a practice that strengthens the inverse association. Supermarkets and other multi-product retailers have long found that store traffic can be increased by sale prices on high profile brands, the everyday prices of which are most familiar to consumers. Once lured to the store, the shopper also purchases other items at regular margins. Chevalier and Curhan (1976) found that supermarkets tended to "over perform" when leading brands went "on deal" – with price cutting, advertising and display that exceeded the terms of the manufacturer's deal. Thereby, leading advertised brands increase their sales and margins at the expense of their smaller, horizontal competitors and of their retailers, alike (Steiner, 1984).[7]

Cost-side influences also reduce RGMs in intensively advertised product classes. For over 100 years, after advertising has been successfully introduced into a previously unadvertised class of goods, retail price cutting on the leading advertised brands has erupted and their RGMs got squeezed. Subsequently, high-cost stores were forced to downplay or to discontinue stocking the most demanded brands and so lost market share to more efficient types of retailers. The consequent reduction in the long-run cost of distribution further depressed RGMs.

C. THE FRINGE MANUFACTURER IN AN ADVERTISED INDUSTRY: THE CASE OF GRANDPA'S PINE TAR SOAP

To understand why the nature of the inverse association facing advertised brands and fringe brands in the same industry is markedly different, we describe and interpret the competitive situation facing Grandpa's Pine Tar Soap. Shortly after World War II this unadvertised fringe brand, first marketed in 1878, was still attempting to compete against the advertised brands of the Big 3 soap makers – Procter & Gamble, Lever Brothers and Colgate-Palmolive. Soap had been one of the earliest product classes to be advertised in the U.S (Presbrey, 1929), and grocers had long since become accustomed to making "practically no markup" on the advertised brand (Klaw, 1969).

On the standard 10¢ size, fringe brands like Grandpa's Tar Soap were offered to the trade at $6\frac{1}{2}$¢ to 7¢ while the advertised brands of the Big 3 soap makers would be offered at $7\frac{3}{4}$¢ to 8¢ and still sell at retail for a dime.[8] The difference in factory prices actually understates the difference in manufacturer's margins between the two classes of producers due to the extensive scale economies in soap making. Clearly, margins at the two stages were inversely related within the industry – high manufacturer and low retailer margins for the advertised brands and the reverse for the fringe brands.

To this author and others at the Grandpa Soap Company, it appeared that the lower RGM's of the advertised brands helped sustain their higher manufacturing margins by forcing fringe makers to sell at a lower factory price to achieve the same retail price as the advertised soap brands. Unfortunately, there seemed little escape from this situation, for price increases were severely constrained by the very elastic nature of our firm's demand schedule. This high elasticity seemed to emanate not so much from consumer behavior as from the conduct of the distributive trade. Consumers did not desert stores that raised the price of Grandpa's Tar Soap by a penny nor flock to those who cut it a penny.

But when the factory price of a 10¢ fringe soap brand was raised relative to others, the competing brands became more profitable to dealers. Since fringe brands lacked a loyal following with consumers, retailers could quickly substitute among them without materially affecting total departmental sales. Therefore an increase in a fringe brand's factory price caused it to lose retail penetration, and the brand became conveniently available to a smaller universe of consumers. Stores that continued to stock the price-increased brand accorded it less dealer support, further diminishing its sales. Likewise, a cut in factory price boosted a fringe brand's retail penetration and dealer support and increased its sales, although once again consumer utility functions for the brand had not changed.

By contrast, Ivory, Palmolive and other leading brands enjoy a quite stable, $90\%^{+}$ level of retail penetration. The manufacturer's unit sales therefore rise or fall only to the extent that a nearly constant size group of consumers changes its purchases in keeping with the dictates of derived demand. Thus, for the leading brand, DC is close to DK – the schedule which defines what consumers will buy when retail penetration and dealer support are constant. For the fringe brand, DC, and ceteris paribus DM, are far more elastic than DK because retail penetration and dealer support vary inversely with PM. Moreover, we have seen that when E_{DR} rises and a brand's RGM declines, E_{DM} falls relative to its former value and to the E_{DMs} of competing higher-RGM brands. In sum, the three dual stage effects virtually assure that with equal consumer demand elasticities (as defined by DK) a leading brand will have a lower E_{DC} and a still lower E_{DM} than a fringe brand.

D. SUCCESSFUL ADVERTISING OF A PREVIOUSLY UNADVERTISED BRAND, GIRDER AND PANEL BUILDING SETS

We now illustrate, through another real-world example, how the introduction of successful advertising of a previously unadvertised brand reverses the relative values $E_{s\cdot b}$ and $E_{b\cdot s}$ raising the manufacturer's margin and lowering the retailers'. The analysis also demonstrates the crucial role of the three dual stage effects – retail penetration, dealer support and RGM in this process.

In keeping with industry norms, Kenner's Girder and Panel Building Sets had a GDM (Kenner sold to both wholesalers and retailers) of around 50% in 1956 and 1957. In 1958 and 1959 the toy maker used the new medium of television advertising in a limited number of local markets. The sales and GDMs of the building sets changed little in the non-TV markets. But in the TV test markets the advertising created a groundswell of consumer demand. Through the mechanisms previously described, the elasticities of retailer demand schedules (E_{DR}) rose, pervasive retail price cutting erupted and the building sets' GDM fell to an estimated 33%.[9] Despite the plunge in retailers' margins, the growing popularity of the Girder and Panel Sets impelled many new dealers to begin carrying them and encouraged existing outlets to advertise them far more aggressively in the local papers.

In 1960 the Girder and Panel Sets became nationally advertised and their GDM's continued to decline for several years as more discount stores began to stock and feature them. Meanwhile, as output grew, Kenner discovered there were sizable scale economies in tooling, manufacturing and assembly of the Building Sets and in the purchase of television media. Accordingly, as its dealers' margins plunged, Kenner's own margins and profits from the Girder and Panel Line grew materially at the pre-TV level of factory prices, which were not raised for several years.

Figures 1 and 2 depict these events. To simplify the exposition, the manufacturer's marginal and average variable costs and retailers' non-invoice marginal costs have been held constant and the manufacturer is assumed to sell only to retailers.

Figure 1A illustrates the pre-advertising equilibrium. At quantity Q, the factory price PM is B on DM. The 50% RGM (BA/QA) produces a retail price of A along DC. At Q the retailers' $ non-invoice marginal cost per unit is BY, so the retailers' true margin is about 40% (YA/QA). The manufacturer's margin is about 20% (DB/QB).

Figure 1B portrays the situation shortly after the successful TV campaign. The intensified intrabrand competition has pushed the retail price down to E along the brand's new post-advertising consumer demand curve DC1. The quantity sold Q1 has more than tripled. The old factory price had been maintained, but Figure 1B illustrates that it was no longer optimal. Therefore factory price is raised from F to G on the extended manufacturer's demand curve DM1. The new equilibrium post-advertising retail price H is well below its pre-advertising counterpart, and the new equilibrium quantity Q2 is more than twice Q.

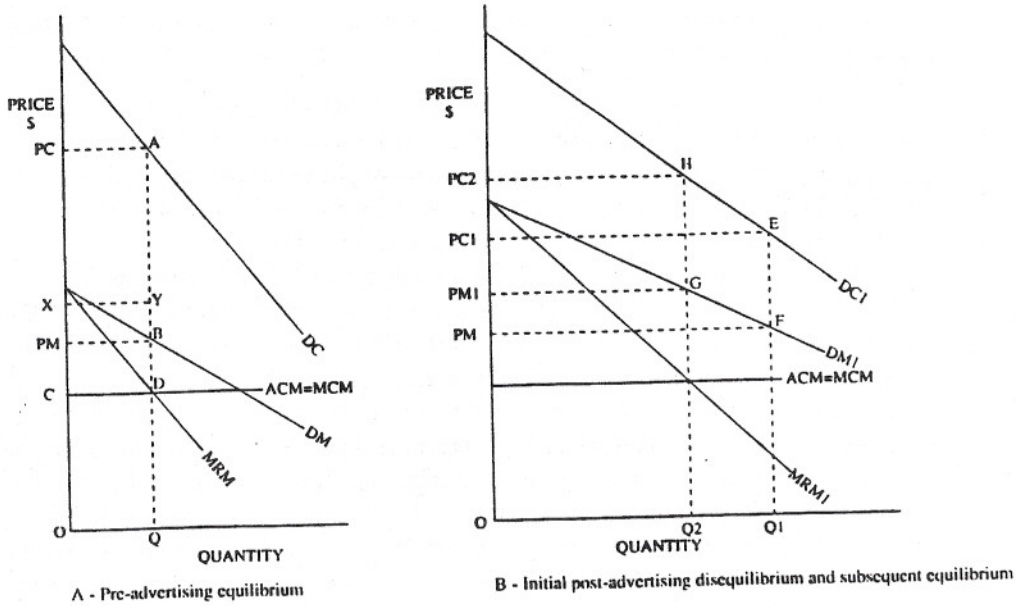


Fig. 1. The effects of successful brand advertising.

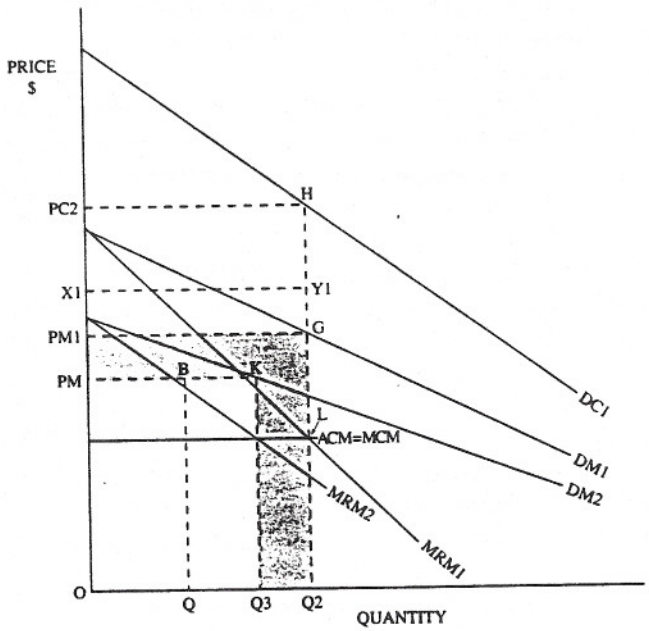


Fig. 2 Vertical and horizontal sources of manufacturer's sales incre

By comparing Figures 1A and 2 we can examine the results for the manufacturer and his retailers. The retailers' true margin has plunged from about 40% prior to advertising to about 21% in the post-advertising equilibrium (Y1H/Q2H in Figure 2). Although, by assumption, the manufacturer's marginal costs have remained constant, the firm's margin has risen from about 27% in Figure 1a to about 40% in Figure 2 (LG/Q2G).

From Figure 2 we can distinguish between the two sources of the manufacturer's sales. The area (O, PM1, G, Q2) is the manufacturer's sales volume in the post-advertising equilibrium and (O, PM, D, Q) in the pre-advertising equilibrium. Had the pre-advertising 50% RGM continued into the post-advertising period, the manufacturer would have faced demand curve DM2 and his sales gain would have been limited to the area (Q, B, K, Q3). This represents the portion of the total sales increase that came from horizontal competitors, such as Erector Sets, and from the purchases of consumers new to the category. But the post-advertising RGM fell to around 33% (GH/Q2H), so the manufacturer now received two-thirds rather than one-half of each retail sales dollar. The larger shaded area (PM, PM1, G, Q2, Q3, K) represents the revenues the manufacturer took from his retailers through his enhanced vertical position.

V. Exceptions to the Inverse Association

We have seen that throughout much of the consumer goods economy the magnitudes of $E_{s \cdot b}$ and $E_{b \cdot s}$ are markedly different, causing margins in manufacturing and retailing to be inversely related. Indeed, an intriguing article by Bradburd suggests that the inverse association between margins at successive stages also characterizes producer goods industries.[10]

However, the inverse association does not extend to product classes where the propensities to switch brands within store and stores within brand are roughly equal. We identify and briefly describe some of these situations below.

A. EQUAL MANUFACTURER/RETAILER POWER IN NON-COMMODITY CLASSES

Mutual dependence. It was Bowman's (1952) insight that RPM arose out of "mutual dependence" between an "insecure partial monopolist" manufacturer and his insecure partial monopolist retailers. When price cutting erupts on a moderately popular brand, its producer is concerned that numerous dealers will succeed in switching consumers to a higher margin substitute brand, resulting in a loss of retail penetration and dealer support. Concurrently, dealers feel that if they do not meet the lower prices, a good many consumers will switch their patronage to a price-cutting store. The mutual insecurity prompts the manufacturer and his retailers to adopt a minimum resale price in the belief that the profits of both parties will be enhanced. If they are right, margins at both stages will rise.

The mixed regimen. In this structure, margins at the two stages are again positively correlated. However, since the parties are both powerful rather than insecure, the margins are moderate (Steiner, 1978a). The mixed regimen occurs in categories where a handful of leading national brands are opposed by strong private label brands of large chain retailers. Within stores there is a relatively high cross-elasticity of demand between the two kinds of brands. This keeps the lid on the factory prices of leading advertised brands. Vigorous intrabrand competition on these items produces low RGMs and results in a moderate level of national brand retail prices. These prices, together with the "reputation premium" (Braithwaite, 1929; Borden, 1942) that famous advertised brands command with consumers, forces competing private label goods to be retailed at considerably lower prices. The foregoing dynamics may require a 20–40% private label market share. Yet the process often commences at a far lower share, as recently illustrated in the cigarette and disposable diaper industries.[11]

The replacement tire business has historically exemplified these characteristics. In the 1970s and before the 4-firm concentration ratio was over 70%, yet tire producers' rates of return were below average for U.S. manufacturing industries. The private label market share was in the 30–40% range and the average industry gross distribution margin a modest 30–33%.[12]

Bilateral monopoly. Although this structure is rare in real world consumer goods markets, we note it for completeness and because it has received much analytical attention. With one manufacturer's brand and one retail store there is neither inter nor intrabrand competition. Since both $E_{b \cdot s}$ and $E_{s \cdot b}$ are zero, we would predict high and roughly equal margins in manufacturing and retailing. Interestingly this same result is generated through a different analytical approach that examines the effects of dual marginalization.

B. TRUE COMMODITY CLASSES

In categories where consumers recognize that the goods are fungible across stores and across brands both $E_{s \cdot b}$ and $E_{b \cdot s}$ will have very high values.

RGM's will be low because grading of commodity goods performs the same role that manufacturer's advertising does in the intrabrand market of non-commodity goods. It enables consumers readily to recognize that, for example, the grade A large eggs in different stores are identical. The inability of advertising and other branding efforts to differentiate brands creates a highly competitive manufacturing environment, so producers' margins will also be thin unless there are non-marketing related entry barriers.

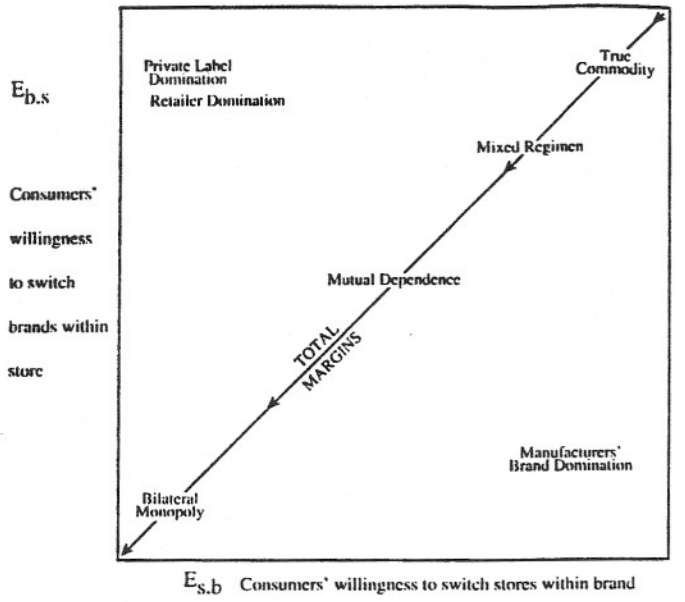


Fig. 3. Industry structures located by magnitudes of $E_{b.s}$ and $E_{s.b}$.

C. LOCATING INDUSTRY STRUCTURES BY REFERENCE TO $E_{b\cdot s}$ AND $E_{s\cdot b}$

Figure 3 evolved from a suggestion by Michael Lynch that I should attempt to locate industry structures by the relative values of $E_{b\cdot s}$ and $E_{s\cdot b}$. The locations shown are intended to be illustrative rather than precise.

Along the NE/SW diagonal total margins increase to the southwest. On any perpendicular (NW/SE) to the diagonal, total margins are equal, with the retailers' share of the total increasing to the northwest and the manufacturers' share to the southeast.[13] At the northeast corner is situated the true commodity industry in which both stages approach perfect competition and total margins are the lowest. At the southwest corner with the highest total margin is the bilateral monopoly structure in which $E_{b\cdot s}$ and $E_{s\cdot b}$ are both zero.

VI. Implications

We now identify 4 of the numerous areas which require rethinking in light of the inverse association between margins at the two levels.

Explaining manufacturers' margins and profits. Both individual brand RGMs and mean industry RGMs vary widely. Only recently have a few economists recognized that differential RGMs can create entry and mobility barriers (Albion, 1983;

Nelson and Hilke, 1991). The manufacturer of a high RGM brand must accept a lower factory price than competing brands with thinner RGMs if his brand is to sell at the same price to consumers.

I therefore propose that RGM be included as an independent variable in regressions seeking to explain manufacturers' margins, prices and profits. This would capture the reality that a manufacturer's market power is a joint function of its competitive standing with its retailers and with rival producers. The addition of RGM is predicted to raise the regression's R^2, and the coefficient on RGM is expected to be sizable, negative and significant.

Thus, correcting for other variables, where RGM is below average, both within and among industries, manufacturers' margins will be higher. As the analysis below will indicate, this is not a prediction about aggregate industry profits or rents but about their distribution among and between manufacturers and retailers.

Industry welfare analysis. In keeping with the single stage paradigm, welfare calculations for consumer goods industries reflect values at the manufacturing stage. What transpires in the downstream distribution markets is ignored, although in non-food industries somewhat more value is added in the downstream wholesale/retail markets than by final consumer goods manufacturers (Steiner, 1991c, note 29, p. 49). Since costs and margins in distribution are not measured, there is no concept of distributors surplus. Total industry surplus is simply the sum of manufacturers and consumers surplus.

Changes in productivity and in the vigor of competition at one stage can be amplified, offset or outweighed by corresponding changes at the other. These dynamics can't be captured unless both stages of an industry are examined.[14] Obviously, single stage models yield their most misleading results in industries where there is a strong inverse association between margins in manufacturing and retailing.

Welfare effects of manufacturers' brand advertising. In a single stage world rising factory prices, associated with intensified advertising, imply rising retail prices. In a dual stage environment, if $RGMs fall by more than the rise in factory prices as advertising is increased, retail prices will fall. I have shown that this can occur with constant costs in manufacturing, although the effect is more pronounced when returns to scale are present. Hence, some amending of the conclusions of Comanor and Wilson (1974) and Porter (1976) seems indicated. In both studies output is valued in manufacturers' selling prices. In some of the industries where consumers are judged to be worse off due to intensive advertising,they may be better off.

Moreover, the conclusion in Comanor and Wilson that "relative advertising expenditures appear to be more important than relative prices in allocating sales among industries" (p. 239) may also require revision. Due to the negative correlation between advertising intensity and RGMs, using manufacturers' prices and

demand schedules as surrogates for consumer prices and demand schedules leads to biased results.

First, the influence of price on the allocation of consumer demand among industries is overstated. Actual consumer prices in heavily advertised industries are lower relative to those in other industries than they appear to be in this methodology. Next, the influence of manufacturers' advertising on the interindustry allocation of demand among consumers is exaggerated. As shown in Figures 1 and 2 and in the accompanying discussion, when advertising drives down a brand's RGM, DM shifts out relative to DC. A good part of the increase in the manufacturer's sales comes from retailers and not from consumers. Likewise, the RGM effect causes manufacturer level market demand schedules in strongly advertised industries to lie closer to consumer level market demand schedules than in less intensively advertised industries. Therefore, advertising has a weaker sales allocating influence among the consumers of different classes of products than among the manufacturers of these products.

Porter's work is significant for its attempt to incorporate the role of the retailer. Still, manufacturers' advertising is seen as welfare diminishing on the grounds that it "leads to allocative inefficiency and elevation of manufacturers' rate of return" (p. 236). Porter concludes that advertising's deleterious effects are greater in convenience goods, in part because the coefficient on advertising in his regressions explaining profits is far stronger than in non-convenience goods. If, as I suspect, convenience goods have lower RGMs, ceteris paribus, this conclusion would also require amending.

The paradox of the advertising/response function. In a well-known summary of the evidence on the advertising/response function, Simon and Arndt (1980) reported that studies "linking physical measures of sales impact to physical amounts of advertising consistently indicate diminishing returns to advertising . . .". With two exceptions, studies relating sales in dollars to dollars of advertising also showed diminishing returns to advertising. Despite this apparently strong evidence, the authors discovered that the great majority of advertising practitioners believed that the advertising/response function was S-Shaped, with a substantial period of increasing returns before an inflection point is reached and diminishing returns set in. As vice president of advertising for Kenner Products, this author shared that belief.

Once again, analyzing the evidence in a dual stage framework unlocks the paradox. Diminishing returns from the start may characterize the function when the response is measured in units or valued in the prices consumers pay. But if concurrently a brand's factory price is rising relative to its retail price, there is often a period of increasing returns when output is valued in the prices manufacturers receive. These dynamics also explain why the optimal advertising budget for a manufacturer in a dual stage world is higher than in a single stage environment

where retailers' margins are not depressed by increases in advertising intensity (Steiner, 1981, 1987).

The competitive process. The inverse association between margins at successive stages has important implications for our understanding of the dynamics of competition. In antitrust law and in economics competition is a process that takes place solely among firms at the same level.

But we have seen that when intensified competition depresses margins at one stage, margins at the other stage are likely to rise. As a brand's popularity grows, the manufacturer's revenues are augmented from two sources – a fall in the brand's RGM and a rise in its market share. A "vertical" gross margin dollar taken from retailers is just as good as a "horizontal" gross margin dollar taken from rival manufacturers. Thus, competition has both a vertical and a horizontal dimension.

Notes

* Economic Consultant, Washington D.C. I gratefully acknowledge the valuable comments on an earlier draft by Ralph Bradburd, William Comanor, Douglas Greer, Michael Lynch and Thomas Overstreet, Jr.

[1] In many consumer goods industries, manufacturers sell their output to distributors, as well as directly to retailers. This does not require a "triple stage" model, since the wholesale/retail markets can ordinarily be combined into one without introducing material distortion.

[2] Retailers are hard pressed to calculate profit per brand, since they cannot easily allocate overhead against the thousands of individual items they stock. However, a brand's dollar RGM per square foot of selling space or per dollar invested in inventory are good estimates of its contribution to store overhead and profit.

[3] Elsewhere (Steiner, 1991a) I've used the term Vertical Market Share to represent the portion of the retail price going to the manufacturer (1-RGM) and to retailers (RGM).

[4] Interbrand competition among stores is the least direct of the three forms of brand competition. Therefore both the mean magnitude of the cross-elasticity that defines this form of competition and its range are lower than those of $E_{s\cdot b}$ and $E_{b\cdot s}$. In Section 4 we will show that the values of $E_{b\cdot s}$ and $E_{s\cdot b}$ vary widely, depending principally on how well the brands are known.

Virtually all interbrand competition is forced to take place among stores rather than within them where exclusive dealing arrangements are prevalent. In light bulbs, a moderate priced category, these arrangements had caused interbrand cross-elasticities to be very low. Subsequently, in some markets a major grocery chain added a second (less expensive) light bulb line, thereby introducing interbrand competition within its stores. Demand elasticities were quickly invigorated and retail light bulb prices fell sharply in these markets (Steiner, 1985).

[5] Margins at the two stages are predicted to be positively associated in true commodity industries (see Section V). The authors' intention was to include in their sample only categories in which the goods were thought to be more or less physically homogeneous. However, in numerous cases (e.g., corn flakes) consumers definitely do not regard the rival brands as fungible, so they are properly classified as non-commodity categories.

[6] Although the A/S ratio has been widely used as a surrogate for advertising intensity, it is difficult to fathom why sales belong in this definition. If in the aggregate firms in 2 industries spend the same amount to purchase identical media schedules, the industries would seem to have the same advertising intensity. Yet using A/S produces the conclusion that if Industry A's sales are twice Industry B's, it is only half as intensively advertised. Of course spending the same budgets would be the profit maximizing strategy for firms in both industries if the variable margins in Industry B were twice those of A.

It is particularly inappropriate to use the A/S ratio when there are large interindustry differences in size and in average product price, as in Weiss, Pascoe and Martin (1983). To illustrate, in the 1974 FTC Line of Business Report more media dollars were spent advertising passenger cars than in any other industry, save one. However, passenger cars have only a 0.7% A/S ratio and therefore fell to 113th place when industries were ranked by the A/S ratio.

Advertising academics find that alternate measures of advertising intensity are not closely correlated and recommend using several of them. But if only one is selected, the A/S ratio is not the parameter of choice (Lancaster, Batra and Miracle, 1982; Hovland and Lancaster, 1985). Finally, the FTC's media advertising data which were used in both studies, are noisy and subject to a substantial large-firm bias. LNA is a far superior source. It collects quarterly outlays in the 6 major media for all brands and sums them by very discrete product classes.

[7] In 1982 70% of factory shipments were "on deal" in the typical health and beauty aid category (Quelch, 1982). There is evidence that this strategy may no longer be as effective. Procter & Gamble, the world's largest advertiser, recently announced it was drastically reducing or eliminating deals on its brands and instead was lowering their everyday factory prices. The extent to which this new policy will be adopted by other leading convenience good manufacturers is not yet clear (*Wall Street Journal*, 1992, P. 1, A-11).

[8] In 1949 P&G offered its best selling medium size of Ivory Soap to the trade at $7.75 in 100 case lots and $8.00 in smaller quantities. Although it was referred to as the "10¢ size", aggressive grocers often offered it on special for less. By contrast, few customers ordered Grandpa's Pine Tar Soap in 100 case lots or cut its retail price below 10¢.

[9] Parenthetically, Kenner's French licensee had a virtually identical experience on its Spirograph toy when the prohibitions against toy commercials on TV were suddenly removed just in time for the 1975 Christmas season. The French licensee quickly prepared a commercial and bought television time for Kenner's Spirograph, a staple that had been in its line for some years. During the TV campaign Spirograph's retail price plunged by around 40% and its sales volume increased an amazing 7-fold over the levels of prior Christmas seasons.

[10] Using input-output tables and manufacturers' price/cost margins from the Census of Manufactures, Bradburd (1982) sought primarily to probe the association between cost-importance and price/cost margins in producer's goods. However his regressions also included a variable for the downstream industry's price/cost margins. Based on derived demand theory, Bradburd had expected a positive correlation between the margins of the firms in the selling and buying industries. To his surprise, there was a sizable and statistically significant negative correlation. Bradburd observed "One possible explanation is that there is a fixed amount of monopoly rent that can be extracted in the entire vertical structure of production for any good" (p. 409).

[11] For years the private label unit market share in cigarettes was close to zero. Suddenly, by 1984 it had spurted to around 4%. Alarmed at this evidence of a rising interbrand cross-elasticity of demand, the major tobacco manufacturers themselves became private label suppliers and also introduced lower priced branded lines. By early April 1993 the combined unit market share of all types of "discount cigarettes" had reached 36% (Wall Street Journal, 1993). At this juncture Philip Morris took the revolutionary step of cutting the price of Marlboro, the industry's dominant brand, by some 20% in hopes of rebuilding Marlboro's eroding market share. Shortly thereafter, rival R. J. Reynolds, the second ranking cigarette maker, cut prices on its Winston and Camel brands.

With Pampers, the pioneer brand in the category and Luvs, Procter & Gamble had a 50.5% share of the disposable diaper business in 1988. Its share plunged to 41.9% in 1992, while the private label share rose from 14.1% to 20.4% (and Kimberly Clark's share increased from 33.7% to 36.9%). In April 1993 Procter & Gamble announced a 5% price cut on Pampers, its third within a year, and a 16% price cut on Luvs. (Advertising Age, 1993).

[12] For background on the tire business, see FTC Staff Report (1966), Scherer and Ross (1990, p. 281), Cook and Schutte (1967, pp. 169–182), Comanor and Wilson (1974, pp. 135, 188–199). Also see various issues of Modern Tire Dealer, for example January 1979, pp. 32–33, for national brand/private label market shares.

[13] As drawn, total margins are roughly equal in retailer and private label domination, mutual dependence and manufacturers' brand domination. Intuitively, I believe that under most cost and demand conditions total margins will be higher in mutual dependence but cannot at this juncture prove that conjecture.

[14] A dual stage welfare diagram is provided in Steiner (1985). Vertical productivity indexes for the toy and women's outerwear industries were constructed in Steiner, 1978c.

References

Advertising Age (April 19, 1993), Vol. 64 No. 17, 43, Chicago: Crain Communications.

Albion, M. (1983) *Advertising's Hidden Effects*, Boston: Auburn House.

Albion, M. and Farris, P. (1981a) *The Effect of Manufacturer's Advertising on Retail Pricing*, Report 81–105. Cambridge: Marketing Science Institute.

Albion, M. and Farris, P. (1981b) *The Advertising Controversy*. Boston: Auburn House.

Borden, N. and Marshall, M. (1959) *Advertising Management, Text and Cases*, Boston: Harvard Business School.

Borden, N. (1942) *The Economic Effects of Advertising*, Chicago: Richard D. Irwin.

Bowman, W. (1952) 'Resale price maintenance – A monopoly problem', *Journal of Business* **25**, 141–155.

Bradburd, R. (1982) 'Price/cost margins in producer goods industries and 'The importance of being unimportant', *The Review of Economics and Statistics* **LXIV**, 405–412.

Braithwaite, D. (1928) 'The economic effects of advertisement', *Economic Journal* **XXXVIII** (149), 16–37.

Chain Store Age/Supermarkets, (various years) 'Annual sales manual', New York, NY.

Chain Store Age, General Merchandise Edition, New York, NY.

Chevalier, M. and Curhan, R. (1976) 'Retail promotions as a function of trade promotions: A descriptive analysis', *Sloan Management Review* **18** (1), 19–32.

Comanor, W. (1986) 'The political economy of the pharmaceutical industry', *Journal of Economic Literature* **24**, 1178–1217.

Comanor, W. and Wilson, T. (1974) *Advertising and Market Power*, Cambridge: Harvard University Press.

Connor, J. and Weimer, S. (1986) 'The intensity of advertising and other selling expenses in food and tobacco manufacturing: Measurement, determinants, and impacts', *Agribusiness*, **2** (3), 293–319.

Cook, V. and Schutte, T. (1967) *Brand Policy Determination, Marketing Science Institute*, Boston: Allyn and Bacon.

Dardis, R. and Skow, L. (1969) 'Variations for soft goods in discount and department stores', *Journal of Marketing* **33**, 45–50.

Discount Merchandiser, (various years) 'The true look of the discount industry', New York, NY.

Discount Store News, New York, NY.

Farris, P. (1993), personal communication.

Federal Trade Commission, Bureau of Economics (1974, 1975, 1976) *Annual Line of Business Reports*, Washington, D.C.

Federal Trade Commission (1966) *Staff Report on the Manufacture and Distribution of Automotive Tires*, Washington, D.C.

Fogg-Meade, E. (1901) 'The place of advertising in modern business', *Journal of Political Economy* **9**, 218–242.

Greer, D. (1991) *Industrial Organization and the Public Policy*, 3rd Edition, New York: MacMillan & Co.

Grether, E. (1935) *Resale Price Maintenance in Great Britain*, Berkeley: University of California Press.

Grether, E. (1937) *Price Control under Fair Trade Legislation*, New York: Oxford University Press.

Haring, A. (1935) *Retail Price Cutting and Its Control by Manufacturers*, New York: Ronald Press.

Harris, B. (1979) 'Shared monopoly and the cereal industry: An empirical investigation of the effects of the FTC's antitrust proposal', MSU Business Study, Michigan State University, E. Lansing, MI

Hovland, R. and Lancaster, K., (1985) 'Do concentrated markets spend more on advertising? Evidence on nondurable consumer goods industries', *Journal of Macromarketing* **5**, (2), 17–31.

International Mass Retailing Institute, Washington, D.C.

Klaw, S. (1959) 'Lever's artful dodging', *Fortune* **59** (June), 124–127.

Lancaster, K., Batra, R. and Miracle, G. (1982) 'How the level, intensity and distribution of advertising affect market concentration', *Proceedings of the 1982 Conference of the American Academy of*

Advertising, Knoxville: Alan D. Fletcher, College of Communications, University of Tennessee, pp. 53–59.

Leading National Advertisers, Inc. (LNA), 'Class/Brand YTD', New York, NY.

Lynch, M. (1986) *The 'Steiner Effect', A Prediction from a Monopolistically Competitive Model Inconsistent with any Combination of Pure Monopoly or Competition*, FTC Bureau of Economics Working Paper No. 141.

Marshall, A. (1920) *Industry and Trade*, London: MacMillan & Co., Ltd.

Modern Tire Dealer, Vol. 60, No. 7, Akron.

National Retail Merchants Association (various years), 'Merchandising and operating results of department stores', New York, NY

Nelson, P. and Hilke, J. (1991) 'Retail featuring as a strategic entry or mobility barrier in manufacturing', *International Journal of Industrial Organization* **9**, 533–544.

Palamountain, J. (1968) *The Politics of Distribution*, New York: Greenwood Press.

Porter, M. (1976) *Interbrand Choice, Strategy and Bilateral Market Power*, Cambridge: Harvard University Press.

Presbrey, F. (1929) *The History and Development of Advertising*, Garden City: Doubleday, Doran & Co.

Preston, L. (1963) *Profits, Competition and Rules of Thumb in Retail Food Pricing*, Berkeley: Institute of Business and Economic Research, University of California.

Procter and Gamble Archives, Edward Rider Archivist, Cincinnati, OH.

Quelch, J. (1982) *Trade Promotion by Grocery Product Manufacturers: A Managerial Perspective*, Cambridge: Marketing Science Institute Report, pp. 82–106.

Ravenscraft, D. (1983) 'Structure in profit relationships at the line of business and industry level', *Review of Economics and Statistics* **65**, 22–31.

Reekie, W. (1979) *Advertising and Price*, London: The Advertising Association

Scherer, F. and Ross, R. (1990) *Industrial Market Structure and Economic Performance*, Boston: Houghton Mifflin.

Schmalensee, R. and Willig, R. (1989) *Handbook of Industrial Organization*, Amsterdam, North Holland.

Simon, J. and Arndt, J. (1980) 'The shape of the advertising response function', *Journal of Advertising Research* **20**, 11–28.

Steiner, R. (1973) 'Does advertising lower consumer prices?', *Journal of Marketing* **37**, 19–26.

Steiner, R. (1978a) 'A dual stage approach to the effects of brand advertising on competition and price', in John F. Cady (Ed.), *Marketing and the Public Interest*, Cambridge: Marketing Science Institute, pp. 127–150.

Steiner, R. (1978b) 'Learning from the past – Brand advertising and the great bicycle craze of the 1890s', in Permut, S. (Ed.), *Advances in Advertising Research and Management*, The American Academy of Advertising, pp. 35–40.

Steiner, R. (1978c) 'Marketing Productivity in Consumer Goods Industries', *Journal of Marketing* **42**, 60–69.

Steiner, R. (1981) 'Judging the welfare performance of manufacturers' advertising', *Journal of Advertising* **10**, 3–13.

Steiner, R. (1984) 'Basic relationships in consumer goods industries', in *Research in Marketing* **7**, 165–208.

Steiner, R. (1985) 'The nature of vertical restraints', *Antitrust Bulletin* **30**, 81–135.

Steiner, R. (1987) 'Point of view: The paradox of increasing returns to advertising', *Journal of Advertising Research* **27**, 45–63.

Steiner, R. (1991a) 'Intrabrand competition – Stepchild of antitrust', *The Antitrust Bulletin* **XXXVI** (1), 155–200.

Steiner, R. (1991b) 'Manufacturers' Promotional Allowances, Free Riders and Vertical Restraints', *The Antitrust Bulletin*, **XXXVI** (2), 383–411.

Steiner, R.(1991c)'Sylvania economics – A critique', *Antitrust Law Journal* **60** (1), 41–59.

Steiner, R. (1993) 'Jeans: Vertical restraints and efficiency', in Duetsch, L. (Ed.), *Industry Studies*, Englewood Cliffs: Prentice Hall.

Wall Street Journal (1992) August 11, A-1, A-11.

Wall Street Journal (1993) April 5, A-1, A-6.
Weiss, L. Editor (1989) *Concentration and Price*, Cambridge: The MIT Press.
Weiss, L., Pascoe, G. and Martin S. (1983) 'The Size of Selling Costs', *The Review of Economics and Statistics* **65**, 668–672.
Wills, R. and Mueller, W. (1989) 'Brand pricing and advertising', *Southern Economic Journal* **56** (2), 383–395.
Yamey, B. (1952) 'The origins of resale price maintenance: A study of three branches of retail trade', *The Economic Journal*, September, pp. 523–545.

Why economists are wrong to neglect retailing and how Steiner's theory provides an explanation of important regularities

BY MICHAEL P. LYNCH*

AUTHOR'S NOTE: *I trust my indebtedness to Robert Steiner is evident in the text. I would also like to thank, without implicating, Russell Porter and Morris Morkre for discussions and criticism of the conclusions expressed in this paper, Bert Foer and Greg Gundlach for advice and editing and Bill Curran, Editor-in-Chief of* **The Antitrust Bulletin**, *for granting permission to post this paper on AIA's website.*

[This paper is a slightly expanded version of the paper published in **The Antitrust Bulletin**, Vol. XLIX, No. 4, Winter, 2004]

* Independent Economist

For more than three decades Robert Steiner has been arguing that economists have neglected both retailers and the competition between retailers and their manufacturer suppliers.[1] The most extreme form of neglect is to act as though retailers do not even exist; manufacturers or importers are assumed to sell directly to consumers. Steiner calls the latter the "single stage" thinking in opposition to his "dual stage" theory in which both retailers and manufacturers play an important role in determining consumer prices and the quantities and types of goods sold. Clearly one consequence of this "single stage" view is that manufacturers have no influence on the retail margins at which their products are sold to consumers, nor do retailers have any influence on the factory price they pay to their suppliers. A further consequence: any change in prices at the manufacturing level is assumed to be passed through to the consumer dollar for dollar. Another way of putting this is that retailers who on average mark-up the products they buy from manufacturers by more than 40% are assumed by economists to apply a zero percent mark-up to any price changes. The retailer's dollar gross margin per unit ($rgm) is assumed to be unaffected by changes at the manufacturing level. In contrast, Steiner, based in part on his own business experience and in part on his own and others study of certain empirical regularities, asserts that manufacturers, through brand advertising, can directly affect retailers by forcing them to compete harder and to lower their margins on leading national brands. His early experience at Kenner Products provided a dramatic example. In 1958, Kenner was one of earliest toy companies to use the new medium of television to advertise their toys. The results exceeded even Steiner's youthful expectations. Not only did unit sales of the Girder & Panel sets increase many fold, but to his surprise retailers and wholesalers reduced their margins on the product so that the gross distribution margin fell from 50% to 33%. Since Kenner maintained its original factory price, the result was that retail prices fell by about 25%. After two years, Kenner raised their factory price, but the post-TV advertising retail price of the Girder & Panel sets remained well below their pre-TV levels.[2] This was not an isolated, anomalous, case. Steiner has shown that the advent of television advertising in the toy industry in the 1950s led to a substantial decline in the retail price of toys, despite the fact that manufacturers raised their prices! To an economist steeped in the model of perfect competition, this claim, that a rise in factory prices is associated with a fall in retail prices, appears fantastic, yet it is clear that it happened and not just in the toy industry. But Steiner also argues that retailers do not only dance to the tune played by manufacturer advertisers. Steiner further argues that there are impressive empirical

---

[1] See his early papers, *Does Advertising Lower Consumer Prices?* Journal of Marketing, vol. 37 #4 (1973). Also Reprint # 37, American Enterprise Institute (1976), *The Prejudice Against Marketing,* Journal of Marketing, vol.40 #3 (1976), *A Dual-Stage Approach to the Effects of Brand Advertising on Competition and Price*, in John Cady, Editor, *Marketing and the Public Interest*, Proceedings of 1977 Symposium Conducted by Marketing Science Institute in Honor of E.T. Grether. MSI Rejport No. 78-105.

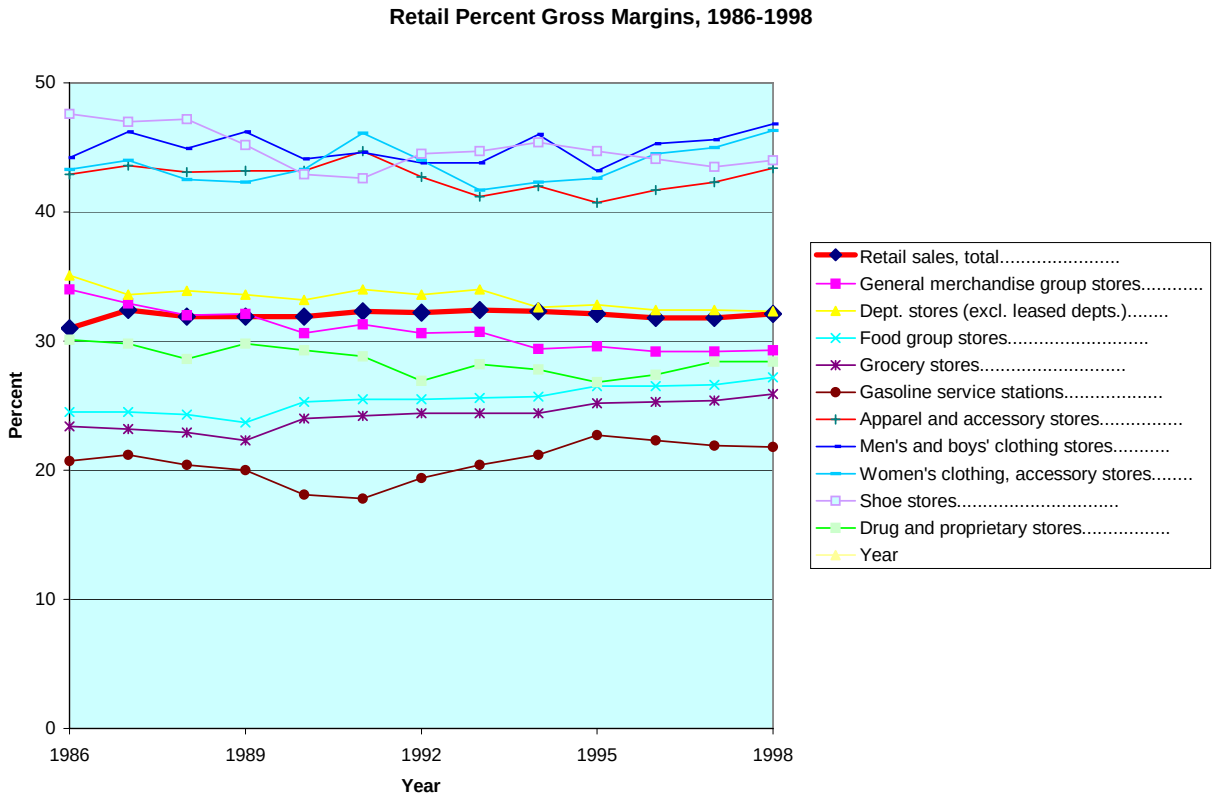
[2] In my 1986 attempt to model Steiner's theory, I referred to the inverse association of margins at the retailing and manufacturing levels as the "Steiner Effect." Here I will use the term "strong Steiner Effect" to refer to situations where the fall in retail margins more than offsets any increase in factory price.

regularities in retailing that have important consequences for how prices at the factory level are, or are not, reflected at the retail level that have been virtually ignored by economists. Further, Steiner argues that economists ignore vertically competitive interactions between retailers and manufacturers. Manufacturers, through successful brand advertising can force retailers to become more competitive and to reduce their margins on popular brands. Retailers are not passive. Through their control of shelf space allocation, display position, through promotion of their own store brands, retailers can pressure manufacturers to lower factory prices. If Steiner is correct, then his views have important implications for measuring the welfare effects of market power in manufacturing and retailing, for estimating passthroughs of manufacturing level price changes to the retail level, and for antitrust analysis including merger analysis and vertical restraints.

Neglect of Retailing?

There is no doubting the economic significance of retailing. As shown in Figure 1, the percentage retail gross margins (%rgms) for all retailers combined account for almost one third of the price of every product bought by consumers. For later reference, it is also worth noting how remarkably stable this overall %rgm is. According to the Annual Census of Retail Trade, the %rgm for all retail stores averaged 32% over the period 1986 – 1998. It varied little: a low of 31%, a high 32.4% with a coefficient of variation of only 1.2%. Specific retail types showed more variation with department stores and the apparel categories having the lowest variations at 2.5%, and gasoline stations the highest at 7.3%. Gasoline stations also had the lowest average %rgm at 20.6%.

Figure 1


Retail Percent Gross Margins, 1986-1998
Percent
Year
0
10
20
30
40
50
1986
1989
1992
1995
1998
Retail sales, total.....
General merchandise group stores.....
Dept. stores (excl. leased depts.).....
Food group stores.....
Grocery stores.....
Gasoline service stations.....
Apparel and accessory stores.....
Men's and boys' clothing stores.....
Women's clothing, accessory stores.....
Shoe stores.....
Drug and proprietary stores.....
Year

Nearly one of out of every six employees on a non-farm payroll works in wholesale or retail trade. Currently, the world's largest corporation is a retailer employing more than one million workers in the United States alone and had annual sales of $244.5 billion in 2003.[3] Yet if, in some far future, the *New Palgrave's Dictionary of Economics* is the only work on the economy to survive, future historians/economist would not even know retailing existed.[4] The *New Palgrave's* is the economist's encyclopedia *par excellence*. It was organized by, edited by and written by economists "…to define the state of the discipline by presenting a comprehensive and critical account of economic thought.[5]" In four massive volumes totaling more than 4,000 pages, there is not a single article on retailing or wholesaling. Nine articles are concerned with "distribution" but these are solely concerned with either how income is divided among the factors of production, or how it should be divided (the ethics of income distribution) or with probability distributions. Although retailing was less important early in the nineteenth than in the twentieth century[6], historian/economists of the far future would be slightly better informed if only the original *Dictionary of Political Economy* edited by R. Inglis Palgrave (1894-1899) had survived, rather than its successor. In the old Palgrave's Dictionary, the first of its kind in English, the historian/archaeologist would find several discussions of wholesale and retail trade, including a discussion of the advantages/disadvantages of operating at larger scales. They would be still better informed if the surviving encyclopedia was not exclusively devoted to economics, but had a broader focus on "social science." *The Encyclopedia of the Social Sciences,* edited by the economists Edwin R. Seligman and Alvin Johnson, contains articles on anthropology, history, political science, psychology, sociology, statistics, law, education and social work in addition to economics. It was published in 15 volumes between 1930 and 1935. It contains numerous articles on retail trade, retail credit and one on wholesaling. Although its successor, *The International Encyclopedia of the Social Sciences,* 1968, edited by sociologist David Sills with statistician W. Allen Wallis as chairman of the editorial advisory board, has no article containing the words "retailing" or "wholesaling" in the title, it does in fact contain an excellent discussion of both under the somewhat eccentric title "Internal Trade." by economist Richard Heflebower.[7] In that article, Heflebower notes, for the first time by an economist, the curious neglect of retail/wholesale trade by economists, despite its economic importance, and calls retailing/wholesaling " …the stepchild of economic analysis." He goes on to say,

> " But in the development of price and output theory, economists have dealt primarily with the structure and performance of the raw-material and manufacturing industries (or with transport and electric power) and have treated these as if manufacturers sold finished products directly to consumers or, alternatively, as if the distributive trades were analytically neutral. Except in an occasional empirical study (e.g., Adelman 1959, pp. 109-149, 248-274), the distributive trades have been neglected in their roles as buyers

[3] See Wal-Mart's 2003 10-K filing at the SEC.
[4] *The New Palgrave: A Dictionary of Economics,* edited by John Eatwell, Murray Milgate and Peter Newman, 1987.
[5] ***Ibid.*** p. . See also George Stigler's fairly critical review, "**Palgrave's Dictionary of Economics**," *Journal of Economic Literature*, Vol. 26 (4) pp. 1729-36 1988.
[6] In 1900, wholesale and retail trade accounted for about 10% of the total labor force, compared to about 17% of total non-farm labor force in 1987. See …
[7] "Internal Trade,"*International Encyclopedia of the Social Sciences*, Volume 7, pp. 493 -499.

> from, or as potential entrants into, supplying into supplying industries (whereby they influence significantly the performance of earlier-stage markets), or as resellers that affect the information provided about, and consumers' choices among, goods made by rival manufacturers. This lack of economic analysis and the general confusion about the economic function of the distributive trades have contributed in no small part to the question asked in the book Does Distribution Cost Too Much?, which is answered affirmatively there and elsewhere". [8]

Heflebower's observations, if anything, seem even truer today than they were more than three decades ago. Overall, I conclude that economists as a profession have shown surprisingly little interest in the distributive trades, despite their large and growing economic importance.

Both Textbooks and Empirical Studies Typically Either Ignore Retailing or Assume 100% Pass-through of Manufacturer or Importer Price Changes

Economists routinely attempt to assess how much of an industry-wide upstream cost increase will be passed through to consumers. To do so, one would think they would have to consider how wholesalers and retailers react to an increase in their costs. Economists are not famous for agreeing with one another, but on this question there is a surprising uniformity of practice in the way economists deal with pass-throughs to consumers. Most textbooks, whether the question arises in the context of an excise tax[9], a unit tariff[10], or of manufacturer price increases in antitrust litigation[11], provide the same analysis and the same answer. The pass-through to consumers will be *100% or less*. The exact percentage pass-through depends on the relative magnitudes of the elasticities of supply and demand. The more elastic supply relative to demand, the closer the pass-through to 100%. Either 100% of the cost increase will be passed-through, or some of the cost increase will be "absorbed" by firms in a competitive industry. In no case will more than a 100% of the increase be passed on. The use of the word "absorption" in this context can be misleading. Sellers in a perfectly competitive industry don't "absorb" *any* cost increase in the usual sense of that word, i.e., accept lower profits rather than raise price sufficiently recover the full cost increase. By assumption there are no excess profits that can be reduced in the first place. What is really happening is completely a reflection of the assumption that supply curves slope upward, or, that firms are subject to *diseconomies* of scale. The reason competitive firms may pass on less than 100% of a cost increase is that, at the smaller scale induced by the price increase, their average and marginal costs, exclusive of the tax increase, are assumed to go down. So in the sense of accepting lower profits, competitive firms don't "absorb" any of a tax increase. If price rises by less than the amount of the tax increase it is only because firms are assumed to be subject to diseconomies of scale, so unit costs are lower at the lower output induced by the cost increase. Economies of scale, or decreasing unit cost with increases in output,

---

[8] *Ibid.,* p. 493.
[9] E.g. Musgrave & Musgrave, Public Finance in Theory and Practice, 3rd Ed., 1980, 261-264.
[10] E.g., Paul Krugman and Maurice Obstfeld, *International Economics: Theory and Policy, 5th Edition, 2000, pp. 190-191.*
[11] William Landes and Richard Posner, *Should Indirect Purchasers Have Standing To Sue Under the Antitrust Laws? An Economic Analysis of the Rule of Illinois Brick.*U. of Chicago Law Review, V. 46 (#3), Spring 1979.

are excluded by assumption because they are inconsistent with perfectly competitive equilibrium.

In practice, the answer economists generally give is even simpler: 100% will be passed through to consumers. It is as though neither wholesalers nor retailers exist. I'll provide three brief examples that could be almost indefinitely multiplied.

International trade

Economists often attempt to assess the impact of tariff, quota changes of anti-dumping actions on consumers and hence on retail prices and quantities. A Google internet search will quickly reveal a plethora of quantitative models to do this ranging from small partial equilibrium models built using the GEMS modeling language covering only a few sectors and two regions to grand Computational General Equilibrium (CGE) models such as the Global Trade Analysis Project (GTAP) model. That these models are widely used (or at least cited) in policy deliberations can be confirmed by an examination of ITC decisions and reports, or FTC or GAO reports etc. Not even the grandest of these models, however, contains any explicit treatment of retailing, nor any discussion of why retailing is omitted, nor of how to adjust the model's producer prices to obtain retail prices.[12] The only example I have found that explicitly mentions retailing is in an ITC Report that attempts to measure the economic effects on the U. S. domestic economy of eliminating all significant import restraints. In this case, the ITC used both its own CGE model and partial equilibrium models for those sectors where data availability prevented the use of the CGE model.[13] The ITC CGE model is quite elaborate, with more than 70 major sectors. One of the interesting features of this model is that it "… explicitly accounts for upstream and downstream production linkages."[14] Retailing/wholesaling is one sector. Chapter 3 of this report is especially interesting because it attempts to measure the effects on U.S. consumers of removing import restrictions on apparel and textiles. Apparel and textiles are sold at relatively high retail %rgms of 50% or more. In 1995, for example, a pair of men's denim jeans had an average manufacturing/wholesale price of $14.20, but sold at retail for $32, a %rgm of almost 56%.[15] Men's cotton dress shirts with an average manufacturing/wholesale price of $11.24, sold at retail for $31.38, a %rgm of 64%.

---

[12] For example, the 63 page description of the GTAP model does not contain the word "retail." See Hertel, Thomas W. and Marinos E. Tsigas **Structure of GTAP,** on the web at: www.**gtap**.agecon.purdue.edu/resources/download/86.pdf

[13] United States International Trade Commission***, THE ECONOMIC EFFECTS OF SIGNIFICANT U.S.IMPORT RESTRAINTS,*** Third Update 2002, Investigation No. 332-325, June 2002, Publication 3519

[14] *ibid.,* D-4.

[15] See table 1, p.8 in Jane Bondurant and Don Ethridge, "PROPORTIONS OF THE RETAIL DOLLAR RECEIVED BY COTTON INDUSTRY SEGMENTS: SELECTED CONSUMER GOODS," available at: www.aeco.ttu.edu/Publications/ 1998%20Beltwide/CER-98-19.pdf.

On the basis of its model, the ITC staff concludes that abolishing all tariffs and import quotas on apparel would lead to more than a 26% increase in imports and a 17% decrease in the "composite" price (an average of the prices paid by importers and the prices received by domestic apparel producers).[16] Now the question is, how does this 17% fall in producer prices translate into retail prices for apparel? The table does contain a row for wholesale/retail trade. The "composite price" shown for this sector is essentially the $ wholesale and retail margins per $ of final sales. This price is shown to increase slightly by .4%. So after restrictions are removed, for example, men's dress shirts, which had wholesaled for $10 would now cost only $8.30. Distribution margins would rise, however, from $10 to $10.04, so the new price paid by consumers would be $18.34. A 17% reduction in producer/import prices would lead to a 8.3% fall in retail prices. [The retail price with a 100% pass-through would be $18.30.] Despite taking wholesaling/retailing into account, the ITC model, like most others, essentially assumes 100% pass-through of producer prices to the consumer level. If, however, retailers preserve their %margins of say 50%, then the actual retail price would be $16.60 (double the invoice cost – the keystone formula). Retail prices would fall by 17%, or at more than twice the rate estimated by the ITC. If the assumption of constant retail %rgms, rather than constant $rgms, is closer to the truth, then economists generally substantially underestimate the consumer benefit of reducing trade restrictions and underestimate the consumer harm of increasing trade restrictions.

Oil Prices

The Energy Information Agency (part of the Department of Energy) uses an elaborate model to predict prices of gasoline, heating oil etc.[17] At the refiner level, the model is quite elaborate, containing three separate linear programming modules to represent refining practices in different areas. These modules are used to compute the marginal costs of both traditional and new reformulated petroleum products designed to meet new State and Federal emission standards. The model even takes capacity expansion into account. The model's treatment of the pass-through of refiner prices at the retail level, however, is very simple,

> The costs of distributing and marketing petroleum products are represented by adding fixed distribution costs to the marginal and refinery fixed costs of products. The distribution costs are applied at the Census Division level (Table 60) and are assumed to be constant throughout the forecast and across scenarios

In other words, this otherwise elaborate model assumes a 100% pass-through of refiner costs at the retail level. In this case, however, at least for gasoline, there is substantial evidence that 100% pass-through is in fact a reasonably good approximation of the facts.[18] Between 1983 and 1998, for example, the average wholesale price of regular unleaded gasoline fell from 89.5 cents per gallon to 49.9 cents per gallon.[19] Retail prices

---

[16] See ITC Report*., op.cit.,* Table 3-6 on p. 67-68.

[17] A comprehensive description of the EIA NEMS PMM model can be found on the web at http://www.eia.doe.gov/oiaf/aeo/assumption/petroleum.html.

[18] See John Cook, EIA, *Price Changes in the Gasoline Market: Are Midwestern Gasoline Prices Downward Sticky? February, 1999* www.**eia**.doe.gov/pub/oil_gas/petroleum/analysis_publications/**price_changes**_gas_**market**/pdf/**price**_change.pdf

[19] See , Energy Information Administration/Preliminary Petroleum Marketing Annual 2003,

(excluding sales taxes) fell from 98 cents a gallon to 63.4 cents. A prediction based on the assumption that retail $rgms would remain constant would have been closer to reality (58.4 cents per gallon), than one based on the assumption that %rgms would remain constant (54.6 cents per gallon). Gasoline, however, is the only commodity I have found that is consistently better modeled by an assumption of unchanging $gm than by an assumption of unchanging %rgms.

Cigarette Litigation Costs

Two FTC Staff studies on the effects of the proposed federal tobacco litigation settlement provide another example.[20] The FTC Staff's base line scenario assumed that the proposed settlement costs would lead cigarette manufacturers to increase their factory prices by about 57 cents per pack by year 5 post settlement. Retail prices were also assumed to rise by 57 cents per pack, that is, retailers would pass-through the factory price increase dollar for dollar. The Staff's entire rationale for this assumption is contained in the following sentence.

> "The gains or losses from a pass-through different from 100 percent will accrue entirely to the manufacturing sector, so long as the wholesale and retail distribution of cigarettes is competitive, with distributors obtaining no more than a competitive rate of return for providing their services."[21]

Steiner[22] criticized the Staff Report writing that since the FTC assumptions implied a 45% increase in the factory price and a substantial reduction in overall unit sales (11%) that "…it seemed inconceivable to this former consumer goods manufacturer that wholesalers and retailers would not boost their dollar margins by a substantial amount." With an unchanged $rgm per pack, and about 2.6 billion fewer packs sold, retailers would have far fewer gross margin dollars to cover non-invoice costs such as the foregone interest cost of carrying inventory and shrinkage and the fixed costs of retailing. Although the proposed federal settlement never went into effect, the major cigarette manufacturers reached a settlement with the states and raised their factory prices very substantially between Nov. 1, 1998 and Nov. 1, 1999. Steiner's careful estimate of the actual pass-through led him to conclude that the % trade margin, rather than falling to about 13% as implied by a 100% pass-through, remained at about 17%. [Third Relevant Market, pp. 755 – 757.] Two distinguished economists, both at one time Directors of the Bureau of Economics at the FTC, replied to Steiner's criticism, saying

---

Table 6. U.S. Refiner Motor Gasoline Prices by Grade and Sales Type (Cents per Gallon Excluding Taxes).

[20] See FTC Staff Reports, *Competition and the Financial Impact of the Proposed Tobacco Industry Settlement* (September 1997*)* and *Evaluation of the Tobacco Industry Analysis* Submitted to Congress on October 8, 1997 (November 10, 1997). Congress had asked the FTC to predict the economic effects of the proposed tobacco industry settlement

[21] P. 26 of the September, 1997 Report cited ***supra*** in note 19.

[22] See the exchanges between Albert Foer, Robert Steiner and Jeremy Bulow and Jonathan Baker, December 20, 1999 at http://www.antitrustinstitute.org/recent/48.cfm. Also see Robert Steiner, "TheThird Relevant Market*",* ***The Antitrust Bulletin****,* Fall 2000, pp. 745 – 758.

> "… an industry-wide cost increase will be completely passed-through to consumers, without an additional markup, if the industry is competitive and firms have constant marginal costs. This is, after all, how the distribution sector of the cigarette industry looks: in most localities, cigarette consumers can choose among a large number of convenient alternative sellers when purchasing that product, and there is no reason to think that incremental costs would vary for most wholesalers and retailers were product sales to decline."

Notice, nothing is said about how retailers could continue to cover their fixed costs with the same $rgm and far fewer unit sales. Only incremental costs are mentioned.

Why Do Economists Believe that Retailing can be Neglected?

Why do economists so often behave as though manufacturers sell directly to consumers, as though retailing makes no essential difference in the determination of retail prices? The basic answer, as we have already seen, is that economists believe that retailing can be acceptably modeled as a perfectly competitive industry with constant marginal costs. As another economist, Ronald Cotterrill, wrote,

> "Now that I have shown that manufacturers pass on 100 percent (or more) of this overcharge. The remaining link in the price transmission process is to explicitly incorporate the away-from-home and food retailing industries into the model. Away-from-home purveyors are straight forward because these firms operate in competitive markets. Any increase in their costs is passed through at the 100 percent rate in the general as well as the actual price transmission model."[23]

If retailing can be modeled as a perfectly competitive industry, then the "derived demand theorem" justifies analyzing the manufacturing sector alone without concern for the details of the retail sector. Kristian Palda, in an award winning University of Chicago dissertation/book, explained the practice as well as anyone. Palda, who was primarily interested in quantifying the effect of advertising expenditures on sales, had run into a problem. In his regression analysis, he needed to "hold constant" factors other than advertising that might have affected the sales of "Lydia Pinkham's Vegetable Compound." The first factor that would occur to any economist is the price paid by wholesalers and retailers for the product. The problem was there had been only four changes in the factory prices from 1905 to 1960. So there was little variation. And what little variation there was showed no consistent pattern. For example, between 1917 and 1918, factory prices increased by 24%, yet dollar sales grew by almost 50%; during 1947 the average factory price was 7% higher than it had been the year before, but dollar sales fell by 12%. The lack of variation in factory prices precluded their use in a regression, but Palda thought there was a more fundamental problem.

---

[23] **Estimation of Cost Pass Through to Michigan Consumers in the ADM Price Fixing Case,** Food Marketing Policy Center, University of Connecticut**,** Food Marketing Policy Center**,** Research Report No. 39**,** November 1998.

> "A more fundamental problem is raised by the nature of the relation between the factory and the retail price of the vegetable compound. ***While price to distributors is the relevant price in the analysis of the company's sales variations, it could be normally expected that it is but a faithful reflection of the retail price***. ***This is so because demand on the wholesale level is typically envisaged as being a derived demand.*** However, if, wholesale price movements are in no way systematically related to changes in the retail price, attention should be shifted to prices at retail level. It appears that such unusual circumstances, in which prices at retail moved independently of prices at wholesale, prevailed over long stretches of the company's history."[24] [Emphasis added]

Shifting attention to retail prices, however, brought to light another problem. Despite the fact that the company had a strict policy of no trade discounts, no quantity discounts, no discount for cash etc., retail prices varied greatly. In 1939, for example, wholesalers (62% of sales), retail chains (25%), independent retailers (10%) and mail order houses (3%) all paid the same price – 83.33 cents per bottle of the liquid tonic.[25] Wholesalers were expected to apply their normal markup and charge retailers $1 a bottle and retailers, in turn, would sell to consumers at $1.50. Yet a 1939 survey "found that 99.88 per cent of sales were made at prices below $1.50 and that 80 per cent of retail sales were made at prices below the product's normal wholesale price of $1.00."[26] Lydia Pinkham, a pioneer in early heavy advertising and whose advertising to sales ratio exceeded 50% between 1908 and 1935, was a very popular product and retailers, complaining bitterly, found that they had to sell it at much less than their normal retail gross margins. So retail prices, far from being a "faithful reflection" of wholesale prices, showed great variation despite the fact all wholesalers and retailers paid the same invoice price. Moreover, both wholesalers and retailers sold the product at less than their traditional margins and, sometimes, even at a loss.[27] Palda clearly believed that wholesale (or factory) prices were usually a faithful reflection of retail prices; that the Lydia Pinkham case was a rare anomaly. I believe this to be the crux of the issue. To the extent that wholesale prices are a faithful reflection of retail prices, that is, to the extent that Marshall's derived demand theorem applied to retailing is roughly in accord with reality, then Steiner's "dual stage" theory is a needless and mistaken complication. To the extent that wholesale prices are not a faithful reflection of retail prices, however, economists need to pay a lot more attention to the details of the wholesale and retail distribution of a product, if they want to translate a change in a factory price into a retail price. We now know, thanks to the abundance of scanner data, that the Lydia Pinkham case is not a rare anomaly, in the sense that retail prices often vary when wholesale prices don't. Indeed, retailers

---

[24] Kristian S. Palda, *The Measurement of Cumulative Advertising Effects*, 1964, p. 39

[25] Neil Borden, *Advertising: Text and Cases*, 1950, p. 204. It is thanks to litigation that so much of what would ordinarily be considered highly proprietary business information is known. In the mid-1930's the two different families heir to the company bitterly disagreed on company policy. The Massachusetts Supreme Court decided the case in 1937 and made public detailed records of the company's operation.

[26] *ibid.,* p. 205.

[27] Over 6% of sales were at prices between 59 and 81 cents, i.e., at below retail invoice cost, clearly "loss leaders." *ibid.,* p. 205.

generally follow one of two major pricing strategies: "Hi-Lo" pricing or "Every Day Low Pricing."[28] Supermarkets and drugstore chains generally follow the first strategy; Wal-Mart, Home Depot generally follow the second. Many, perhaps most, retail price changes for Hi-Lo stores are the result of the retailer temporarily reducing his margin and not because the manufacturer has changed his factory price.[29] For such stores, retail price changes do not simply mirror factory price changes, nor are factory price changes simply mirrored at retail level. This evidence suggests that the derived demand theorem is often inapplicable to retailing.

Actual Retail Industry Very Different from the Assumptions underlying the Model of Perfect Competition

If retailing, for most purposes, could be reliably modeled as a perfectly competitive industry, there would be no harm, and much simplification gained, from ignoring the separate effects of retailing. Then the derived demand theorem would lead to reliable predictions. Retail price would equal the factory price plus transportation plus the marginal cost of the retailer. Since falling marginal costs are ruled out in a perfectly competitive world, the retailer's own marginal cost must be constant or rising with unit sales. If the latter, then any change in factory price will be fully and faithfully passed on to consumers. But for many reasons, it seems unlikely that the MPC model either has can be a satisfactory model of retailing.

For one thing, the assumptions that go into the model are vastly different from the actual world of retailing. Now it is true, as Milton Friedman pointed out almost half a century ago, that the appropriate test of a theory is how its predictions of the phenomenon of interest stack up against reality, not how realistic any one of its assumptions may be.[30] Newton's first law of motion (a body in motion will continue in motion at constant velocity in a straight line unless a force intervenes to change that motion) underlies one of the most successful scientific theories in all human history. Yet any four year old pushing a wooden block on a rug knows that it is not true, or at least not always true. So,

---

[28] See, e.g. Levy, Daniel, Mark Bergen, Shantanu Dutta, and Robert Venable (1997) "The Magnitude of Menu Costs: Direct Evidence from Large U.S. Supermarket Chains," Q*uarterly Journal of Economics*; 112, pp. 791-825, Daniel Hosken, David Matsa, and David Reiffen, "How do Retailers Adjust Prices?: Evidence from Store-Level Data"**,** FTC website**,** Jan. 2000 and Daniel Hosken and David Reiffen PRICING BEHAVIOR OF MULTIPRODUCT RETAILERS, Federal Trade Commission Revised: May, 2001. Varian's elegant model does, I think, capture the essence of Hi-Lo pricing, but is not consistent with the co-existence of Hi-Lo and EDLP SeeVarian, Hal R. (1980) "A Model of Sales", *American Economic Review*; 70, pp. 651-9.

[29] "We establish a number of interesting facts about retail prices in the U.S. First, most products appear to have a "regular price." Using the BLS data, we find that for the 20 categories of products in our sample, products are priced at exactly their annual modal price 62% of the time. Moreover, in every category, products are priced at their annual mode at least 40% of the time." Daniel Hosken, David Matsa, and David Reiffen, supra note 28.

[30] "The Methodology of Positive Economics," *Essays in Positive Economics*, U. of Chicago Press, 36-37, 1953

very successful theories can rest on assumptions that are not always true. But it's also true that every assumption known to be false about the domain of the real world you're trying to model, produces false predictions.[31] So while it's true that every useful theory makes "simplifying" – known to be false – assumptions about the world, the art comes in choosing simplifying assumptions that aren't too false for the domain you're interested in. The motions of the moon or the path of a rocket in the atmosphere is not materially changed by the reality of friction. But can the same be said for the assumptions of the PC model applied to retailing?

The assumptions of the PC model as applied to retailing that I find most troubling are that it's "spaceless", that retailers face infinitely elastic demand curves, that retailers sell only one product, and most disturbing of all, that there are no economies of scale in the provision of retail services or for consumers who must travel to reach stores. Less troubling, but often false, is the assumption that the retailer sells all units of the product at the same price and that retailers all pay the same price to manufacturers or wholesalers for each unit. Least troubling is the assumption that there is free entry and that free entry leads to zero excess profits. Not only is the vast number of retailers consistent with the notion that there are low barriers to entry, but entry in retailing often doesn't require building a new store. It can happen simply by reallocating existing shelf space to a product, or a product category that is expected to yield more retail gross margin dollars than an existing product or category.

Space

Recognizing that different stores have different locations and are different distances for different customers and that transportation is costly, leads quickly to the rejection of the notion that every retailer faces an infinitely elastic demand curve. Stores located closer to given consumers will have an advantage over stores that are further away.

One Product Retailers

One of the great retailing innovations of the 20th century, the A&P economy store introduced by John Hartford in 1912 carried only 300 SKUs.[32] In 2002, a typical conventional grocery store carried about 22,000 SKUs, a food/drug combo 62,000 and a "supercenter" 125,000.[33] Even a traditional convenience store carried 3,500 SKUs. Gasoline stations, in the past, were probably the closest thing in reality to the single product retailer.[34] Specialized gasoline stations are, however, increasingly losing out to gasoline station/convenience store combos and to competition from Wal-Mart and other broad array low-price retailers.[35] The importance of the single product assumption is that

[31] See Paul Samuelson on the "F-Twist", Comments On Ernest Nagel's "Assumptions in Economic Theory", #129 in Vol. II of his Collected Papers.

[32] William Walsh, *The Rise and Decline of the Great Atlantic & Pacific Tea Company*, pp. 28 – 30.

[33] See http://www.bishopconsulting.com/store_format.cfm.

[34] This may explain, in part, why the assumption of a 100% pass-through is more satisfactory for gasoline than any other commodity I have studied.

[35] See the roundtable discussion in the Jan., 2001 issue of the National Petroleum News entitled ".The year 2000 was a very STRESSFUL YEAR for most in the petroleum/convenience-store industry. Several PEAKS in the price of gasoline sent margins to RECORD LOWS, forcing marketers to rely on other areas to keep stores PROFITABLE" at www.petroretail.net/npn/2001/0101/0101cvr.asp.

it rules out pricing some products low or even at a loss to attract customers who will buy a "basket" of different products. It also precludes any analysis of why some stores carrying a more limited array of goods, e.g. convenience stores, can nevertheless attract customers who could have gone to stores carrying a much broader array of goods. Limiting the model to single product retailers also precludes analyzing economies of scope which are certainly important for consumers who can engage in "one-stop" shopping. The economies of scope for consumers who buy entire baskets at the same store on the same shopping trip simply reflect the transportation economies of scale. The cost of transporting 30 items in several shopping bags by automobile is likely to be little more than transporting one item. The economist Geoffrey Heal has emphasized the importance of economies of scale in transportation in an elegant paper that doesn't seem to have attracted much attention from the profession. [see Heal, Geoffrey, "Spatial structure in the retail trade: a study in product differentiation with increasing retums," THE BELL JOURNAL OF ECONOMICS, Autumn 1980, vol. 11, no. 2, 565 - 583.] The brilliant book by Robert W. Bacon [*Consumer Spatial Behaviour: A Model of Purchasing Decisions over Space and Time,* Clarendon Press Oxford; 1984] not only stresses the importance of recognizing economies of scale in transportation in explaining shopping behavior, but also tackles the problem of multi-product retailers and how the concurrent existence of convenience and supermarkets can be explained. Walter Oi[36] explains the growth of the distributive trades largely through economies of scale …

Economies of Scale: Retail Stores

In 1912 John Hartford's A&P economy stores occupied only about 600 sq. ft. of space. In 2002, a conventional grocery store occupied 25,800 sq. ft. of space, a food/drug combo 55,700 sq. ft. and a "Supercenter" occupied 195,000 sq. ft. or more than two orders of magnitude more than the old economy store. Not all stores in 2002 were large. "Convenience" stores, of which there were many, occupied only 2750 sq. ft. or only four times the size of the old A&P stores. The fast growth of Supercenters and Food/drug combo stores suggests that not only are there significant economies of scale, but they have been increasing over time. On the other hand, a very interesting phenomenon is the persistent success and growth of stores at different scales. A Supercenter is almost 71 times larger than a typical convenience store, yet both can prosper in the same market area. Clearly, there is no simple relationship between size and ability to compete for retail stores.

Economies of Scale: Transportation

As suggested above, economies of scale in transportation are also important, perhaps even more important, than economies of scale in the operation of retail stores in understanding the complex reality of modern retailing. Think of the entire manufacturing/wholesaling/retailing system as a network. There are tens, or even hundreds, of thousands of factories and farms scattered over the world shipping products

[36] Oi, Walter Y., " The Indirect Effect of Technology on Retail Trade" in *The Impact Of Technological Change On Employment And Economic Growth*, Edited by Richard M. Cyert and David C. Mowery, 1988, 329-375

in bulk to thousands of wholesalers and to retail warehouses. The wholesalers and retail warehouses in turn ship smaller quantities to hundreds of thousand retail outlets. Millions of consumers transport themselves to these retail outlets and transport very small unit volumes of goods purchased to their homes. Whether we're dealing with oil or natural gas pipelines, the telecommunications system, municipal water systems, cable television/internet systems, economies of scale in transportation is a crucial determinant of where pipelines are placed, and how many are laid and at what capacity they are built. Another constant in all these networks is the "problem of the last mile", the problems posed for the system of transporting small quantities to the ultimate consumers. In retailing, consumers themselves solve this "last mile" problem by providing their own transportation. Much of wholesaling/retailing seems to be a product of balancing relatively large shipments to warehouses and retail outlets against longer trips by consumers who also benefit from transportation economies of scale.

Why Don't Consumers Generally Buy Directly from Manufacturers?

One reason is that the cost of shipping a single unit from the factory that makes it will be much higher than the unit shipping cost to a retailer who orders many at a time. Suppose you want to buy an "Easy-Bake Oven and Snack Center" which has a shipping weight of 4 pounds. If you wanted to get it within 3 – 5 days, you could ship via Global Express Mail at a cost of $34.90.[37] If you're willing to wait 4 - 6 weeks, the cheapest mode would run $13.30. Although I don't know the Shanghai factory price, I do know that Toys-R-US online everyday low price in early June of this year was $19.99 If, instead of one oven, you ordered ten, the per unit shipping cost would fall from $34.90 to $18.80 for Express – Mail and from $13.30 to $6.37 for surface shipping There are substantial economies of scale in transportation and this is certainly one reason and one important service retailers provide to their customers. An interesting case in point that illustrates the economies of scale in transporting large shipments and the use of especially low prices on well known items to attract customer traffic is as follows. Wal-Mart ran a chain-wide ad the day before Thanksgiving in 2003 offering the Easy-Bake Oven at $8.88. They purchased hundreds of thousands of ovens from Hasbro and undertook to transport them from Shanghai at their own expense.

Absent economies of scale in retail store services and economies of scale in transportation, it is hard to see what service, if any, could not be supplied more cheaply by the manufacturer (especially with a "satisfaction guaranteed" return policy) than by a retailer. Using a model that rules out economies of scale of any kind to analyze retailing, strikes me as akin to studying the prostitution industry using a rigorous model that assumes people get no pleasure having sex.

Economies of Scale Totally Incompatible with the Modern Theory of Perfect Competition

The modern theory of perfect competition, whether partial equilibrium models or the CGE general equilibrium models such as that used by the ITC, requires a very strong adherence to the assumption that there are no economies of scale. As the ITC staff put it,

> A central assumption in many general equilibrium analyses is that production exhibits *constant returns to scal*e. This assumption results in an

[37] I am actually quoting the USPS rate to ship to China using their online international postage calculator.

> almost horizontal domestic supply curve for a given product, which means that price changes tend to have minimal effects on supply when the domestic product has a large market share. In contrast, the pervasive assumption in partial equilibrium modeling is that production exhibits *decreasing returns to scal*e. These differing assumptions are rooted in the fundamentally different questions asked of each model. Partial equilibrium models generally are used to simulate current policy changes on a specific market. In these cases, an upward-sloping supply curve is appropriate because the analysis covers a limited time horizon. Comparative static welfare analysis in general equilibrium models is usually focused on simulating an alternative equilibrium in which all resources have adjusted to their most productive uses. This perspective requires long-run supply elasticities, which are reflected in the horizontal supply curve.[38]

When I was a student, we were taught that a competitive equilibrium was compatible with fixed costs, indeed with fixed plant costs of different sizes. An industry could be composed of identical firms each with U-shaped average cost curves. The competitive equilibrium occurs at the point where (the rising) marginal cost met the minimum point on the average cost curve. Jacob Viner's valiant attempt to square the demands of the theory of perfect competition with the common sense reality of fixed plant costs seemed to work and was a staple of textbooks for many years.[39] Alas, Telser in developing Edgeworth's theory of the core has shown that Viner's construction does not work.

> Perhaps the most surprising result comes from the application of this theory to the widely used analysis of an industry consisting of identical firms with U-shaped average cost and increasing marginal cost. Viner (1931) is the best-known reference and is the major source of most of the textbook treatments of this problem. For the Viner industry there is a unique firm output rate at which the average cost is a minimum and equal to marginal cost. …. Most economists believe that this standard case, the Viner industry, does give a competitive equilibrium. This is false. The core is almost always empty for the Viner industry, so that there is no competitive equilibrium.
>
> …
>
> Many mathematical treatments of the existence of an equilibrium assume cost conditions equivalent to constant returns to scale. This implies finitely elastic supply schedules. The analysis in this chapter shows that a perfectly elastic supply schedule not only is sufficient for a competitive equilibrium but also is necessary.[40]

So the modern, mathematically rigorous theory of perfect competition allows no fixed costs, no set-up costs, no indivisible factors, none.[41]

The importance of production economies of scale in explaining the observed concentration of people and production across geographical space is widely appreciated by economists. As T.C. Koopmans, one of the founders of modern general equilibrium

---

[38] ITC – Report, *supra,* note 20 at p. 3.

[39] "Cost Curves and Supply Curves," Reprinted in *A.E.A.Readings in Price Theory, pp. 198-232. Originally published in 1931.*

[40] See Lester Telser, *Economic Theory and the Core*, ch.3, pp. 88 – 137.

[41] Frank Machovec has written a very interesting account of the PC model came to dominate modern economics and displace the earlier process-oriented view of competition. See *Perfect Competition and the Transformation of Economics*, Routledge, 1995.

theory has observed, if there were no economies of scale in production[42], there would be little or no transportation of produced goods even in a world where people and resources are in different spatial locations.

> Just as indivisible commodities can be introduced in a spaceless model, so also can locational problems be studied without recognizing indivisible commodities. Much of the literature on location problems is of that character. The data that define the location problem in this case are the geographical distributions of various mineral deposits, geographical variations in the suitability of land for the production of various crops, and costs (or input requirements) of transportation between pairs of points. If, as in most of these studies, location problems are looked at from the point of view of the individual firm, the geographical distribution of markets is an additional datum. In models, descriptive or normative, which are concerned with the location of all economic activities, this distribution becomes itself a variable.
>
> As long as indivisibility of commodities is not recognized, these models fail to grasp the essential character of the problems posed by urban conglomerations. The manner in which the various activities of a metropolitan area are (or could best be) arranged in space has very little relation to mineral deposits or grades of agricultural land, although the particular collection of activities represented in any such area may well be strongly influenced by the character of these factors in the surrounding country. ***If we imagine all land to be of the same quality, both agriculturally and in amount and accessibility of mineral resources, then an activity analysis model of production that includes the proportionality postulate would show a perfectly even distribution of activities to be most economical. Each square inch of area would produce the same bundle of commodities from its own resources by the same bundle of activities, and all transportation would thus be avoided.*** [emphasis added].
>
> …
>
> Again, there is no doubt about the existence and importance of transportation cost or of intermediate commodities. One may conclude from these observations that, in regard to the allocation problems raised by indivisible commodities, with or without locational distinctions, theoretical analysis still has not yet absorbed and digested the simplest facts establishable by the most casual observation. This is a situation readymade for armchair theorists willing to make a search for mathematical tools appropriate to the problems indicated. Since the mathematical difficulties have so far been the main obstacle, it may be desirable in initial attempts to select postulates mainly from the point of view of facilitating the analysis, in prudent disregard of the widespread scorn for such a procedure."[43]

So if all manufacturing were subject constant returns to scale, manufacturers would not only sell directly to consumers, they would actually produce the item in each consumer's home. Transportation costs would be irrelevant and there would be no retailing. Of course manufacturing is subject to economies of scale, factors are indivisible etc. But even with manufacturing economies of scale, there might be little or no wholesale or retail trade if there were no economies of scale in transportation, i.e., if the unit cost of shipping items between two points was independent of the number of items shipped. If consumers could buy an item from the factory and pay to have it shipped to their home, why would they incur the travel costs from home to retail store and back again when the retail price would have to include the shipping costs borne by the retailer between the

[42] Koopmans argues that all economies of scale are due to "indivisibilities," hence he uses the latter term to refer to the former.

[43] ***Three Essays on the State of Economic Science***, pp. 153-154, McGraw-Hill, 1957

factory and the retail outlet? Total transportation cost would be reduced by eliminating retailers since the item would be shipped only once from factory to home, instead of from factory to store, then consumer to store and finally consumer with item to home. Economies of scale in transportation cost could account for the existence of retailers even if there were no economies of scale in operating a retail store.

The assumptions discussed above are clearly false, and they produce false implications about the real world of retailing.

Implications of the MPC Model of Retailing that Seem Patently False

Implication #1: The retail price is equal to the retailer's marginal cost which consists of the price paid to the manufacturer for the item plus his constant marginal cost of distribution. This is clearly false, and it has to be false for every retailer who must survive without an outside subsidy. All retailers have fixed labor and occupancy costs, in addition to non-invoice variable costs must be covered for them to survive.[44] Economists in search of everyday examples of instances where price must be above marginal cost to recover costs, don't need to leap into their cars and drive in search of a Hotelling bridge. A visit to the corner grocery store, cleaners etc. will serve as well, and will save on transportation costs.

Implication #2: All retail prices for the same item are the same at all stores. Infinitely elastic demand curves imply prices must be the same at all stores. Many studies have shown that prices at stores located very close to one another have significantly different prices at the same point in time.[45]

Implications of the PC Model of Retailing Inconsistent with Many Empirical Findings

Implication #3. When marginal costs are assumed to be constant, pass-through rates are 100%. In fact, numerous studies have found pass-through rates both substantially lower than 100% and some substantially higher than 100%. A recent study by Besanko et al. using Dominick's data, for example, found that "…70 percent of the estimated product-zone level pass-through rates are smaller than 1, while 30 percent are greater than 1. Approximately 14 percent of the estimated rates are statistically larger than 1. ..."[p.22]. Furthermore,

> We find that although category average own pass-through rates range from as low as 0.22 for toothpaste, to as high as 5.58 for beer, on average the estimated pass-through rates for this large supermarket chain are much higher than the percentage claimed by manufacturers in the Cannondale (2001) study for all retailers. There are substantial

[44] Despite its age, Bob R. Holdren's, *The Structure of a Retail Market and the Market Behavior of Units*, Prentice-Hall, 1960, chapter 3, remains, for an economist, one of the best accounts of the cost structure of a retailer. For the two stores examined in detail, non-invoice variable costs are about 2% of the cost of goods sold. See Tables 7 & 8.

[45] See, among many others, the careful studies by Allen Jung, published in the *Journal of Business in 1960s,* Howard Marvel on gasoline pricing, 1976, Pratt, John W., David A. Wise and Richard Zeckhauser, "Price Differences in Almost Competitive Markets," with, *Quarterly Journal of Economics*, May 1979, 189-211

> differences in pass-through estimates across retail stores and categories, and between products within categories. We find that 87 percent of own pass-through estimates are positive. As many as 14 percent of the own pass-through rates are significantly greater than one, implying that in these cases, on average the retailer offers a larger discount to the consumer than the retailer receives from the manufacturer, for the time period of our study. These findings challenge the "empirical generalization" that *most* products display passthrough much smaller than one (Blattberg et al. 1995), although our results are consistent with empirical findings of Armstrong (1991) and Walters (1989).
> … A notable finding from our analysis is the extent of statistically significant cross pass-through effects. As many as two-thirds of the estimated cross pass-through rates are significant. This implies that the retailer responds to a trade promotion for one brand by changing retail prices of multiple products in the category. We also show that the estimated own pass-through rates are biased if such cross effects are ignored. Interestingly, the cross pass-through effects of a given brand's wholesale price change are positive for some competing products in the category, and negative for others. [footnotes omitted][46]

The significant number of non-zero cross pass-through rates is one more indication that retailing will not be understood until economists explicitly recognize that retailers typically sell many products, not one.

Implication # 4: Retail margins depend only on marginal retailer costs and are independent of the elasticity of consumer demand. There should be no significant correlation of the type described by Steiner's second inverse association (the "Steiner Effect"), that the more prominent, the more advertised, the brand, the lower its retail margin and the higher its manufacturing margin. Yet there is substantial evidence of a negative association between the two, as discussed below.

Empirical Regularities that, though not Inconsistent with the PC model, are Unexpected and Puzzling.

There is massive evidence for Steiner's first inverse association; retailers choose relatively low margins on leading nationally advertised brands compared to lesser brands and store brands. The only explanation for this regularity available from the PC model of retailing is that the marginal cost of distributing the leading brands must be lower than for lesser and store brands. But there seem to be no obvious differences in marginal retailing costs that would explain this difference. Faster turnover of LBs has been offered as an explanation of the lower margins retailers put on them, but this explanation depends on retail economies of scale, and so is ruled out for the PC model. Thus the observed inverse association presents at least a puzzle for this theory, if not a contradiction. Many of the most striking changes observed over the last 150 years in retailing, e.g. the vast increase in the size of the average retail outlet and in the number of items stocked, seem inexplicable if viewed through the lens of a model that rules out economies of scale.

Many economists seem to use the PC model of retailing as a "default" position. Unless someone has shown that the model can't be used in a particular case, they will essentially

[46] Besanko, David, Jean-Pierre Dube, Sachin Gupta, "Retail Pass-through on Competing Brands", July 2002, p.7.

invoke it as a rationale for ignoring retailing altogether. Yet, economists have, by and large, ignored retailing, its regularities and Steiner's explanation. I argue economists' reluctance to seriously consider retailing is part of a broader aversion to economies of scale. As Brian Arthur has written,

> Ideas that invoke some form of increasing returns are now acceptable in economics indeed they have become highly fashionable. But this was not always so. As recently as the mid-1980s, many economists still regarded increasing returns with skepticism. In March 1987 I went to my old university, Berkeley, to have lunch with two of its most respected economists. What was I working on? Increasing returns. "Well, we know that increasing returns don't exist," said one. "Besides, if they do," said the other, "we couldn't allow them. Otherwise every two-bit industry in the country would be looking for a handout." I was surprised by these comments. Increasing returns did exist in the real economy, I believed. And while they might have unwelcome implications, that seemed no reason to ignore them.

Little makes sense in retailing, including its existence, unless you understand that there are economies of scale, especially in transportation. Yet many theoretical economists especially have long shunned models involving economies of scale, largely because they have not found a general model that can both take economies of scale into account and that yield theorems of the scope and interest produced by models of perfect competition and monopoly. Despite the widespread knowledge of the extent and importance of economies of scale, mainstream economists have, until recent years, steadfastly refused to acknowledge them in their models. Rather these models have been premised on universal constant returns to scale. James Buchanan cites both technical and ideological reasons for this otherwise puzzling state of affairs.[47] He also sees a widespread acknowledgement of "…the aridity and emptiness of highly formalized general equilibrium analysis…" and a rediscovery and a return to increasing returns. I have argued that there is no theoretical or empirical support for this position and much empirical evidence against it. I turn now to a discussion of Steiner's two inverse associations and how he explains them

Steiner's Two Inverse Associations and How He Explains Them

Steiner's theory offers an explanation of two otherwise puzzling empirical regularities. First, retailers usually price "leading national brands" [LBs] so that they yield lower dollar and percentage gross margins[48] [$rgms and %rgms] than on competing fringe

[47] "The Return to Increasing Returns: An Introductory Summary," Chapter 1 in *The Return to Increasing Returns,* Edited James M. Buchanan and Yong J. Yoon, U. of Michigan Press, 1994. See also Elhanan Helpman, "Increasing Returns, Imperfect Markets and Trade Theory", Chapter 7, Vol. 1, *Handbook of International Economics,* Jones & Kenen (Eds.) North-Holland, 1984.

[48] The dollar gross margin ($gm) is the retail price less the manufacturer price. The %rgm is the $gm divided by the retail price.

brands or on "store brands" [SBs].[49] This is the first inverse association: the more prominent the brand, the lower its retail margin. Second, manufacturers of LBs tend to have high manufacturing margins compared to fringe brands and to SBs.[50] This is the second inverse association: the more prominent the brand, the lower its retail margin and the higher its manufacturing margin. The data in Table 1 illustrates both inverse associations for soft drinks as sold in a Chicago area supermarket called Dominick's Finer Foods over the period from 1987 to 1997.

[49] There is now a very large a growing body of empirical studies that show LB %rgms are low compared to to margins on fringe brands and store brands. See, e.g., Albion, Mark, *Advertising's Hidden Effects: Manufacturers' Advertising and Retail Pricing,* Auburn House Publishing, 1983, Chintagunta, Ailawadi etc. and the Barsky et al. work cited in the text.

[50] Data bases providing manufacturer margins and retail margins on the same on the same products are few and far between. So there is less quantitative evidence on the second association than on the first. However, the relationship has been shown for food products, toys, prescription drugs and apparel items. See Steiner "Inverse Association…" pp. 721 – 723.

Table 1

Manufacturer and Retail Prices and Margins for Soft Drinks - Dominick's Data Base 1987 - 1997

Leading Brand	Ratio LB Mfr Price to SB Mfr Price	Leading Brand Retail Price*	Store Brand Retail Price*	% Discount from LB Price	Dollar Gross Margin on Leading Brand	% Gross Margin on LB	Dollar Gross Margin on SB	% Gross Margin on SB	Dollar LB Mfr Margin	% LB Mfr Margin
	1	2	3	4	5	6	7	8	9	10
Coke Classic	2.35	$2.70	$1.96	27.5%	$0.35	13.0%	$0.96	48.9%	$1.35	57.4%
Schwepps Ginger Ale	2.17	$2.50	$1.92	23.1%	$0.33	13.0%	$0.92	47.9%	$1.17	53.9%
Pepsi Cola N/R	2.23	$2.63	$1.91	27.5%	$0.40	15.3%	$0.91	47.6%	$1.23	55.2%
Pepsi Cola Diet N/R	2.25	$2.57	$1.82	29.1%	$0.32	12.3%	$0.82	45.0%	$1.25	55.6%
Barq's Root Beer	2.00	$2.52	$1.81	28.1%	$0.52	20.6%	$0.81	44.8%	$1.00	50.0%
Schweppes Tonic N/R	2.28	$3.12	$2.30	26.5%	$0.84	27.0%	$1.30	56.5%	$1.28	56.1%
Schweppes Diet Tonic N/R	2.29	$3.14	$2.32	25.9%	$0.85	27.0%	$1.32	57.0%	$1.29	56.3%
R.C. Cola	1.88	$2.48	$2.09	16.0%	$0.60	24.2%	$1.09	52.0%	$0.88	46.8%
A&W Rootbeer Reg	1.78	$2.03	$1.75	13.8%	$0.25	12.3%	$0.75	42.8%	$0.78	43.8%
Sunkist Orange	1.46	$2.39	$2.18	9.1%	$0.93	39.0%	$1.18	54.1%	$0.46	31.5%
Schwepps Ginger Ale	2.06	$3.07	$2.34	23.7%	$1.01	32.9%	$1.34	57.3%	$1.06	51.5%
Canada Dry Ginger Ale	2.36	$3.47	$2.34	32.4%	$1.11	32.0%	$1.34	57.3%	$1.36	57.6%
A&W Rootbeer SF	1.72	$2.05	$1.73	15.3%	$0.33	16.0%	$0.73	42.3%	$0.72	41.9%
Diet Coke	2.25	$2.57	$1.81	29.6%	$0.32	12.3%	$0.81	44.6%	$1.25	55.6%
Column Medians	2.20	$2.57	$1.94	26.2%	$0.46	18.3%	$0.94	48.4%	$1.20	54.5%
Column Minimum	1.46	$2.03	$1.73	9.1%	$0.25	12.3%	$0.73	42.3%	$0.46	31.5%
Column Maximum	2.36	$3.47	$2.34	32.4%	$1.11	39.0%	$1.34	57.3%	$1.36	57.6%

* Assumes all SB mfr prices equal to $1.

Source: Barsky, Robert, Mark Bergen, Shantanu Dutta, Daniel Levy, "What Can the Price Gap between Branded and Private Label Products Tell Us about Markups?" Presented at the NBER Conference on Research in Income and Wealth:Scanner Data and Price Indexes,.September 15-16, 2000 [Revised: September 13, 2001], Table A-2, Only those brands with 300 or more observations were used. Available at: : papers.nber.org/papers/W8426

Column 1 shows the ratio of the invoice cost of the leading brand to the invoice cost of the store brand. I assume that the invoice cost is equal to the manufacturer price in each case. Barsky's et al. tables provide only ratios of prices. I have chosen to scale all prices shown in Tables 1 and 2, by taking the manufacturers' price for the store brands in each row to be $1. The table exhibits three regularities. First, the price of the SB is always lower than, i.e. discounted relative to, the price of the LB. The median "discount" from the LB retail price is 26%, with a high for Canada Dry Ginger Ale of more than 32% to a low of only 9% for Sunkist Orange (see column 4). Second, retail margins, both percentage and dollar, are higher on the SBs than on the LBs. Comparing column 6 to column 8 row by row, illustrates the first inverse association. For example, the percentage gross margin (%rgm) on Coke Classic is 13% versus 49% on the comparable store brand. The median %rgm for the leading national brands is 18% versus 48% for the store brands. The %rgms on SB soft drinks is generally more than twice as high as that on LB soft drinks. Moreover, the lowest %rgms (12% to 15%) are on the most popular brands and most heavily advertised brands, Pepsi and Coke. The highest %rgms (20%-39%) are on lesser known brands such Sunkist Orange, RC Cola and various Schweppes products. Similarly, comparing columns 5 and 7 row by row shows that dollar gross margins ($rgms) are higher on SBs than on LBs. For example, the $gm on Coke Classic is 35 cents versus 96 cents on the comparable SB. The median $gm on SBs is 94 cents versus 46 cents on the LBs. Note that the $rgms on SBs are considerably higher, despite the fact their retail prices are lower than the LBs they are designed to compete with (compare columns 2 &3). So both dollar and percentage retail gross margins are higher on soft drink store brands than on the leading brands they are designed to compete with.

Third, if we assume that the manufacturers' prices on SBs are a decent approximation to the marginal cost of the comparable LB, then Table 1 also provides evidence for the second inverse association.[51] Coke Classic, for example, would have a manufacturing margin of over 57% ($1.35/$2.35). Although we don't have any direct evidence on the manufacturing margin earned by the SB supplier, its marginal cost would have to be much less than half of Coke's to have as high a margin. Given that the SB supplier likely operates at a smaller scale than Coke, it seems very unlikely that its marginal costs are substantially lower than Coke's. Note also that the best known brands have the highest estimated % manufacturing margins (%mms). All the Pepsi and Coke products have %mms in excess of 55%, whereas RC Cola is around 47%, A&W Root Beer 42 – 44% and Sunkist Orange is again the lowest at 32%.

[51] This entails two separate and certainly controvertible propositions. First, that the quality of the LB and SB are essentially equivalent and second, that the manufacturer price of the SB, which presumably at least covers fixed costs, is a decent approximation of the marginal cost of the LB. On these questions see Barsky et al., *supra,* at pp. 12 – 24. They conclude "In summary, we believe there is enough evidence to suggest that using private label product prices to infer national brand costs is a reasonable assumption in this industry. There is reason to believe, therefore, that this measure of markup can be appropriate for at least some categories and products in this industry. Further, since the private label will have some markup, and the nationally branded products have advantages on size and scale in production, packaging and negotiation on input prices, we believe that private label product prices provide a conservative measure of these costs."

Table 2 shows that the same three regularities appear in 17 of 18 categories analyzed by Barsky et al.[52] The median SB discount from the LB is 25% (column 4). Except for frozen entrees, both %rgms and $rgms are higher on SBs than on LBs, as Steiner's first inverse association suggests. Third, column 1 is arrayed in increasing order of the ratio of the LB factory price to the SB factory price. If we assume that the SB factory price is an approximation of the marginal cost of the LB product, then LB manufacturers %mms range from a low of 23% for detergents to a high of 157% for crackers.
Columns 9 and 10 show dollar and unit shares of SBs in 17 categories based on national data gathered by a trade association. Though not necessarily representative of actual SB shares at Dominick's, they do allow us to get some idea of the unit sales weighted average %rgm in each of the 17 categories (column 11). These range from about 3% for detergents to almost 33% for canned soup. As seen in column 11, sales weighted %rgms differ by an order of magnitude across different categories. Not surprisingly, the range for LBs margins is similar to the weighted averages shown in column 11. LB %rgms range from a low of 3% for laundry detergents to a high of 32% for canned soups. The median is around 12% for cookies and crackers. For SBs, %rgms range from a low of 7% for laundry detergents to a high of 44% for soft drinks. For comparison, the overall %rgm for Dominick's in 1993 was about 22%.[53]

SB unit shares (column 10) range from a low of less than 1% in frozen entrees to a high of 38% in frozen juices, with a median penetration of almost 13%. There is no obvious relation between SB penetration and the category %rgm. For example, frozen juices which have the highest SB penetration also have the second highest %rgm at almost 29%. On the other hand, cheese, analgesics and fabric softeners, the only other categories to have more than 20% SB penetration, all have relatively weighted %rgms at 16%, 8.1% and 8.3% respectively. The lowest category %rgm is for laundry detergents, followed by dish detergents and cereals. All three categories show relatively low SB penetration (6 to 9%) and relatively low ration of LB manufacturer prices to SB manufacturer prices.

[52] Barsky et al. report on 19 categories in their paper. I have omitted one of these (toothbrushes) because the margins, especially on the SBs, are much higher than in any other category. Both inverse associations are very strong in this category, so its inclusion would strengthen the argument made here.
[53] See Dominick's Supermarkets Inc. SEC 10-K filing for 1996, p. 17.

Table 2
Manufacturer and Retail Prices for Leading National Brands and Store Brands for Eighteen Categories from Dominick's Supermarkets, 1989-1997: Store Brand Shares for 17 Categories, 1992-1996, Nationwide

	Ratio LB to SB Manuf. Prices	LB Retail Price	SB Retail Price	% Dis-count from LB Retail Price	LB Dollar Gross Margin	LB % Gross Margin	SB Dollar Gross Margin	SB % Gross Margin	SB $ Share of Cat-egory	SB Unit Share of Cat-egory	Cat-egory %Gross Margin
	1	2	3	4	5	6	7	8	9	10	11
Category											
Laundry Detergents	1.23	$1.27	$1.07	15.3%	$0.0369	2.9%	$0.07	6.9%	3.6%	5.8%	3.1%
Canned Tuna	1.23	$1.41	$1.24	12.3%	$0.18	13.0%	$0.24	19.4%	9.8%	12.8%	13.9%
Frozen Entrees	1.27	$1.49	$1.14	23.1%	$0.22	14.5%	$0.14	12.5%	0.8%	0.9%	14.5%
Bottled Juice	1.38	$1.57	$1.25	20.6%	$0.19	12.3%	$0.25	19.9%	15.0%	16.7%	13.6%
Cheese	1.38	$1.59	$1.27	20.0%	$0.21	13.0%	$0.27	21.2%	27.4%	32.5%	15.7%
Dish Detergents	1.42	$1.52	$1.13	25.4%	$0.10	6.5%	$0.13	11.8%	5.7%	7.4%	6.9%
Cereals	1.43	$1.53	$1.21	20.6%	$0.10	6.5%	$0.21	17.7%	6.2%	9.0%	7.5%
Frozen Juice	1.44	$1.89	$1.60	15.3%	$0.45	23.7%	$0.60	37.4%	31.5%	38.1%	28.9%
Oatmeal	1.52	$1.69	$1.41	16.7%	$0.17	9.9%	$0.41	28.9%	13.7%	19.1%	13.5%
Canned Soup	1.61	$2.35	$1.75	25.4%	$0.74	31.5%	$0.75	43.0%	5.8%	9.2%	32.6%
Tooth Paste	1.61	$1.77	$1.30	26.5%	$0.16	9.1%	$0.30	23.2%	1.5%	2.3%	9.4%
Snack Crackers	1.68	$1.88	$1.38	26.5%	$0.20	10.7%	$0.38	27.7%	na	na	na
Fabric Softeners	1.73	$1.83	$1.21	34.2%	$0.10	5.7%	$0.21	17.1%	16.2%	21.7%	8.1%
Cookies	1.77	$2.00	$1.34	32.9%	$0.23	11.5%	$0.34	25.5%	11.5%	13.8%	13.4%
Analgesics	2.01	$2.09	$1.28	38.7%	$0.08	3.8%	$0.28	22.0%	17.9%	24.6%	8.3%
Soft Drinks	2.09	$2.47	$1.79	27.5%	$0.38	15.3%	$0.79	44.0%	7.7%	18.1%	20.5%
Grooming (Razors)	2.17	$2.32	$1.33	42.9%	$0.15	6.5%	$0.33	24.6%	6.5%	10.8%	8.5%
Crackers	2.57	$2.96	$1.48	50.0%	$0.39	13.0%	$0.48	32.3%	8.0%	12.4%	15.4%
Col. Min	1.23	$1.27	$1.07	12.3%	$0.04	2.91%	$0.07	6.86%	0.8%	0.9%	3.1%
Col. Max	2.57	$2.96	$1.79	50.0%	$0.74	31.51%	$0.79	44.04%	31.5%	38.1%	32.6%
Col. Median	1.57	$1.80	$1.29	25.4%	$0.19	11.11%	$0.29	22.62%	8.0%	12.8%	13.5%
Col. Average	1.64	$1.87	$1.34	26.3%	$0.23	11.65%	$0.34	24.18%	11.1%	15.0%	13.8%

Source: Barsky, Robert, Mark Bergen, Shantanu Dutta, Daniel Levy, "What Can the Price Gap between Branded and Private Label Products Tell Us about Markups?" Presented at the NBER Conference on Research in Income and Wealth:Scanner Data and Price Indexes,.September 15-16, 2000 [Revised: September 13, 2001], Available at: : papers.nber.org/papers/W8426. Store Brand Shares from PLMA and IRI Data, courtesy of PLMA. Prices in columns 2 & 3 have been scaled from ratios by assuming that the manufacturer price for the store brnad is $1 per unit in each category. Columns 9 & 10 from the Private Label Manufacturers Association.

Steiner's Explanation for the Inverse Associations

Because of advertising, consumers recognize that the LB at one store is identical to the LB at another store. In contrast, consumers do not know if a store brand at one store is identical, except in name, to a store brand at a different store. Consumers are more likely to choose a store to shop on the basis of the availability and price of the LB, than on the basis of the availability and price of a given store's SB.[54] Thus, the elasticity of demand, *as seen by a retailer*, will be higher on an LB than on fringe and store brands. To maximize profits, retailers will set retail prices using the so-called "Lerner Rule" so that the %rgm on each product will equal the inverse of the elasticity of demand for that product as perceived by the retailer. Thus, the first inverse association is explained. The second follows from the assumption that successful national advertising leads to a lower elasticity of demand *as seen by an LB manufacturer* relative to an SB manufacturer.[55] Now the Lerner Rule implies that each manufacturer will set his factory price so that the %mm will be equal to the inverse of the relevant elasticity of demand. Thus, the second inverse association is explained. Neat. Economical. But the above does not constitute a formal model and leaves some important questions unanswered. The informal account above and Steiner's diagrammatic analysis treats LB and their corresponding SBs as though they are independent of one another, retailers price each product as though they were only selling one product rather than choosing prices to maximize profits from the category. The necessity for retailers to cover their fixed costs is not explicitly considered, so it is not clear how retailers can survive a substantial cut in margins on LBs. Although in virtually every historical account of actual margins, Steiner discusses the often important role played by economies of scale[56], they play little or no role in his diagrammatic treatment of his theory. Nor is it clear, under what conditions, if any, a profit maximizing LB manufacturer would set his post-advertising factory price at a level that would produce a post-advertising LB retail price below the original pre-advertising price of the product.

Toward A Satisfactory Formal Model of Steiner's Theory

In the absence of a formal model, economists have legitimate concerns over whether the propositions are consistent with each other and with profit maximization on the part of manufacturers and retailers and this is undoubtedly another reason why Steiner's work has not received the attention it merits. My own 1986 attempt to provide a formal model of Steiner's theory was I think a step in the right direction, but suffered two serious

[54] In my 1986 paper I added that consumers might choose which store to shop at by comparing different stores' LB prices, even if they intended to buy few, if any, LBs. See Lynch, "The Steiner *Effect: A Prediction from a Monopolistically Competitive Model Inconsistent With Any Combination of Pure Monopoly or Competition*," Working Paper No. 141, Bureau of Economics, Federal Trade Commission, August 1986, p.4.

[55] I omit two important elements of Steiner's explanation of an LB manufacturer's demand curve, the retail penetration of the brand and the retail dealer support of the brand. The omission is because I have not succeeded in modeling them, not because I think they're unimportant.

[56] See, e.g., Marketing Productivity in Consumer Goods Industries – a Vertical Perspective, *Journal of Marketing*, vol. 42, #1 (1978).

flaws.[57] First, it failed to formally impose the zero profit condition in the model, so it was not actually an example of a monopolistically competitive model. Second, and more serious, it was limited to single product retailers and so could not begin to capture the relationship between LB and SB retail prices. I now have succeeded in producing a model that overcomes both these deficiencies and that is complex enough to capture some of Steiner's theory, yet almost simple enough to solve analytically and the one part that can't be solved analytically can be simulated on a spreadsheet.

Monopolistic Competition: Not Even Wrong?

Before describing my attempt to model Steiner's ideas, I need to deal (briefly!) with a more general issue as to the type of model selected and what purposes I expect a model to serve. I have argued strongly against using the PC theory to model retailing. But what theory should be used? Given what I have said about the unusual ease of entry into retailing and the vast number of retailers at many scales, the theory of monopolistic competition (MC) seems to be many ways the simplest and most appropriate theory to use and indeed my model is of that variety.[58] Yet Milton Friedman's criticism of such models was and is very strong. In essence he was saying that although such models seem to make more "realistic" assumptions than the PC model, they are so general that any real world outcome is consistent with some version of a monopolistically competitive model. In the indignant words of the great physicist, Wolfgang Pauli, such a theory is "not even wrong."[59] John Sutton writes,

> The low point in the fortunes of imperfect competition was reached in the 1961 Review of Economic Studies, in which Archibald replied to two of his critics, Milton Friedman and George Stigler…Archibald showed that no non-trivial comparative static results held good, as to the effects of changing demand parameters, or costs (via taxes, etc.); nor did it help matters to confine oneself to special cases of the model obtained by assuming 'no advertizing' or 'no quality choice'. Archibald's paper was widely accepted as the definitive statement of the emptiness of imperfect competition. Archibald's own conclusion, however, was positive, and

[57] See *supra* note 51. I should also mention Lal and Narasimham's sophisticated model presented in **"**The Inverse Relationship Between Manufacturer and Retailer Margins: A Theory,**"** Management Science, Vol. 15, No. 2, 1996. The authors reference Steiner's work and may seem to be providing a model of Steiner's theory. If that was their intention, then I don't think they have succeeded, though their model is interesting in its own right. For example, although manufacturer advertising leads to lower retail margins, in their model consumers gain nothing from the lower retail margins, because the "focal" good manufacturer raises the factory price by exactly the amount the retail margin falls. Their model rules out the "Strong Steiner Effect" by its assumptions.

[58] The model employs a Nash-Cournot equilibrium using "Bertrand" prices.

[59] It is not only economists that are at times enamored by theories "not even wrong." "String theory not only makes no predictions about physical phenomena at experimentally accessible energies, it makes no precise predictions whatsoever**….** There is, however, one physical prediction that string theory does make: the value of a quantity called the cosmological constant (a measure of the energy of the vacuum). Recent observations of distant supernovae indicate that this quantity is very small but not zero. A simple argument in string theory indicates that the cosmological constant should be at least around 55 orders of magnitude larger than the observed value. This is perhaps the most incorrect experimental prediction ever made by any physical theory that anyone has taken seriously." See Peter Woit,**"**Is String Theory Even Wrong?" American Scientist, Volume 90, No. 2 March-April 2002, p.110 at ***http://www.americanscientist.org/***

> indeed appeared a very natural response. What he argued was that only by putting more structure into the theory could one hope to get predictions out. The direction which he favored …involved specifying consumer tastes over potential product varieties and endogenously determining the (number and possibly type of) products on offer, and in this way pinning down within the model demand conditions facing each firm at equilibrium. [60]

The Steiner Effect is inconsistent with PC theory and consistent with MC theory, but the list of false implications I have given above suggests that is not hard to find PC implications that are inconsistent with the facts of retailing. So if the virtue of the PC model is that it does generate refutable predictions, then its vice, at least for retailing, is that its predictions are so often refuted. However, if an MC model makes no refutable predictions, then I think such a model is worse than wrong. The model I have constructed follows Archibald's path – I put in more structure – structure based at least in part on the "stylized" facts I'm trying to model. For example, the most consistent empirical regularity regarding LBs compared to SBs, is that the retail price of the LB is virtually always higher than the retail price of the SB designed to compete with it. So rather than starting with two general demand functions, I model the choice between the two for consumers in a manner that assumes that no consumer would buy a unit SB unless its price were below the price for the LB.

An Outline of My Model of Steiner's Dual StageTheory

Steiner emphasizes the importance of two choices made by consumers: which store to shop at and which product to buy (LB or SB) at the chosen store. The more likely consumers are to switch stores to get a particular brand rather than switch brands within a store, the greater the market power of the brand manufacturer. Thus to model Steiner's theory consumers must be able to choose between at least two stores and each store must carry at least two different products in a category. Thus there must be some competition between multi-product retailers.[61]

How Consumers Choose Where to Shop

Consumers must first choose where to shop. The key assumption is that they choose on the basis of LB prices only, though other nonprice factors, e.g. an individual consumer's distance from the store, also matter. The rationale is that consumers know that LBs are the same at different stores, whereas SBs may vary in quality. Consumers may also believe that comparative prices on LBs are a good guide to comparative prices of all the items carried by different stores. In any case, it is assumed that consumers know LB prices before choosing which store to visit. As

---

[60] Sutton, John (1989) "Is Imperfect Competition Empirically Empty," *The Economics of Imperfect Competition and Employment: Joan Robinson and Beyond,"* edited by George Feiwel, NYU Press, 225-240.

[61] The only work known to me that provides a theory of how consumers located at different points in space and who shop for multiple items is the brilliant book by Robert Bacon, ***Consumer Spatial Behavior: A Model of Purchasing Decisions over Space and Time***, Oxford, 1984. Models of shopping at single product retail stores can be found in Geofffry Heal,

shown previously,[62] if the consumer's choice between any two stores depends only on the ratio of the LB prices at those two stores (and a couple of other non-restrictive assumptions), then the share or fraction of the N consumers attracted by the ith store is given by a relatively simple function,

$$s_i = \frac{m_i p_i^{-\mu}}{\sum_1^R m_k p_k^{-\mu}}$$

for i =1, ..R where **R** is the number of retail stores in the market and μ is a parameter that measures the "visibility" of the LB product among stores, and m_i is a parameter measuring non-price factors affecting the attractiveness of store i to consumers. If all stores had the same prices, for example, then store *i*'s share of the retail market would m_i divided by the sum of all the ms. In general, the parameter m_i can reflect special conditions for retailer i (such as being within a more densely populated area than some other stores). I will later assume that the visibility parameter can be increased through advertising. Thus the "visibility" of the LB brand is closely related to the willingness of customers to switch retail stores in order to get the LB. I will later assume that N, the number of consumers in the market who wish items in the product category can be increased through manufacturer advertising.

How Consumers Choose Which Product to Buy at a Store

Consumers will buy at most one unit of either the LB or the SB product, but not both. Which brand a consumer buys depends on the retail prices of the brands at the chosen store and on the consumer's wiliness to pay a "reputation premium" for the LB. For example, suppose a consumer is willing to pay $1 more for the LB and the prices of the SB and LB are $1.50 and $2.25 respectively. The consumer would buy the SB in this case. I assume that consumers differ in the reputation premia they are willing to pay and that these reputation premia are uniformly distributed. The maximum premium any consumer is willing for the LB is designated r^*_{max}. For example, suppose the minimum premium is $0 and the maximum is $2 for the customers choosing a particular store. If the retailer sets the retail price of the SB at $1 below the LB, then half of the customers would choose the SB and the other half the LB. I will later assume that reputation premia can be affected by manufacturer advertising. These assumptions imply that the quantities sold at the at the ith store will be,

$$(1)\quad q_{LBi} = s_i N\left[\frac{r_{\max} - (p_{LBi} - p_{SBi})}{r_{\max}}\right]$$

and,

$$(2)\quad q_{SBi} = s_i N\left[\frac{p_{LBi} - p_{SBi}}{r_{\max}}\right]$$

[62] See Lynch, *supra,* footnote 2.

How Retailers Choose Prices for the Product s They Offer. I assume that retailers choose prices to maximize their profits from the category. The marginal cost to the retailer of selling a unit of the LB or the SB is simply the price m_{LB} or m_{SB} paid to the manufacturer per unit. Each retailer is assumed to have has fixed costs, f, that need to be covered if he is to stay in business. Gross margin dollars from LB sales will be simply be the dollar retail gross margin per unit ($rgm) times the number of units sold and similarly for the SB. Retail profit is given by,

$$\pi_i = (p_{LBi} - m_{LB})q_{LBi} + (p_{SBi} - m_{SB})q_{SBi} - f = \$gm_{LBi}q_{LBi} + \$gm_{SBi}q_{SBi} - f_i$$

Setting the partial derivative of the above with respect to p_{SBi} equal to zero yields a surprisingly simple equation,

$$\frac{\partial \pi}{\partial p_{SBi}} = 0 \Rightarrow \qquad p^*_{SBi} = p_{LBi} - \left(\frac{m_{LB} - m_{SB}}{2}\right) = p_{LBi} - \overline{\Delta m}$$

The equation above not only exhibits the first inverse association, it also allows one to immediately calculate the profit-maximizing retail price of the SB, given the retail price of the LB and the two manufacturer prices. Table 3 below shows the predictions from this equation applied to the soft drink data contained in Table 1 along with the actual prices chosen by Dominick's. The predicted prices are surprisingly close to the actual prices. This may be just a lucky fluke, but it does demonstrate that the model is no "empirically empty."

A second equation for each retailer is obtained by setting the partial derivative of the retailer's profit function with respect to p_{LB} equal to zero. The resulting equation, unfortunately, is not simple and so I won't write it down here. A third equation for each retailer is obtained by assuming that free entry implies zero profits for each, or that total gross margin dollars for the category for each retailer is just equal to his fixed costs. Next, in the hope of gaining simplicity and insight, I assume that symmetry (e.g. all retailers have the same fixed costs, the same ms etc.) implies that all retailers choose the same equilibrium prices. I now have three equations to determine the two retail prices and the number of retailers (R^*) that insures zero profits. Solving this system leads to a quadratic equation for p^*_{LB}. The expression obtained is again rather complex and I will omit it here.[63]

How Manufacturers Choose Their Factory Prices and Advertising Budget.
I assume there are many potential suppliers of the SB product and that the SB factory price is set by the condition that manufacturing gross margin dollars just equal the SB's fixed manufacturing cost. SB manufacturers don't advertise. I assume that the LB

[63] For a fuller, more technical account see "Toward a Formal Model of Steiner's Dual-Stage Theory of Manufacturing/Retailing", forthcoming

manufacturer chooses his factory price to maximize his profits. The LB manufacturer can infer, for any given SB factory price, and given retailer fixed costs, the quantity he will sell at any LB factory price. Setting the partial derivative of the manufacturer's profit function equal to zero, produces another fairly simple equation that determines his profit-maximizing price.

$$\frac{\partial \pi_{LB}}{\partial m_{LB}} = 0 \Rightarrow m^*_{LB} = r_{\max} + \left(\frac{m_{SB} + mc_{LB}}{2} \right)$$

Thus the manufacturer will set his factory price equal to the maximum reputation premium any consumer is willing to pay plus the average of the sum of his marginal cost and the SB factory price.

The optimal advertising budget is found by setting the partial derivate of π_{LB} equal to zero. This is complex, because I assume LB advertising can increase the number of consumers (N) who will buy in the category and the maximum reputation premium any consumer reached is willing to pay for the LB, and the visibility parameter, μ. In all cases I assume that it is advertising expenditures (A) raised to some power less than 1 (i.e., diminishing returns to advertising) that determines N, $r^*_{\max}$ and μ. This is the point at which I can no longer obtain an analytical solution. Instead, I have built a spreadsheet model, incorporating all the inequality constraints, and I then use Excel's "Solver" function to obtain results for specified parameters.

Implication #1: Retailers Will Set Higher Dollar and Percentage Gross Margins on SBs Than on LBs

In the model, it is very clear why profit-maximizing retailers will always set higher $rgms on SBs. Take the Coke Classic and its competing store brand as shown in the first row of Table 1 as an example. For every unit of Classic Coke sold, the retailer keeps 35 cents as a contribution toward covering fixed cost including profits. Suppose, contrary to fact, he priced the SB at $1.35 instead of the actual price of $1.96. Then the $rgms would be same on both brands. The retailer would presumably sell more units of his store brand, but he would not increase his category gross margin dollars by one penny. Why? Because every additional unit of the SB sold means one fewer unit of the LB sold. If the retailer priced the SB at $1.34, he would actually reduce category profits despite higher SB unit sales. If he set the SB price at $1.36, however, he would increase category profits so long as SB sales didn't fall to zero. At the higher price, some SB customers will switch to Classic Coke, but those who continue to choose the SB will yield a net increase of one penny per unit sold to the gross margin dollars earned in the category. It is the ***difference*** between the $rgms that counts. There is a limit to how much the retailer can raise the price of the SB. In the model, it is assumed that at an SB price of $1.96, the SB share of unit sales would fall to zero. So the model says that if the SB is carried at all, a profit maximizing retailer will always put a higher dollar margin on the SB than on the LB. Since the retail price of the SB must always be lower than the retail price of the LB, it follows that the %rgm on the SB will always be higher than on the LB. Thus the first inverse association is explained.

The model, however, does more than just confirm the first inverse association; it determines the exact amount by which the SB must be discounted from the LB. Because of the reputation premium Coke Classic commands, the SB must always sell at a discount to it. In the model, how much of a discount depends on the maximum reputation premium any customer would pay. The actual discount chosen by Dominick's is shown in Table 1 to be 84 cents per unit. Under the assumptions given above, the model provides a surprisingly simple formula for calculating the category profit maximizing price for the SB, given the retail price for the LB. The model implies that the optimal SB price is equal to the LB retail price less the average difference between the manufacturer prices. For given manufacturer prices (here $2.35 and $1) and a given Classic Coke retail price of $2.70, the profit maximizing retail price for the SB will be $2.03. The actual price chosen by Dominick's was $1.96. Table 3 shows the predicted versus the actual prices for all the soft drink brands shown in Table 1.

Table 3

Predicted versus Actual Store Brand Soft Drink Prices, Dominick's Data, 1987 - 1997

Leading Brand	Predicted	Actual	Pred. less Actual	%Error
Coke Classic	$2.03	$1.96	$0.07	3.41%
Schwepps Ginger Ale	$1.91	$1.92	-$0.01	-0.48%
Pepsi Cola N/R	$2.02	$1.91	$0.11	5.43%
Pepsi Cola Diet N/R	$1.94	$1.82	$0.12	6.23%
Barq's Root Beer	$2.02	$1.81	$0.21	10.25%
Schweppes Tonic N/R	$2.48	$2.30	$0.19	7.52%
Schweppes Diet Tonic N/R	$2.49	$2.32	$0.17	6.76%
R.C. Cola	$2.04	$2.09	-$0.04	-2.14%
A&W Rootbeer Reg	$1.64	$1.75	-$0.11	-6.72%
Sunkist Orange	$2.16	$2.18	-$0.01	-0.57%
Schwepps Ginger Ale	$2.54	$2.34	$0.20	7.73%
Canada Dry Ginger Ale	$2.79	$2.34	$0.45	15.96%
A&W Rootbeer SF	$1.69	$1.73	-$0.05	-2.83%
Diet Coke	$1.94	$1.81	$0.13	6.89%
Column Average	$2.12	$2.02	$0.10	4.10%
Column Median	$2.02	$1.94	$0.12	5.83%
Column Minimum	$1.64	$1.73	-$0.11	-6.72%
Column Maximum	$2.79	$2.34	$0.45	15.96%

Source: Barsky, Robert, Mark Bergen, Shantanu Dutta, Daniel Levy, "What Can the Price Gap between Branded and Private Label Products Tell Us about Markups?" Presented at the NBER Conference on Research in Income and Wealth:Scanner Data and Price Indexes,.September 15-16, 2000 [Revised: September 13, 2001], Available at: : papers.nber.org/papers/W8426.

The model suggests that it is the difference between the LB and SB factory prices that is crucial to whether a store carries a private label and, if so, what margin it will carry. The laundry detergent category, for example, has the lowest %rgm of any of the 18 categories

shown in Table 2. The difference between the LB and SB factory prices is the also the lowest of any of the categories. There is some connection between the two. The LB factory price is only 23 cents per unit above the SB factory price. But at the observed discount of 20 cents per unit, the $rgm on the SB can be only 3 cents above the $rgm on the LB. Now the $rgm on the LB will be determined primarily by inter-store competition. The visibility of a brand like Tide is very high and so the elasticity of demand as seen by the retailer for LBs like Tide is high as evidenced by the low 4 cent unit dollar margin and low 3% on these LB detergents. The high inter-store elasticity dictates a low retail margin on the LB and the relatively low LB/SB factory price ratio combined with a hefty discount necessary because of the LB reputation premium forces both LB and SB margins to be relatively low.

In contrast, the canned soup category has the highest %rgm of the 18 categories shown in Table 2, yet the ratio of LB to SB factory prices 1.61, which is right at the median value for this ratio. Despite the presence of such famous brands as Campbell's, the % and $ rgms on the LB soups are quite high at 25% and 74 cents respectively. The SB soups are sold at a discount of over 25% or 60 cents per unit. This implies that the $rgm on the SB soups can only be a razor thin one cent above the margin on the LB soups at 75 cents. So in the soup category, both the LBs and the SBs have high %rgms at 25 and 43% respectively.

Toothpaste has the same LB to SB factory price ratio as canned soup, roughly the same percentage discount at 26%, yet the category has a relatively low overall %rgm of about 9%. The main reason is that the low %rgm on the LB toothpastes, presumably due to high inter-store elasticities, leaves little room for high SB margins and the small unit share of the SBs means their relatively high %rgm don't have much overall effect on the category.

The frozen juice category is interesting because it has the highest unit share for SBs at more than 38%. Despite the high generics share, both LBs and SBs support relatively high %rgms and the category margin is the second highest. This data raises some question as to how to define Steiner's "mixed regimen" and under what conditions it is better for consumers.

I conclude that both the %rgms of the LBs in a category and the ratio of LB to SB factory prices are important determinants of overall category margins. If both are relatively low, then category margins will be low. If both are high, then category margins will be higher than the storewide average. Mixed cases can produce either relatively high or low category margins. In the Dominick's database, LB factory prices are commonly more than 50% higher than their generic rivals and this is often enough to provide relatively high %rgms on SBs. Moreover, some LBs have high retail markups, despite their well known names. In these cases, category margins may well be higher than the store-wide average.

Implication #2: Profit Maximizing Retailers and Manufacturers Produce the Steiner Effect and the Strong Steiner Effect Especially If Advertising Leads to significantly higher retailer elasticities, and large Increases in Unit Sales and both Manufacturers and Retailers Have Economies of Scale.

As illustrated in Charts 1 and 2, simulations show that the model can reproduce the type of results that actually occurred when Steiner first advertised the Kenner Girder and Panel sets on television. The strongest requirements to achieve this happy result are that advertising lead to substantially higher retailer elasticities, to large increases in unit sales and that there are both manufacturing and retailing economies of scale.

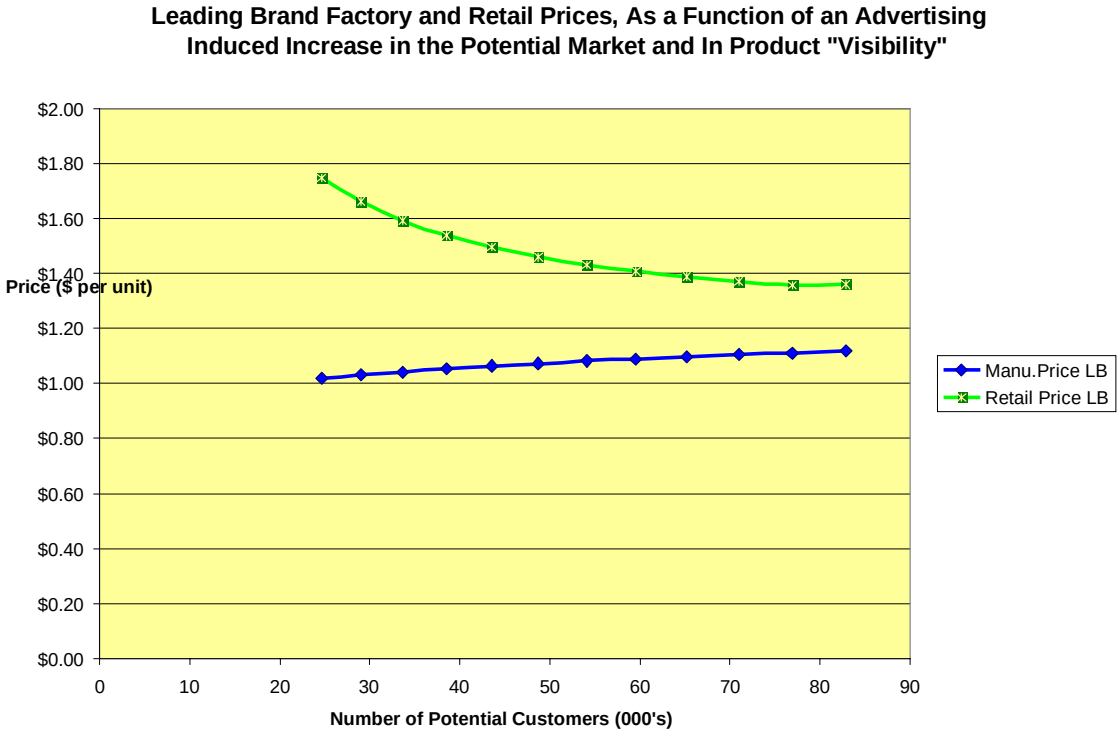


Chart 1

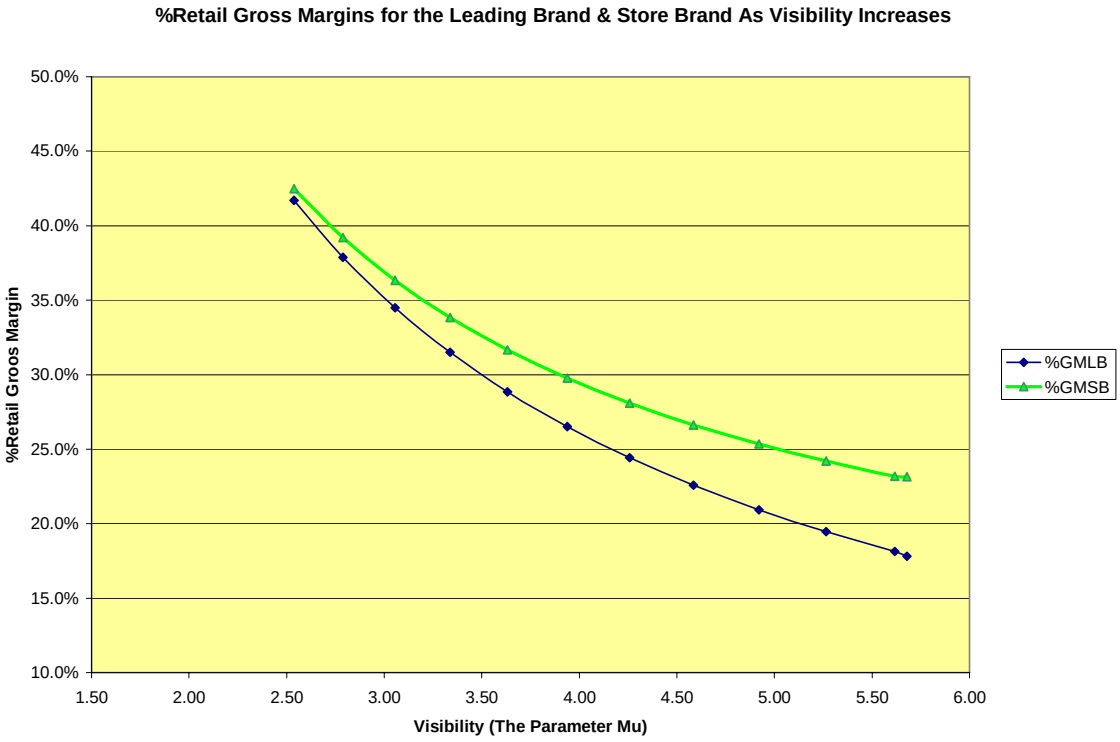

%Retail Gross Margins for the Leading Brand & Store Brand As Visibility Increases
50.0%
45.0%
40.0%
35.0%
30.0%
25.0%
20.0%
15.0%
10.0%
%Retail Groos Margin
1.50
2.00
2.50
3.00
3.50
4.00
4.50
5.00
5.50
6.00
Visibility (The Parameter Mu)
%GMLB
%GMSB

Chart 2

Implication #3: Pass-Through Rates of Uniform Upstream Price Increases Will Always Be Greater Than 100%

To my surprise, the model implies that general upstream cost increases will always be passed -through to consumers at a rate of more than 100%. Suppose, for example, that the cost of production rises by 10 cents per unit for both the SB and the LB manufacturer. The SB factory price will rise by 10 cents. Equation (5) implies that the LB factory price will also rise by 10 cents. Given that the difference in the factory prices is unchanged, equation (3) implies that the difference between the two retail prices will be unchanged. Equation (4) now implies that the LB factory price will be marked-up by a factor of $\mu/(\mu-1)$. The SB retail price will rise by exactly the same dollar amount as the LB retail price, since the differential between the two remains unchanged. Thus the pass-through rate for a uniform upstream cost increase will exceed 100% by one divided by $\mu - 1$. For example, if $\mu = 3$, then 150% of the upstream cost increase will be passed through to consumers. Although not exact, simulations suggest that category percentage margins will stay nearly constant through uniform upstream cost increases. I did not set out to build this property into the model, rather it is entailed by the other assumptions. I think this is a nice example of how formal model can reveal connections not known by the model builder.

Broader View of Steiner's Theory

The model sketched above applies to one category, whereas real retailers carry hundreds of categories. Although there are some new complications to be faced when extending the model to many categories, it does suggest the following broad view of how retail and manufacturing margins are determined. Each category contains at least one LB. The higher the visibility of the LB, the lower the LBs retail margin. LBs in different categories have different visibilities and this account, in part, for differences among category margins. Whether a low margin on an LB, leads to low category margins depends on ratio of the LB and SB factory prices and reputation discount that SB must offer to compete with the relevant LB

There is little one can say of a general nature concerning the overall efficiency of the equilibria generated by this model. I am afraid that one of the prices to be paid in moving to more realistic model recognizing economies of scale is the absence of any sweeping statements saying X is clearly the most efficient arrangement. The best we can do, and it is difficult enough, is to analyze specific situations in hopes of finding more or less efficient modes of organization.