Supermarkets as a Natural Oligopoly*

Paul B. Ellickson
Duke University

September 20, 2004

Abstract

This paper uses a model of endogenous sunk cost (ESC) competition to explain the industrial structure of the supermarket industry, where a few powerful chains provide high quality products at low prices. The predictions of this model accord well with the features of the supermarket industry documented here. Using a novel dataset of store level observations, I demonstrate that 1) the same number of high quality firms enter markets of varying sizes and compete side by side for the same consumers and 2) quality increases with the size of the market. In addition to documenting a local structure of competition consistent with the ESC framework, I demonstrate that the choice of quality by rival firms behaves as a strategic complement. This key finding, which is consistent with an ESC model of quality enhancing sunk outlays, eliminates several alternative explanations of concentration in the supermarket industry, including most standard models of cost-reducing investment and product proliferation. These results suggest that the competitive mechanisms sustaining high levels of concentration in the supermarket industry are inherently rivalrous and unlikely to lead to the emergence of a single dominant firm.

Keywords: endogenous sunk costs, vertical product differentiation, oligopoly, retail, supermarkets, market concentration, dartboard, complementarity.

JEL Nos: L13, L22, L81

*This paper is an updated version of Chapters 3 and 4 of my MIT dissertation. My advisors, Susan Athey, Scott Stern, and Richard Schmalensee provided invaluable guidance and advice. I also thank Elizabeth M. Bailey, Bryan Ellickson, Howard Smith, and seminar participants at Berkeley, Chicago, Harvard, Stanford, UCLA, Yale, and the NBER Summer Institute for many useful comments. All remaining errors and omissions are my own. All correspondence to: Paul B. Ellickson, Duke University Economics Department, 305 Social Sciences, Durham, NC, 27708. Email: paul.ellickson@duke.edu.

1 Introduction

This paper proposes and tests a model of the organization of the supermarket industry. Using a data set that includes detailed information on every store, I demonstrate that the structure of the supermarket industry is remarkably uniform throughout the United States. Because the advantage of one supermarket chain relative to another largely reflects efficiency in distribution, the uniformity in industrial structure manifests itself most clearly in regions larger than a single metropolitan area, roughly the scale of a mid-sized state. The industrial organization of each of these 51 spatially defined markets is a natural oligopoly in which a small number of firms (between 3 and 6) capture the majority of sales, independent of the population served by the market. The number of firms does scale up with the size of the market (as measured by population or revenue), but the expansion is limited to a fringe of firms operating stores of much lower quality than those of the oligopolists.

The theoretical explanation for why we see a natural oligopoly among supermarkets is based on John Sutton's (1991) theory of endogenous sunk costs. As markets grow, local rivalry drives firms to expand their sunk investments, limiting the number of firms that can profitably enter even the largest markets. Although the number of oligopolists does not change, the quality of the service they provide expands with market size: the expenditure on sunk costs does buy something. It is this competition to provide higher quality service that determines market structure. While other theories might lead us to expect that firms would monopolize certain regions or find other ways to isolate themselves along "horizontal" dimensions of product space, this is not what happens in practice. Instead I find that the oligopolistic firms compete head to head for their consumers at the local level, responding to quality increases by nearby rivals with increases of their own. These results are consistent with the vertical model of endogenous investment proposed in this paper, but sharply contradict alternative explanations of industrial structure.

I begin by presenting an endogenous sunk cost (ESC) model of retail competition, based on Sutton (1991). In this model, supermarkets compete for customers by offering a greater variety of products in every store. Variety is increased by investing in proprietary firm level distribution systems, a sunk investment. Because variety is a purely vertical form of product differentiation, firms that fail to match the quality increases of their rivals cannot survive. Therefore, as markets grow, existing firms must incur higher costs if they are to remain in the industry, and this escalation in costs discourages entry by other firms. Consequently, markets both large and small are served by roughly the same small number of high quality firms.

However, this simple version of the Sutton theory doesn't fit the facts exactly. Like many retail industries, the supermarket industry includes a fringe of small firms in addition to the dominant chains, and the size of the fringe does grow with market size. Fortunately, as Sutton also realized, this complication is easily accommodated by allowing for a competitive fringe of firms that do not compete in quality. For these firms, sunk costs are not endogenous, and consequently their number expands with the size of the market. The existence of this fringe brings a useful side benefit: a natural control group of exogenous sunk cost firms to contrast with the oligopolists that invest endogenously.

I evaluate the implications of the ESC model using a census of store level observations. Focusing first on the 51 regional markets described above, I substantiate the claim of natural oligopoly: roughly the same number of high quality firms enter each market, regardless of size, and the quality of their stores expands with the size of the market. Furthermore, this escalation of quality is exhibited only by firms investing in distribution, supporting the claim that store level quality is linked to firm level investment.

To demonstrate how the ESC framework can be further distinguished from alternative models, I then shift the focus to store level competition. Natural oligopoly at the regional level has little meaning for individual consumers if firms monopolize local markets. Focusing on zip-code level markets, I find that the dominant firms do not carve out local monopolies. Supermarkets neither cluster by owner nor serve distinct niches. Instead, each firm faces high quality rivals in the vast majority of local markets in which they compete, as a proper interpretation of the ESC model should require. This rivalrous nature of local competition provides a sharp contrast with alternative models of competition. In particular, to support persistent concentration in a horizontal setting, firms could soften price competition by monopolizing continuous regions of product space as in Schmalensee (1978). Alternatively, they might dampen competition by isolating themselves in product space. However, supermarkets do not behave this way. This is not to suggest that horizontal differentiation plays no role in competition, but simply that models that rely solely on the horizontal dimension cannot account for features of retail markets that the ESC model easily captures.

Finally, I develop and test another even more restrictive implication of the ESC framework that sets it apart from rival theories. Most models of strategic investment imply strategic substitution between rival firms (Bagwell and Staiger, 1994; Athey and Schmutzler, 2001). I show that in an ESC model, investments may be either substitutes or complements. In contrast, strategic complementarity is inconsistent with most standard models of capacity competition, horizontal product differentiation, cost-reducing investment, and product proliferation. Focusing once again on store level investments, I find that quality is indeed a strategic complement, providing the final and most restrictive test of the ESC framework.

Although this is the first study to use an ESC model to explain competition between retail firms, subsequent authors have adopted a similar approach in a number of settings. Dick (2003) applies the methodology developed here (and in Ellickson (2001)) to the banking industry. She finds evidence of endogenous investment in branch quality among large regional banks. Berry and Waldfogel (2003) examine the newspaper and restaurant industries and Latcovich and Smith (2001) study online book sellers.

The paper is organized as follows. Section 2 provides a brief history of the supermarket industry, arguing that endogenous investments have played a central role throughout its evolution. Section 3 presents a formal model of supermarket competition. After deriving the natural oligopoly and escalation results, I identify the conditions that yield strategic complementarity in investments. Section 4 describes the dataset and demonstrates how distribution networks can be used to identify distinct geographic markets. The first set of empirical results are presented in Section 5. I first establish that these regional distribution markets are indeed served by roughly the same number of dominant firms. Next, I connect this oligopoly result to the ESC mechanism emphasized here by showing that quality expands with the size of the market. The store level analysis is presented in Section 6. After demonstrating that the dominant firms compete head to head, I find that quality is a strategic complement. Section 7 concludes.

2 The Evolution of the Supermarket Industry

The evolution of the supermarket industry is marked by three major innovations: the rise of chain grocery stores in the early 1900's, the introduction of the supermarket format circa 1950, and the adoption of automated distribution and procurement systems in the 1980's and 1990's. Although each innovation involved substantial sunk investments, the focus of those investments has shifted over time. While the earliest innovations were mostly aimed at reducing costs, more recent investments in technology favor increasing quality.

The rise of the chain stores, led by the Atlantic and Pacific Tea Company (A&P), was primarily driven by the efficiencies of integrating backward into both wholesaling and manufacturing. The physical stores these firms operated were essentially identical to those of the non-chains (Adelman, 1958). However, by manufacturing their own products and distributing them through their own vertical networks, the chain stores eliminated several layers of middlemen, substantially lowering per-unit costs. These cost savings were passed on to consumers through lower prices.[1] Not surprisingly, the escalating levels of investment quickly led to a shakeout, as the smaller independent grocers could no longer com-

[1] Several price studies performed in the late 1920s and early 1930s found that chain store prices were 4.5-14% lower than their independent counterparts (Tedlow, 1990).

pete. As the chain store format spread, concentration rose sharply. Between 1919 and 1932, the share of the top 5 firms in the U.S. increased from 4.2% to 28.8% of total grocery sales (Tedlow, 1990), with A&P alone accounting for 16.3%.

The introduction of the supermarket format solidified the competitive advantage of the chain stores by introducing a second, store level component of sunk investment. Supermarkets were 5 times larger than a typical grocery store, carried far more products, and required consumers to serve themselves. Because of their unique, single-floor, space-intensive design, supermarkets are not easily converted into other uses. Located in suburbs to economize on land prices, these new stores sold nationally branded goods and advertised heavily, adding yet another layer of firm level investment (Tedlow, 1990). Significantly, supermarkets now competed on the basis of both variety *and* price, linking stores to firms through advertising and distribution. By the late 1960s, supermarkets had become the dominant format, accounting for the vast majority of retail food sales. Grocery stores continue to exist only as a competitive fringe.

Although advertising helped drive the diffusion of the supermarket format in the 1950s, its importance has declined in recent decades, being replaced by the extensive investment in distribution and information technology pioneered by mass merchandisers like Wal-Mart and Target, who use advanced distribution software to carefully manage their inventories (Messinger and Narasimhan, 1995). Almost all the major supermarket firms invest[2] in proprietary information technology and logistical systems aimed at increasing variety while minimizing storage and transportation costs. Although there are parallels to the cost-reducing investments of the 1920s, the modern focus is on managing and expanding the variety of products offered in each store, rather than simply lowering per-unit costs[3]. The following model is motivated by the importance of these investments.

3 An Endogenous Sunk Costs Model of Supermarkets

This section introduces a theoretical model of supermarket competition based on Sutton's (1991) endogenous sunk cost (ESC) framework. I emphasize two main implications of this model. First, supermarkets form a natural oligopoly in which the number of firms is largely independent of market size. Second, this oligopoly is sustained by escalating investments in quality enhancing sunk costs. These implications

[2] Unfortunately, specific data on investment intensities is difficult to find. However, according to industry sources at the Food Industry Center and the Food Marketing Institute, *total* distribution costs average about 4-6% of sales, with transportation and warehousing each accounting for about 2% (personal communication with author).

[3] The explosion in both product variety and store size over the last two decades is striking. The number of products offered per store increased from 14,145 in 1980 to 21,949 in 1994 (Messinger and Narasimhan, 1995). Over the same period, average store size increased by about 1,000 square feet per year (Progressive Grocer). The costs of increasing brandwidth are both fixed and sunk. The expenses associated with increasing store size and expanding automated logistics systems far outweigh the burden of hiring extra employees to stock the shelves. Dedicated distribution centers are not easily converted to alternative uses and proprietary IT systems are not readily transferred to other firms.

will be tested in Section 5. I then turn to the nature of strategic interaction. After establishing that Sutton's model implies that competing investments are strategic substitutes, I introduce an alternative specification in which these investments are complements. This more restrictive implication is evaluated in Section 6.

3.1 A Vertical Model of Competition

In this model of retail competition, supermarket chains are vertically differentiated, differing only in their level of quality z, which represents the brandwidth or variety provided in each of their stores. On the demand side, I assume that a wider choice set, prices held fixed, appeals to all consumers, allowing supermarkets to draw from a broader customer base. Utility is given by

$$u(x_1, x_2, z) = (1-\alpha)\ln(x_1) + \alpha \ln(z x_2) \tag{1}$$

defined over two goods, a Hicksian composite commodity x_1 and the quality differentiated good x_2 that is the focus of our analysis. Each of M identical consumers is endowed with Y units of good 1, which is a numeraire ($p_1 = 1$). Therefore, ignoring any distribution of profits, each consumer has wealth Y. I let $p(z)$ denote the price of a differentiated good of quality z.

Focusing on chain level investment, I assume there are N identical firms, where firm j uses input $F(z_j) + cq_j$ of the composite good to produce quantity q_j of the differentiated good of quality z_j. The sunk cost of quality $F(z)$ represents a firm level investment in distribution technology. Competition is modeled as a three stage game. In the first stage, firms choose whether or not to enter and incur a sunk entry cost σ, assumed to reflect the minimum efficient scale of a small chain of minimal quality. In the second stage, firms choose a level of quality z, requiring sunk cost $F(z)$. In the third and final stage, firms compete in the product market, which is modeled as Cournot.[4] Using this basic framework, I will now illustrate both the exogenous and endogenous sunk cost cases.

3.1.1 The Exogenous Sunk Cost Case

In the exogenous sunk cost case, quality is fixed. Firms still pay the fixed cost of entry σ, determined exogenously by the industry's underlying technology. Without loss of generality, assume all stores offer quality $z_j = 1$ and let $p(1) = p$. Maximizing profit at store j and solving the resulting symmetric first

[4] It is important to emphasize that the natural oligopoly result does not depend on the assumption of Cournot competition: Shaked and Sutton (1983) derive a similar result under Bertrand. Cournot competition is assumed here both for ease of exposition and because it yields a symmetric equilibrium in quality, a feature which accords well with the specifics of retail competition. The stores operated by Circuit City and Best Buy, Wal-Mart and K-Mart, Staples and Office Depot, and the dominant supermarket chains are often difficult to distinguish and are frequently located in close proximity to their competitors. In contrast, static Bertrand models typically lead to asymmetric equilibria: firms either differentiate themselves in quality or geographic space to dampen the effect of price competition (Shaked and Sutton (1983), Ronnen (1991)).

order conditions yields equilibrium quantities and price

$$q = \left(\frac{N-1}{N^2}\right)\frac{\alpha YM}{c} \quad \& \quad p = \left(\frac{N}{N-1}\right)c$$

Assuming entry will occur until profits are driven to zero, and ignoring the integer constraint on N, the equilibrium number of entrants is $N = \sqrt{\frac{\alpha YM}{\sigma}}$, which increases monotonically with the size of the market YM. As demonstrated in Sutton (1991), this *fragmentation result* is robust to several alternative assumptions regarding the impact of horizontal differentiation, the timing of entry, and the type of product market competition. However, the result is broken when the level of quality is determined endogenously.

3.1.2 The Endogenous Sunk Cost Case

Letting quality z_j be a choice variable of the firm and proceeding via backward induction, I analyze the final product market competition stage first. Following Sutton (1991), I focus on identifying a symmetric equilibrium. The equilibrium quantities and prices are identical to the exogenous sunk cost case, although they now hold irrespective of the level of z. To calculate the equilibrium level of quality, I proceed by assuming that a single firm deviates from this symmetric equilibrium to offer quality z_1 while the remaining $N-1$ firms offer quality z. Equilibrium quality is then determined by the following first order condition

$$\frac{\partial \pi(z_1)}{\partial z_1} = 2\alpha YM(N-1)^2 \frac{[(N-1)z_1 - (N-2)z]z}{[(N-1)z_1 + z]^3} - F'(z_1) = 0 \tag{2}$$

To solve for the equilibrium level of quality, I follow Sutton (1991) in specifying the following cost function ($F(z)$)

$$C_j(p_L, c, z_j, q_j) = \sigma + \frac{p_L}{\gamma}(z^\gamma - 1) + cq_j$$

which includes both the exogenous entry cost σ and a second term that depends on the level of quality chosen. p_L is assumed to be the cost of land, since increasing product variety invariably requires expanding the size of the store. To solve for the symmetric equilibrium in quality, I let $z_j = z$ and solve (2) for z yielding

$$z = \left(\frac{2\alpha YM(N-1)^2}{N^3 p_L}\right)^{\frac{1}{\gamma}} \tag{3}$$

Since both quality and fixed costs grow proportionately with market size YM (they are constant in the exogenous case), it is not surprising to find an equilibrium where the number of firms does not expand with the size of the market.[5] This *non-fragmentation result* is established by imposing a zero profit condition and solving for the equilibrium number of firms.

Since entry in the first stage will drive profits to zero, ignoring integer constraints on the number of firms, the zero-profit condition is then given by

$$\left(\frac{p_L - \gamma\sigma}{\alpha Y M}\right) N^3 - 2N^2 + (4+\gamma)N - 2 = 0 \tag{4}$$

The fact that the number of firms will not increase indefinitely with the size of the market follows immediately from equation (4). In the limit, as market size YM increases to infinity[6], the lead term drops out, leaving a quadratic polynomial with root[7]

$$N = 1 + \frac{1}{4}\gamma + \frac{1}{4}\sqrt{8\gamma + \gamma^2} \tag{5}$$

which depends only on γ and is finite for all finite γ. Since the maximum number of entrants is finite, this equilibrium is referred to as natural oligopoly (Shaked and Sutton, 1983). From this simple framework, I have now identified a robust testable implication (natural oligopoly) as well as the mechanism that sustains it (escalation in quality). However, before taking the model to the data, I must first confront the presence of a competitive fringe.

3.1.3 Accounting for the Fringe

Like many retail industries, the supermarket industry includes a fringe of small firms in addition to the dominant chains. These small stores, remnants of the earlier grocery era, are essentially large convenience stores. They carry a limited range of products and choose mostly urban locations. Moreover, as I demonstrate below, they do not make the necessary chain level investments in distribution that would allow them to compete with the full-line supermarkets. This sharp distinction suggests treating these two groups as separate submarkets, differentiated on the basis of whether they compete in endogenous sunk investments. Sutton (1991, pp. 65-66) formalizes this two-tiered structure by extending his simple vertical model to include two submarkets, only one of which is subject to endogenous investment. In the high quality segment - supermarkets - firms make endogenous investments in quality and compete in a natural oligopoly. In the low quality segment - grocery stores - costs are determined exogenously and the number of entrants grows in proportion to the size of the market. This two tiered structure

[5] However, since quality also *decreases* monotonically with the cost of land, it is clearly important to control for the price of land when empirically evaluating the model. This is particulary relevant for distinguishing the endogenous sunk cost hypothesis from an alternative hypothesis in which *exogenous* sunk costs (i.e. land prices) simply expand with the size of the market. Without controlling for the price of land, these two hypotheses cannot be distinguished.

[6] For finite values of YM, the solution to the zero profit condition (4) depends on the sign of the lead term. In particular, whether the equilibrium number of entrants approaches the limit from above or below depends on whether $p_L - \gamma\sigma$ is positive or negative.

[7] The second root is always less than 1.

will guide our first set of empirical results by providing a control group of exogenous sunk cost firms to contrast with the natural oligopolists.

3.2 The Nature of Strategic Investment

Before turning to the empirical analysis, I present two simple modifications of Sutton's standard model. The first result establishes that the vertical model of quality investment can be reformulated to emphasize cost reduction. The second shows that the model can be modified so that firm investment decisions become strategic complements.

A Model of Cost-Reducing Investment Cost-reducing investments in distribution drove the diffusion of chain grocery stores in the 1920s. The current emphasis on information technology is likely to yield cost efficiencies in addition to expanding the number of products carried. The ESC model should be able to accommodate either case. This is in fact true. As the following proposition demonstrates, Sutton's vertical model can be reformulated as a model of cost-reducing investment where quality does not enter consumers' utility functions at all.[8] The following proposition establishes that all of the equilibrium properties of the standard model continue to hold in this setting.

Proposition 1 *The standard version of Sutton's ESC model is equivalent to a model of cost reduction.*

Proof in Appendix.

Clearly, whether the quality or cost interpretation will be more appropriate depends on the specific setting. In some applications, such as the semiconductor industry where cost per bit is a decreasing function of the size of the fabrication plant, the choice is obvious. In the case of supermarkets, the distinction is less clear. One way to distinguish between these cases is to focus on their consequences: falling prices or escalating quality. Another is to identify the form of strategic interaction.

Complements versus Substitutes Since escalation in sunk investments drives the natural oligopoly result, it is tempting to conclude that these investments should always be strategic complements. However, this turns out not to be the case. The following result establishes that substitution holds in Sutton's standard ESC model.

Proposition 2 *In the standard version of Sutton's ESC model, quality choices by rival firms are always (locally) strategic substitutes.*

Proof in Appendix.

[8] In the quality-enhancing model, quality and price enter the indirect utility function as a ratio. Since consumers do not distinguish between an increase in the quality-price ratio stemming from an increase in the perceived level of quality and an increase in the ratio due to a decrease in the "price of quality," the central insight of Spence (1976) applies: quality increases are equivalent to price reductions from the viewpoint of both buyer and seller.

In fact, this substitution result holds across most standard models of cost-reducing investment: Bagwell and Staiger (1994) demonstrate that investments in cost-reducing or quality enhancing R&D are strategic substitutes under quite general conditions.[9] Nevertheless, strategic complements seem more consistent with the ESC escalation mechanism sustaining oligopoly. Fortunately, it is relatively easy to modify this model to yield complementarity.

Proposition 3 *When the relationship between quality and price is non-linear, Sutton's ESC model is consistent with strategic complementarity.*

Proof in Appendix.

The example underlying Proposition 3 modifies the standard model so that consumers are willing to pay more for groceries if they are offered greater variety. As a result, increases in quality induce consumers to devote a larger fraction of their income to the quality-differentiated good, substituting away from the outside good which, in this case, is the competitive fringe.[10] Consequently, firms are no longer splitting a fixed pie; both consumption and the level of quality are determined endogenously and the strategic interaction shifts to complementarity. Because complementarity arises so rarely in investment games, it provides the opportunity for a strong empirical test of the ESC mechanism. This exercise will be the focus of section 6.2.

4 Data and Market Definition

The data for the supermarket industry are drawn from Trade Dimension's Retail Tenant Database for 1998. Trade Dimensions collects store level data from every supermarket operating in the U.S. for use in their *Marketing Guidebook* and *Market Scope* publications, as well as selected issues of *Progressive Grocer* magazine. The data are also sold to marketing firms and food manufacturers for marketing purposes. The (establishment level) definition of a supermarket used by Trade Dimensions is the government and industry standard: a store selling a full line of food products and generating at least $2 million in yearly revenues. Foodstores with less than $2 million in revenues are classified as small

[9] Athey and Schmutzler (2001) extend Bagwell and Staiger's results to include several additional classes of models, including Bertrand or Cournot competition with differentiated goods, constant marginal costs and linear demand (e.g. Dixit, 1979), HPD on the line (d'Aspremont et al., 1979) or the circle (Salop, 1979) with quadratic transportation costs, and the VPD model of Shaked and Sutton (1983).

[10] Supermarkets have always used a wider selection to induce consumers to substitute away from the corner grocer. More recently, responding to increased competition from take out restaurants, supermarkets have begun investing in prepared food counters targeting consumers who do not have time to cook (Progressive Grocer, 1999). In both cases, the investments are geared toward inducing substitution from an "outside" good. The complementarity result is not confined to the specific example presented here. Analogous findings employing the Shaked and Sutton (1983) Bertrand framework are developed in Ronnen (1991) and Lehmann-Grube (1997). In Ronnen's model, firms offer a staggered set of qualities. In the two-firm case, when the high quality firm raises quality, the low quality firm follows suit and vice versa. Ronnen's result follows from the fact that the market is not fully covered in equilibrium so that changes in quality induce consumers who previously consumed the outside good to join the market. Lehmann-Grube presents results similar to Ronnen's in a model with sequential entry.

convenience stores and are not included in the dataset. Firms in this segment operate very small stores and compete only with the smallest supermarkets (Ellickson (2000), Smith (2002)).

Information on average weekly volume, store size, number of checkouts, number of full and part time employees, whether scanners are in operation, and the presence or absence of various service counters (e.g. deli, seafood) as well as other measures of quality (e.g. ATM, check cashing) is gathered through quarterly surveys sent to store managers. These surveys are then compared with similar surveys given to the principal food broker assigned to each store, which are then verified through repeated phone calls. Market demographics are taken from the decennial Census of the U.S. and price data from the American Chamber of Commerce Researchers Association *Cost of Living Index*.

Testing the natural oligopoly result requires a set of reasonably independent markets that vary significantly in size. Since it is escalating sunk investment that renders further entry into each market unprofitable, it is essential that these costs not spill across markets. Retail industries, which are clearly spatially differentiated, provide a natural setting in which this is arguably the case. The supermarket industry is almost ideal because perishable goods can be shipped relatively short distances. Of course, defining markets accurately requires identifying both the relevant costs and how far they can be spread[11]. The ESC model suggests focusing on distribution networks, since these facilities constitute a primary (and observable) firm level investment.

The task of defining distribution markets is simplified by the fact that supermarket firms cluster their distribution centers in major cities (typically near a railroad spur) and serve surrounding areas from these facilities. For example, all of the major chains operating in Southern California operate warehouses in east Los Angeles. While the radius of operation of a typical distribution center varies geographically, the patterns are remarkably consistent across firms within regions, so that constructing markets simply involves plotting distribution networks and drawing boundaries around them. This is the method used by Trade Dimensions in constructing the 52 marketing areas reported in their Marketing Guidebook. My own analysis (Ellickson, 2001) produced only four changes, resulting in a total of 51 distribution markets.[12] These markets are much larger than MSAs, more closely resembling mid-sized states, and contain an average of over 5 million people and 593 stores.

[11] Previous studies of this industry (Chevalier (1995); Cotterill and Haller (1992)) and the Federal Trade Commission have focused on the MSA as the relevant geographic market, mainly because supermarket chains distribute advertising circulars at the MSA level. However, advertising is only one of the investments that firms make at the chain level and, with the growth of store level promotions (e.g. club cards), its importance has declined in recent years.

[12] The specific details of how these markets were constructed are described in Ellickson (2001), which also establishes the relatively high degree of independence between these markets by matching each store to its primary distribution center and measuring the degree of spillover across markets. In particular, I find that stores supplied by an out of market distribution facility owned by the parent firm account for, on average, less than 10% of total sales. This spillover is uncorrelated with the size of the market.

5 Evidence of Oligopoly in Supermarkets

The empirical results in this section establish the existence of a natural oligopoly and demonstrate that this oligopoly is sustained by escalating investments in quality enhancing sunk costs. I begin by showing that, in most markets, between 4 and 6 firms capture 60 to 70% of sales. Although larger markets do have more firms, the expansion is limited to a fringe of small stores, while the number and realized share of the oligopolists remains relatively stable. The oligopolists provide a distinct, higher quality product than the fringe, operating stores more than twice as large and offering far more services. They are also much more likely to operate their own distribution centers. Both this "quality wedge" and the importance of firm level investment are consistent with a two-tiered model. For the dominant firms, quality expands with the size of the market. The quality offered by firms in the fringe does not. Because the ESC framework can also be interpreted as a model of cost-reducing investment, I also consider the possibility that investments in distribution are focused on lowering costs. I find no evidence that prices fall as markets grow, suggesting that quality enhancement is more relevant.

5.1 Identifying an Oligarchy

The two-tiered structure of the supermarket industry is easily illustrated by constructing the empirical distribution of market shares across all 51 distribution markets. In particular, I constructed Lorenz curves for each market by first ranking firms according to market share and then plotting the cumulative share of sales against the cumulative share of firms. The first three panels of Figure 1 contain Lorenz curves for three individual markets: Spokane (WA), Denver (CO), and Washington (DC). Although the markets contain roughly 1.3, 4.7, and 9.8 million people respectively, the size distribution of firms is remarkably similar across the three. In each market, 5 or 6 firms account for the majority of sales. The remainder is split among a large fringe of very small firms. The main difference between these markets is the size of the fringe, which clearly grows with market size. The lower right panel of Figure 1 presents Lorenz curves for the full set of market, while the appendix contains two figures broken out by population quartile and region respectively. The uniformity in outcomes across markets is striking. For the full set of markets, 60 to 70% of sales are controlled by 4 to 6 firms, while the remainder of the market is captured by an expanding fringe of small firms. Although there is clearly some geographic variation[13], in no region do larger markets tend to have a greater number of *dominant* firms.

This observed structure is clearly inconsistent with exogenous fixed costs: if only 4 or 5 firms fit into the majority of large markets, we should expect to find monopolies in smaller markets, which

[13] For example, in the midwest, where third party distributors appear to be more successful in replicating the efficiencies of vertical integration, the dominant firms typically capture about 50% of the market.

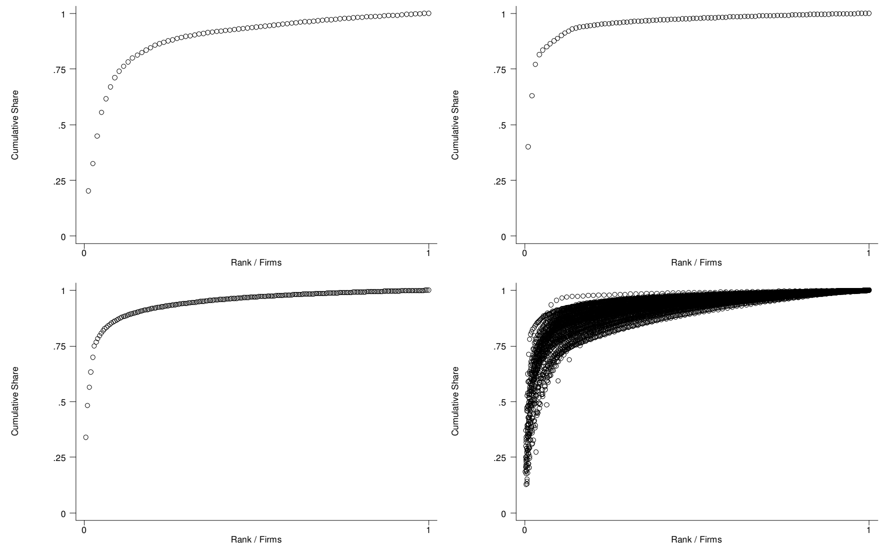

Figure 1: Lorenz Curves

we clearly do not. On the other hand, if minimum efficient scale can be attained by 3 or 4 firms in smaller markets, then larger markets should have many more entrants. The fact that we observe neither outcome suggests that both scale and the number of entrants are determined endogenously. The observed structure coincides with the two-tiered model described above[14]. According to that model, only the top tier of firms will invest in quality. I turn to this prediction next.

5.2 Top Firms Provide a Distinct Product

To justify treating the oligopoly as a separate submarket, I must first establish that its members provide a distinct product from the fringe. To demonstrate this, I split the firms into two groups, dominant and fringe, and calculate average values of several measures of quality for each group. I consider two alternative definitions of "dominant", belonging to the top 6 firms in a distribution market and being vertically integrated into distribution. The top 6 distinction is made separately in each market, so a chain in the top 6 in one market might belong to the fringe in another. Firm-specific levels of average quality[15] are calculated for each firm in each market, resulting in 7,995 firm/market level observations.

[14] Smith (2002), who analyzes supermarket competition using a discrete choice framework, finds a similar two-tiered, oligarch and fringe structure in the United Kingdom. He finds that competition between the tiers is minimal because these firms serve relatively distinct customer groups (primary versus secondary shoppers). A report by the British *Competition Commission* (2000) draws similar conclusions.

[15] The optimal measure of quality would combine a measure of variety with store size, since providing brandwidth requires stocking more products and building larger stores (wide aisles and easily accessible products consistently rate highly in consumer surveys (Progressive Grocer)). Since Trade Dimensions does not record the number of products carried by each store, I use store size (in 1000s of square feet) alone to measure quality. As a robustness check, I present three alternative quality measures constructed from store characteristics: the number of checkouts (cash registers), the number

The results are summarized in the first two columns of Table 1. Consistent with the predictions of the two-tiered model, the top 6 firms offer significantly higher levels of quality along all four dimensions (all differences are statistically significant at the 1% level). Specifically, the dominant firms operate stores that are more than twice the size of stores in the fringe, with over twice as many checkouts. The top 6 firms also operate many more stores and serve a much larger fraction of each market, all differences statistically significant at the 1% level. The majority of the top firms are also integrated into distribution, while only a negligible portion of the fringe firms supply themselves. The last two columns of Table 1 report averages split along this VI/non-VI dimension. VI firms operate stores that are more than twice the size of their non-VI counterparts.

Table 1
Store Characteristics by Firm Type

	Firm Type			
Characteristic	Top 6	Fringe	VI	Non-VI
Size	38.7	16.1	32.4	15.5
	(12.3)	(10.6)	(11.8)	(9.87)
Checkouts	12.6	5.61	10.1	5.37
	(6.70)	(3.28)	(4.67)	(2.62)
Technology	1.58	1.09	1.40	1.09
	(.314)	(.699)	(.444)	(.702)
Features	2.52	1.51	2.06	1.50
	(.755)	(.970)	(.832)	(.947)
Stores	51.6	1.88	185	1.97
	(53.9)	(3.18)	(298)	(4.90)
Percent VI	.737	.036	1	0
	(.438)	(.185)		
Market Share	.116	.002	.693	.307
	(.097)	(.004)		
Observations[16]	306	7689	87	7179

Standard deviations in parentheses.

Taken together, these results suggest that oligopolists provide a significantly different product from firms in the fringe. If scale economies and quantity discounts alone determine market structure, we would not expect the *store* characteristics of lead firms to differ from the followers. We would simply expect the market leaders to have lower prices.[17] However, the quality differential *is* consistent with the two-tiered model of ESC competition emphasized here.

of features present in a store (0-4 among an in-store bakery, restaurant, pharmacy, and deli) and a similar measure for scanning registers and ATM machines (technology).

[16] In the two left columns, observations are at the firm-market level, while in the right two columns they are at the firm level. This accounts for the difference in the total number of observations.

[17] This was precisely the case when A&P dominated the grocery industry in the 1920s (Tedlow, 1990).

5.3 Escalation in Quality

I have established that the oligopolists build larger stores on average than do firms in the fringe. However, the ESC model has a stronger implication: quality provided by the oligopolists (here proxied by store size) should increase with the size of the market. If the ESC mechanism were not in play, we would expect firms to build smaller stores, reflecting the high price of land in large markets. We find, however, that store size increases, but only among the dominant firms.

Table 2 presents several regressions relating average store size in each distribution market to population. The remaining exogenous variables in equation (3) are included as controls. Focusing first on the dominant firms, I find that average store size increases with market size. The first column of Table 2 contains the results of the regression of $\ln(\overline{Store\ Size})$ for the Top 6 firms[18] on the three exogenous variables in equation (3) ($\ln(Population)$, $\ln(Income)$, and $\ln(LandPrice)$), as well as four regional fixed effects (West, South, Midwest/North, East/Atlantic coast). Average housing cost per bedroom proxies for the cost of land.[19] The coefficient on population is positive and significant at the 1% level. The inclusion of market specific demographics in column 2 weakens the result somewhat, although the coefficient on $\ln(Population)$ remains positive and significant at the 5% level.

The existence of the fringe provides a natural control: if the escalation result only applies to firms that invest in distribution, it should not impact firms that do not. Columns 3 and 4 of Table 2 report regression results for non-vertically integrated firms. The coefficients on $\ln(Population)$ are insignificantly different from zero in both specifications (the point estimate is actually *negative* in the second regression). Columns 5 and 6 report similar results for firms that are classified as independent (meaning that they operate less than 11 stores).

[18] $\overline{Store\ Size}$ is constructed as the average store size across all of the stores operated by the Top 6 firms in each market, yielding 51 market level observations.

[19] This is the closest available proxy to the cost per square foot in each store. Cost per bedroom is averaged over all the zip codes that contain a supermarket, weighted by the share of stores in each zip code.

Table 2
Quality Regressions

Firm Type	Top		Not VI		Indep	
Dep Variable	ln(*Mean Store Size*)					
$\ln(Population)$	.061	.041	.023	-.020	.018	-.023
	(.021)	(.019)	(.367)	(.051)	(.025)	(.036)
$\ln(Med.Income)$	.483	.699	.226	.721	.076	.574
	(.200)	(.213)	(.305)	(.360)	(.241)	(.291)
$\ln(LandPrice)$	-.266	-.231	-.099	-.032	-.033	.006
	(.048)	(.048)	(.102)	(.079)	(.092)	(.073)
% Under 18		4.68		7.85		7.74
		(1.41)		(2.04)		(1.83)
% Over 64		2.21		3.97		4.00
		(1.20)		(2.16)		(1.70)
Constant		-2.84		-5.87		-4.90
		(1.94)		3.26		(2.72)
Region FEs	Inc		Inc		Inc	
$\widetilde{R}^2$	.24	.34	.28	.17	.30	.20
Observations	51					

Robust Standard Errors in parentheses.

5.4 Does Distribution Lower Costs?

As I have noted, the ESC model of quality enhancement can be reinterpreted as cost-reduction. In that case we should observe declining prices rather than escalating quality. I have already demonstrated that quality (store size) increases as markets expand in size. Here I will show that prices do not decline.

To evaluate the cost reduction hypothesis, price data from the same quarter as the store census were drawn from the American Chamber of Commerce Researchers Association (ACCRA) *Cost of Living Index*. The data are collected from surveys conducted by local chambers of commerce under ACCRA's guidance. The dataset includes prices for 27 specific grocery products, reported as MSA averages. I have converted these to distribution market averages by weighting the MSA averages in each distribution market by population. Prices are available for 48 of the 51 distribution markets.

The ACCRA provides an index of supermarket prices composed of a weighted basket of grocery products. I constructed an alternative "distribution index", using the same weights, but including only those products which are typically delivered to stores from distribution centers.[20] These "distribution products" are the most likely to reveal the impact of supply chain IT investment.

The first three columns of Table 3 contain regressions of this price index on the same covariates employed in the quality regressions. While the point estimate of the coefficient on $\ln(Population)$ is negative, it is insignificantly different from zero in all three specifications. To control for unobserved

[20] Some products, like Coca-Cola, are delivered directly to stores by the manufacturer, while other products, like produce and milk, are purchased locally. Since they don't pass through the firm's own distribution network, these direct store delivered products are unlikely to reflect chain-specific, distribution level efficiencies.

heterogeneity in costs, I use prices of four "reference products" reported in the same ACCRA survey to create price deflators. The reference products are drawn from industries which I believe to be unlikely to invest in cost-reduction (newspapers, dry cleaners, movie theaters, and pizza parlors) and their prices are included as divisors of the dependent variable. After including the reference price controls, the point estimates for the coefficient on $\ln(Population)$ are all *positive*, and in two cases, significant.[21]

Table 3: Price Regressions
Dependent Variable: ln(Price Index/Deflator)

Deflator	None	None	None	Newspaper	Dry Cleaner	Movie	Pizza
$\ln(Population)$	-.013	-.012	-.007	.011	.091	.007	.032
	(.015)	(.014)	(.017)	(.040)	(.016)	(.015)	(.017)
$\ln(Income)$	-.129	-.138	-.309	-.376	-.769	-.116	-.337
	(.101)	(.120)	(.138)	(.362)	(.195)	(.146)	(.153)
$\ln(LandPrice)$	.194	.190	.205	.041	.070	-.064	.111
	(.046)	(.050)	(.048)	(.101)	(.077)	(.056)	(.059)
% Under 18		-.184	-.464	-2.69	-4.15	-.599	-1.56
		(.745)	(.849)	(2.85)	(1.27)	(1.19)	(1.05)
% Over 64		-.314	-.883	-3.04	-3.70	-.627	-1.48
		(.605)	(.667)	(1.87)	(1.01)	(.681)	(.722)
Constant	-.895	-.698					
	(.657)	(.995)					
Region FEs			Inc	Inc	Inc	Inc	Inc
$\widetilde{R}^2$	.44	.45	.47	.07	.15	.18	.24
Observations				48			

Robust Standard Errors in parentheses.

6 Using Local Competition to Eliminate Alternative Models

I have argued that the ESC framework provides a compelling model of natural oligopoly in the markets served by distribution centers. These large regional markets are dominated by a small set of firms whose quality expands with the size of the market. However, this argument would be seriously undermined if at a finer level of spatial disaggregation supermarkets were local monopolists. Alternative models suggest such a possibility. For example, in horizontal models that feature persistent concentration such as product-proliferation with sequential entry (e.g. Schmalensee (1978)), single firms (or several firms acting as a cartel) produce all the products along a continuous segment of product space, thereby softening competition.[22] In horizontal models emphasizing price competition, firms typically isolate

[21] Regressions with the reference prices included as regressors, rather than as deflators of the dependent variable, yield similar results. Moreover, regressions of each "distribution" product deflated by the (direct store delivered) price of Coca-Cola produced coefficients on $\ln(Population)$ which were uniformly positive (and frequently significant).

[22] Bonano (1987) extends this analysis to include strategic location choice by a monopolist. Further persistence of local monopoly results are established by Prescott and Visscher (1977), Eaton and Lipsey (1979) and Reynolds (1987). In each of these models, competition is *localized* (Schmalensee, 1985) meaning that firms enjoy a monopoly over continuous regions of the product space. Consequently, a finding of head to head competition, where firms compete directly for the same consumers, is inconsistent with most standard models of HPD where equilibria are concentrated. An exception is Eaton and Lipsey (1982), where firms cluster around certain "poles", such as shopping districts and malls.

themselves to dampen its effect. In contrast, the ESC framework implies that firms will compete head to head. Since I have store level data, I am able to verify this prediction and rule out the other models.

Recent models of capacity competition explain the high degree of concentration in retail markets by using a combination of cost-reducing investment and costly consumer search (Bagwell and Ramey, 1994; Bagwell et al., 1997). Vigorous price competition leads to the emergence of a dominant low-cost, low-price leader, rather than oligopoly. In cost-reduction models of this type, as well as the cost-reducing version of the ESC framework, investments are strategic substitutes. However, as I found earlier, the ESC framework is one of the few settings in which these investments can be complements. Therefore, if it holds in the data, complementarity provides a powerful mechanism for distinguishing ESC from alternative models of competition.

6.1 Evidence Against Spatial Differentiation

As before, a firm is defined to be dominant in a distribution market if it is among the top 6 players. We now turn our focus to zip codes. Table 4 presents the average number of dominant *firms* per zip code, the average number of *stores* (of any type) per zip code, and the average number of dominant *stores* per zip code. Since zip codes vary considerably in size (they are much larger in western markets than elsewhere), the results are broken out by region. In each region, the average number of dominant firms is over 1 and close in magnitude to the number of dominant stores, suggesting that local monopoly is relatively rare. If a zip code is large enough to hold more than one store, it usually contains more than one firm. Table 5 verifies this pattern by conditioning on the number of stores operated by dominant firms in each zip code. For zip codes containing two or more stores operated by dominant firms, Table 5 presents the frequencies of each possible market configuration. Again we see that multi-store monopoly is an extremely rare occurrence; when zip codes contain more than one dominant *store*, they usually contain more than one dominant *firm*.

Table 4
Dominant Firms Per Zip Code

Region	Stores (Total)	Stores (Dominant)	Dominant Firms	Observations	Population
New England	1.8	1.6	1.4	609	16379
Mid-Atlantic	2.3	1.7	1.5	2055	20230
South East	2.7	2.3	1.9	2664	17772
East Central	2.2	1.7	1.5	1725	16226
West Central	2.0	1.8	1.5	2189	15565
South West	2.3	1.8	1.6	1287	18218
Pacific	2.8	2.3	1.9	2040	24974
Total Zip Codes			12569		

Table 5
Monopoly: Conditional on the Number of Stores Operated by Dominant Firms

	Number of Dominant stores				
Configuration	2	3	4	5	6+
Monopoly	308	33	11	1	0
Duopoly	1771	480	141	32	6
3-opoly	-	497	298	92	69
4-opoly	-	-	114	93	81
5-opoly	-	-	-	14	39
6-opoly	-	-	-	-	4
Total Markets	4084				

To provide a more formal test, I focus on the narrower question of whether the top firm in a market chooses store locations more spatially clustered than the industry as a whole. This hypothesis can be tested using the "dartboard" index of spatial agglomeration (Ellison and Glaeser, 1997). In this setting, what I mean by "agglomeration" is the concentration of stores by the top firm in relatively few locations. Specifically, for each distribution market I construct the following measure of concentration for the top firm's stores:

$$\gamma = \frac{\sum_i (s_i - x_i)^2 - \sum_i (1 - x_i^2) \cdot \frac{1}{N}}{\sum_i (1 - x_i^2) \cdot (1 - \frac{1}{N})}$$

where s_i is the top firm's share of stores in zip code i, x_i is the share of distribution market population residing in zip code i, and N is the total number of stores in the distribution market. If firms succeed in dividing the market into local monopolies, the top firm should be more clustered than either the industry as a whole or the set of top 6 firms, resulting in a larger value of γ.

Table 6 presents parameter estimates of γ calculated for the top firm, the top 6 firms, and the industry as a whole. I compute γ for each set of firms using three submarket definitions: zip code, county, and MSA.[23] Focusing first on γ calculated for the industry as a whole (store γ), I find that γ is very close to zero in all three submarkets. Since we expect retail firms to locate close to their consumers, this is not surprising.[24] For each definition of local submarket, the estimate of γ for the lead firm (top store γ) is smaller than γ for either the industry as a whole (store γ) or the set of top 6 firms (top 6 store γ), indicating that the store locations chosen by the top firm are less spatially clustered than either the industry as a whole or the full set of dominant firms.[25] These results are clearly inconsistent with product proliferation. Overall, I find no evidence that firms succeed in differentiating themselves

[23] The sample includes all markets in the dataset where γ is defined. Any market which contains only one submarket must be dropped from the sample, so Alaska and Hawaii are not included in the results in column 3 (MSA submarkets).

[24] Ellison and Glaeser find that γ is closest to zero (no excess concentration) in markets where firms must locate close to their end users.

[25] Because γ is a parameter estimate, the standard deviations of γ are much larger for the set of top firm stores, since fewer "darts" are being thrown. Restricting the sample by population to include only large markets improves the precision of the estimates considerably.

spatially from their competitors. Instead, they compete head to head, as the ESC framework implies. Given that firms face rivals in every location, it is natural to then ask how they react to the local actions of their rivals. This will form the final and most restrictive test of the ESC framework.

Table 6
Concentration in Local Markets: The Dartboard

	Submarket		
	Zip Code	County	MSA
Store γ	-.002	-.002	.002
	(.002)	(.002)	(.016)
Top 6 Store γ	-.005	-.005	.002
	(.004)	(.004)	(.020)
Top Store γ	-.015	-.014	-.009
	(.013)	(.014)	(.033)
Observations	51	51	49

6.2 Evidence for Local Rivalry

To identify the form of strategic interaction, I focus on the zip code as a local market,[26] and take quality choice (store size) to be the dependent variable. Starting from the dataset of store level observations in all 51 markets, I once again select out only those stores operated by firms which are in the top 6 in each distribution market. Fringe firms are assumed to be strategically independent from the top 6. A top 6 firm may then face between 1 and 5 other top 6 firms in a given zip code market.

To quantify the strategic interaction between firms, I estimate the reaction functions of competing firms using the following regression:

$$\ln(Size_{ij}) = \alpha_1 \cdot \ln(Avgsize_{\setminus i}) + \alpha_2 \cdot \ln(Avgsize_i) + \sum \alpha_3 \cdot Market_j + \sum \alpha_4 \cdot County + \varepsilon_{ij} \qquad (6)$$

where $Size_{ij}$ is the size of store i in zip code j, $Avgsize_{\setminus i}$ is the average size of store i's competitors in zip code j, and $Avgsize_i$ is the average size of the stores of the firm that owns store i, outside of zip code j. $Market_j$ is a set of (logged) zip code level demographic and market characteristic variables, $County$ a full set of county level fixed effects, and ε_{ij} is an error term. The local demographic and market variables include population, median household income, median age, median home value, and the percent of the population that is urban or Hispanic. Because newer markets undoubtedly contain larger stores, I include a store index code[27] to control for the age of the market and store.

[26] I will also present results for two larger local market definitions, 3 and 4 digit zip codes.

[27] The store index codes were entered sequentially by Trade Dimensions as stores were opened, providing a rough timeline. The codes also contain gaps reflecting the entries for establishments from other retail industries which gives the index some cardinal as well as ordinal properties.

Since the store size chosen by a rival firm is clearly endogenous (owing to the simultaneity of firms' actions), equation (6) cannot be consistently estimated using OLS. Moreover, unobserved factors such as an advantageous location in a shopping district, a disproportionate share of commuters, or idiosyncratic consumer preferences might cause some zip codes to have larger or smaller stores on average. Not all of these effects will be captured by demographic variables, county fixed effects, or the store age index. The importance of unobserved heterogeneity is not unique to this setting, arising as well in the context of peer effects, locational sorting, and entry games (e.g. Bayer and Timmins (2003), Bajari and Krainer (2004)). Following this literature, I proceed by identifying a suitable instrument for competitor's size.

Specifically, I propose instrumenting competitors' store size with their average store size outside of the distribution market.[28] This is a similar approach to the strategy used for prices in both Hausman (1994) and Nevo (1998). There are at least two reasons why a firm's size decisions should be correlated across markets. First, we have already established that the scale economies associated with providing a broad selection of products involve investments in distribution that are shared across stores. Second, the benefits of maintaining a reputation for high quality may extend across markets, as will the returns from advertising. Identification therefore requires that the tendency to provide larger stores in general be unrelated to the idiosyncratic forces driving store size to be large in any particular local market. In the context of peer effects, Bajari and Krainer (2004) demonstrate that instrument validity hinges on finding a covariate of an agent's action that does directly impact the actions of other agents. In the case of supermarkets, it seems reasonable to assume that the reactions of a single store to its competitors actions in that market only depend on the competitor's actions outside that market through its actions in that market. In other words, an individual store only cares about the size portfolio of its competitor's stores through its impact on that competitor's store size in that particular market.

Having constructed an appropriate instrument, the first column of Table 7 can be viewed as the first stage of a two stage regression. The remaining columns present several alternative specifications for equation (6). The second column of Table 7 contains a baseline specification involving only own size and competitor's size. The third and fourth columns test the robustness of this specification by adding first county fixed effects and then zip code demographics. The coefficient on competitor's size remains positive and significant at the 1% level in each specification. The coefficient on ln(Population) is also positive and significant, showing that the escalation result holds at the local level as well. The fifth and sixth column repeat the specification of column 4, using the larger 4 digit and 3 digit zip code market definitions to address issues of selection caused by focusing on markets with at least two top 6

[28] This instrument may be constructed in several ways, using a firm's average outside this zip code but within this market, across all stores outside this zip code (all markets) or across all stores outside this market. Since the results are robust to the choice of alternative, I will focus on the latter.

firms. The size effects are bigger for the larger market definitions, which is not surprising. Again, in every specification the complementarity result is positive and significant at the 1% level.

Together, these regression results provide strong evidence that the quality levels chosen by rival firms are strategic complements. While this result is consistent with several models of ESC that emphasize the demand expanding effect of quality enhancement, it casts significant doubt on a number of competing explanations of local market structure, particularly models of cost reducing investment and product proliferation. Establishing that the ESC framework provides an accurate portrait of local competition helps justify its use in explaining firm level competition as well. Clearly, a similar exercise using firm level investment data would be very informative. However, this evidence on the actual shape of firm's reaction functions in local markets, together with the picture of the competitive structure of local competition presented earlier, suggests that the competitive, rivalrous emphasis of the ESC framework accords well with the observed structure of the supermarket industry.

Table 7
Estimating Reaction Functions

	Market Definition					
	Zip Codes				4 Digit Zip	3 Digit Zip
ln(Competitors' Size)		.176	.205	.155	.282	.259
		(.029)	(.046)	(.046)	(.047)	(.101)
ln(Own Size)	.796	.786	.803	.722	.744	.762
	(.019)	(.019)	(.026)	(.026)	(.023)	(.033)
ln(Population)				.057	.002	.017
				(.012)	(.009)	(.011)
ln(Med. Income)				.049	.124	.119
				(.043)	(.060)	(.088)
ln(Med. Home Value)				-.055	-.113	-.159
				(.033)	(.044)	(.063)
ln(Store Index)				.111	.112	.112
				(.004)	(.003)	(.003)
Constant	.689	.097				
	(.070)	(.118)				
Market Level Controls				Included	Included	Included
County Fixed Effects			Included	Included	Included	Included
Observations	8636				11436	12953

Dependent Variable: ln($Size$). Standard Errors in parentheses.

7 Conclusions

This paper proposes and tests a model of the supermarket industry in which supermarket firms invest in endogenous sunk costs to improve service quality. The model is consistent with a number of facts about this industry. Regional markets of widely varying size are dominated by a small number of firms. This natural oligopoly of supermarket chains, each operating a large number of large stores, dominates a fringe of small firms, each operating a few small stores. The size of the stores operated by the oligopolistic chains expands with the extent of the market. The oligopolistic chains do not carve out separate turf, choosing instead to compete head to head with their rivals, with choice of store size behaving as a strategic complement. No other theory seems capable of explaining these facts.

The same features seem to characterize modern retailing in many arenas, ranging from coffee shops to electronics stores. Whether this conjecture holds up remains an open question.

References

Athey, S. and Schmutzler, A., 2001, 'Investment and Market Dominance, *RAND Journal of Economics,* 32(1), pp. 1-26.

Bagwell, K. and Ramey, G., 1994, 'Coordination Economies, Advertising, and Search Behavior in Retail Markets', *American Economic Review*, 84, pp. 498-517.

Bagwell, K., Ramey, G. and Spulber, D., 1997, 'Dynamic Retail Price and Investment Competition', *RAND Journal of Economics*, 28, pp. 207-227.

Bagwell, K. and Staiger, R., 1994, 'The Sensitivity of Strategic and Corrective R&D Policy in Oligopolistic Structures,' *Journal of International Economics*, 36, pp. 133-150.

Bajari, P. and Krainer, J., 2004, 'An Empirical Model of Stock Analysts' Recommendations: Market Fundamentals, Conflicts of Interest, and Peer Effects', Working Paper, Duke University.

Bayer, P. and Timmins, C., 2003, 'Estimating Equilibrium Models of Sorting Across Locations', Working Paper, Yale University.

Berry S., and Waldfogel, J., 2003, 'Product Quality and Market Size', NBER Working Paper #9675.

Bonano, G., 1987, 'Location Choice, Product Proliferation, and Entry Deterrence', *Review of Economic Studies*, 54, pp. 37-46.

Bulow, J., Geanakopolos, J., and Klemperer, 1985, 'Holding Idle Capacity to Deter Entry,' *Economic Journal*, 95, pp.178-82.

Bulow, J., Geanakopolos, J., and Klemperer, 1985, 'Multimarket Oligopoly: Strategic Substitutes and Complements', *Journal of Political Economy*, 93, pp. 488-511.

Chevalier, J.A., 1995, 'Capital structure and product-market competition: empirical evidence from the supermarket industry', *American Economic Review*, 85, pp. 415-435.

Competition Commission, 2000, 'Supermarkets: A Report on the Supply of Groceries from Multiple Stores in the United Kingdom', The Stationary Office, London, UK.

Cotterill, R.W., and Haller, L.E., 1992, 'Barrier and Queue Effects: A Study of Leading U.S. Supermarket Chain Entry Patterns', *Journal of Industrial Economics*, 40(4), pp. 427-440.

d'Aspremont, C., Gabszewicz, J., and Thisse, J., 1979, 'On Hotelling's Stability in Competition', *Econometrica*, 17, pp. 1145-51.

Dick, A., 2003, 'Market Structure and Quality: An Application to the Banking Industry', Finance and Economics Discussion Series 2003-14. Washington: Board of Governors of the Federal Reserve System.

Dixit, A., 1979, 'A Model of Duopoly Suggesting a Theory of Entry Barriers', *Bell Journal of Economics*, 10, pp. 20-32.

Dixit, A., 1980, 'The Role of Investment in Entry Deterrence', *Economic Journal*, 90, pp. 95-106.

Eaton, B.C. and Lipsey, R.G., 1979, 'The Theory of Market Preemption: The Persistence of Excess Capacity and Monopoly in Growing Spatial Markets', *Econometrica*, 47, pp. 149-58.

Eaton, B.C. and Lipsey, R.G., 1982, 'An Economic Theory of Central Places', *Economic Journal*, 92, pp.56-72.

Ellickson, P., 1999, 'Vertical Product Differentiation and Competition in the Supermarket Industry', Ph.D. Thesis, Massachusetts Institute of Technology.

Ellickson, P., 2001, 'Does Sutton Apply to Supermarkets?', Working Paper, Duke University.

Ellison, G., and Glaeser, E., 1997, 'Geographic Concentration in U.S. Manufacturing Industries: A Dartboard Approach', *Journal of Political Economy*, 105(5), pp. 889-927.

Gabszewicz, J.J. and Thisse, J.F., 'Entry (and Exit) in a Differentiated Industry', *Journal of Economic Theory*, 22, pp. 327-338.

Hausman, J., 1997, 'Valuation of New Goods under Perfect and Imperfect Competition,' in T. Bresnahan and R. Gordon (eds.) *The Economics of New Goods*, University of Chicago Press, 209–237.

Latcovich, S. and Smith, H., 2001, 'Pricing, Sunk Costs, and Market Structure Online: Evidence from Book Retailing', *Oxford Review of Economic Policy*, 17(2), pp. 217-34.

Lehmann-Grube, U., 1997, 'Strategic Choice of Quality When Quality is Costly: the Persistence of the High-Quality Advantage', *RAND Journal of Economics*, 28, pp. 372-84.

Lyons, B. Matraves, C. and Moffatt, P., 2001, 'Industrial Concentration and Market Integration in the European Union', Economica, 68, 1-26.

Messinger, P.R. and Narasimhan, C., 1995, 'Has power shifted in the grocery channel?', *Marketing Science*, 14, pp. 189-223.

Nevo, A., 2001, 'Measuring Market Power in the Ready-to-Eat Cereal Industry,' *Econometrica* 69, 307-342.

Prescott, E. and Visscher, M., 1977, 'Sequential Location Among Firms with Foresight', *Bell Journal of Economics*, 8, pp. 378-93.

Progressive Grocer, 'Annual Report of the Supermarket Industry', selected April issues.

Progressive Grocer, 1998, *Marketing Guidebook*, MacLean Hunter Media: White Plains, N.Y.

Reynolds, S., 1987, 'Capacity Investment, Preemption and Commitment in an Infinite Horizon Model', *International Economic Review*, 28, pp. 69-88.

Robinson, W.T. and Chiang, J., 1996, 'Are Sutton's Predictions Robust?: Empirical Insights into Advertising, R&D, and Concentration', *Journal of Industrial Economics*, 64, 389-408.

Ronnen, U., 1991, 'Minimum Quality Standards, Fixed Costs, and Competition', *RAND Journal of Economics*, 22(4), pp. 490-50.

Salop, S., 1979, 'Monopolistic Competition with Outside Goods', *Bell Journal of Economics*, 10, pp. 141-156.

Schmalensee, R., 1978, 'Entry Deterrence in the Ready-to-Eat Breakfast Cereal Industry', *Bell Journal of Economics*, 9, pp. 305-327.

Schmalensee, R., 1985, 'Econometric Diagnosis of Competitive Localization', *International Journal of Industrial Organization*, 3, pp. 57-70.

Schmalensee, R., 1992, 'Sunk Costs and Market Structure: A review article', *Journal of Industrial Economics*, 60, pp. 125-133.

Shaked, A. and Sutton, J., 1982, 'Relaxing price competition through product differentiation', *Review of Economic Studies*, 51, pp. 3-13.

Shaked, A. and Sutton, J., 1983, 'Natural Oligopolies', *Econometrica*, 51, pp. 1469-1484.

Shaked, A. and Sutton, J., 1987, 'Product Differentiation and Industrial Structure', *Journal of Industrial Economics*, 60, pp. 889-917.

Shaked, A. and Sutton, J., 1990, 'Multiproduct Firms and Market Structure', *RAND Journal of Economics*, 21, pp. 45-62.

Smith, H., 2002, 'Supermarket Choice and Supermarket Competition in Market Equilibrium', Forthcoming in *Review of Economic Studies*.

Spence, M., 1976, 'Product Selection, Fixed Costs and Monopolistic Competition', *Review of Economic Studies*, 43, pp. 217-35.

Spence, M., 1977, 'Entry, Capacity, Investment and Oligopolistic Pricing', *Bell Journal of Economics*, 8, pp. 534-44.

Spence, M., 1979, 'Investment Strategy and Growth in a New Market', *Bell Journal of Economics*, 10, pp. 1-19.

Sutton, J., 1991, *Sunk Cost and Market Structure: Price Competition, Advertising, and the Evolution of Concentration* (MIT Press, Cambridge).

Tedlow, R.S., 1990, *New and Improved: The Story of Mass Marketing in America*, (Basic Books, New York).

A Appendix

Proposition 1 *The standard version of Sutton's ESC model is equivalent to a model of cost reduction.*

Proof. Assume the M identical consumers each have utility

$$u(x_1, x_2) = (1-\alpha)\ln(x_1) + \alpha \ln(x_2) \tag{7}$$

where x_1 is the quantity consumed of the composite good and x_2 the quantity of the differentiated good under analysis. There are N identical firms, where firm j uses input $F(z_j) = \sigma + \frac{c}{z_j} q_j$ of the composite good to produce quantity q_j of x_2. In particular, the firm may invest in fixed costs in order to reduce marginal costs by the fraction $\frac{1}{z_j}$.[29] Although x_2 does not appear to be a quality-differentiated good to consumers, it is clearly differentiated on the input side. Let $p(z)$ be the price of the differentiated good with cost reducing parameter z. Solving for the quantities demanded and plugging into the indirect utility function yields the following profit function for firm j

$$\pi_j = \frac{\alpha Y M}{\sum_{j=1}^{N} q_j} q_j - \frac{c}{z_j} q_j - F$$

Evaluating the associated first order conditions yields equilibrium quantities and price

$$q = \left(\frac{N-1}{N^2}\right) \frac{\alpha Y M z}{c} \quad \& \quad p = \left(\frac{N}{N-1}\right) \frac{c}{z}$$

which now depend on the level of z determined in the second stage. In particular, we find that price falls monotonically as z increases. Since z still expands with the size of the market, larger markets will have lower prices.

Solving for the symmetric level of z yields the first order condition

$$\frac{\partial \pi(z_1)}{\partial z_1} = 2\alpha Y M (N-1)^2 \frac{[(N-1)z_1 - (N-2)z]z}{[(N-1)z_1 + z]^3} - F'(z_1) = 0 \tag{8}$$

which is identical to the quality-enhancing case. In all other respects, the results are identical to the standard model. ■

Proposition 2 *In the standard version of Sutton's ESC model, quality choices by rival firms are always (locally) strategic substitutes.*

Proof. Although the model does not yield an analytical solution for a firm's best response function, by the implicit function theorem it has the same sign as the cross partial derivative of the profit function (2):

$$\frac{\partial^2 \pi(z_1)}{\partial z_1 \partial z} = 2\alpha Y M (N-1)^2 \frac{\left[(N-1)^2 z_1^2 - 2(N-1)^2 z_1 z + (N-2) z^2\right]}{[(N-1)z_1 + z]^4} \tag{9}$$

[29] For example, a supermarket building larger stores faces lower inventory costs per item and a microchip producer building a larger fabrication plant produces chips with a lower cost per bit.

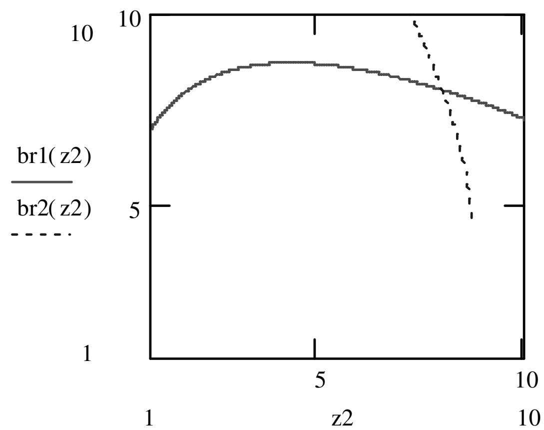

Figure 2: Best response functions

where z_1 represents the quality choice of the deviating firm. Evaluated at $z_1 = z$, equation (9) reduces to

$$\frac{2\alpha YM(N-1)^2}{N^4 z^2}(-N^2+3N-3) \tag{10}$$

which is strictly negative. Therefore, near the equilibrium, quality choices are always strategic substitutes.

Away from equilibrium, quality choices may be either substitutes or complements, as the following example illustrates. Choosing parameters $\alpha = \frac{1}{2}$, $p_L = 1$, $\sigma = 1$, $YM = 512$ and $\gamma = 2$, yields equilibrium $z = 8$ and 2 entrants. Solving for firm 1's best response as a function of firm 2's quality yields

$$br(z_2) = 8\sqrt[3]{z_2} - z_2$$

Figure 2 shows the best response functions of each firm. They are clearly negatively sloped at the equilibrium and the portions over which they are positively sloped occur quite far from the equilibrium. If the fixed cost functions for rival firms are sufficiently different, it is possible for the reaction functions to cross at a point where quality is a complement for the low cost firm and a substitute for the high cost firm (imagine shifting the dotted curve in figure 2 far to the left),[30] but they cannot be complements for both. ■

Proposition 3 *When the relationship between quality and price is non-linear, Sutton's ESC model is consistent with strategic complementarity.*

Proof. Replacing the utility function (1) by

[30]This outcome recalls the Bulow et al. (1985a) model of capacity competition with extremely convex demand functions, where the strategic interactions are also asymmetric and the reaction functions are nearly identical to those presented here.

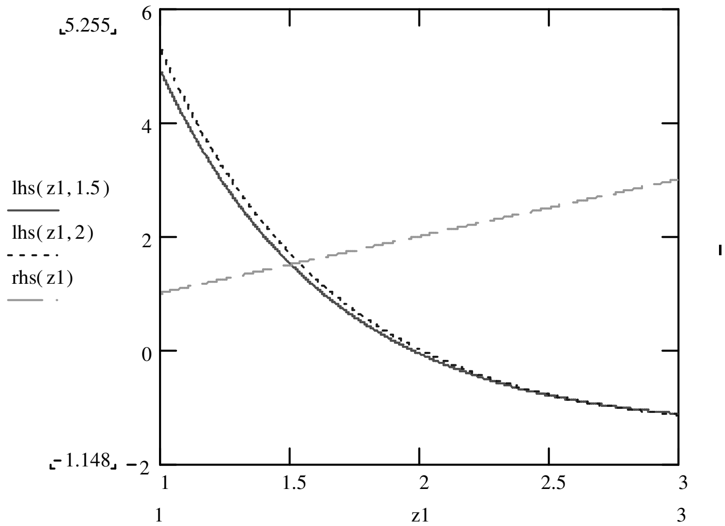

Figure 3: An example of strategic complementarity

$$u(x,z) = (1-\alpha)\ln x_1 + \alpha z \ln x_2 \tag{11}$$

we obtain equilibrium quantities and price

$$q(z) = \left(\frac{N-1}{N^2}\right)\left(\frac{\alpha z}{1-\alpha+\alpha z}\right)\frac{YM}{c} \quad \& \quad p = \left(\frac{N}{N-1}\right)c \tag{12}$$

where quantity now depends on the choice of z determined in the previous stage.

Evaluated at $z_1 = z$, the cross-partial derivative of the profit function reduces to

$$\frac{2\alpha YM(N-1)^2}{(1-\alpha+\alpha z)^2 N^3}\frac{\phi'(z)}{\phi(z)}\left[-(1-\alpha z) + \frac{2N-3}{N}(1-\alpha)(1-z)\right] \tag{13}$$

where

$$\phi(z) = \frac{z}{(1-\alpha+\alpha z)^{1/\alpha}}$$

(compare equation (10)).

The term outside the brackets is strictly positive. Inside the brackets, the second term is negative for all $N > 1$ and $z > 1$, while the first term depends on the level of z, yielding an analog of the income and substitution effect. The following example demonstrates a case in which the effect of the first term outweighs the second. Choosing parameters $\alpha = \frac{1}{2}$, $p_L = 1$, $\sigma = \frac{85}{8}$, $YM = 75$ and $\gamma = 2$, yields 2 equilibrium entrants ($N = 2$) and equilibrium $z = 1.5$. Figure 3 plots the right hand side (rhs) and left hand side (lhs) of the first derivative of profit as a function of z_1. $lhs(z_1, 1.5)$ uses the equilibrium level of z, while $lhs(z_1, 2)$ uses $z = 2$. The effect of an increase in z is to shift $lhs(z_1, z)$ up, increasing the point of intersection and the equilibrium level of z_1. Therefore, at least locally, the slope of the reaction function is positive. Consequently, the optimal response to a rival's quality increase is to increase own quality. Unlike the model of cost-reducing investment presented above, investment by rival firms actually increases the return to own investment, resulting in strategic complementarity. ■

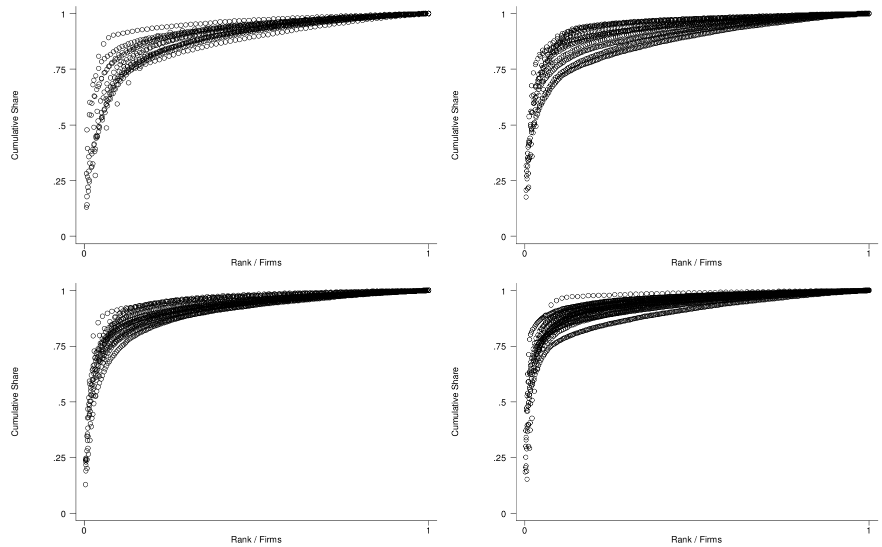

Figure 4: Lorenz Curves (By Population Quartile)

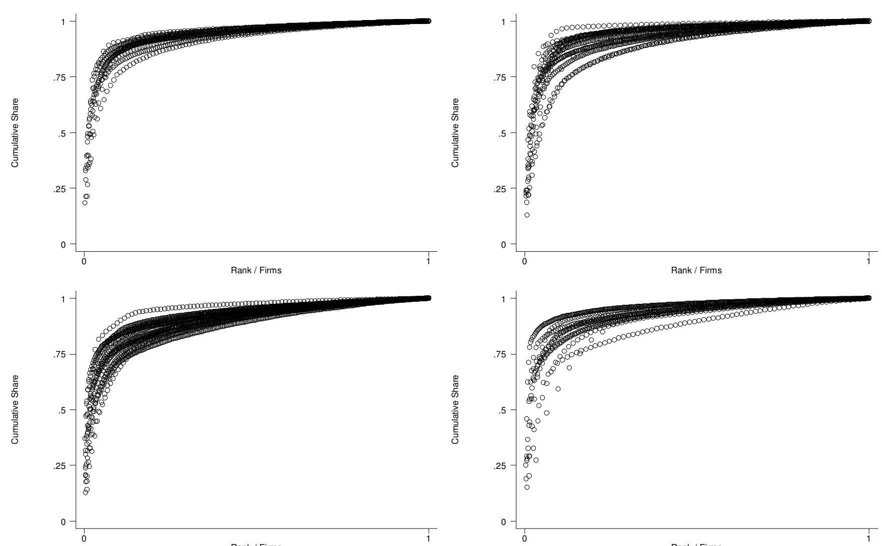

Figure 5: Lorenz Curves (By Region)