# Retail market share and saturation

## Morton O'Kelly

*Department of Geography, Ohio State University, 1036 Derby Hall, 154 N Oval Mall, Columbus, OH 43210, USA*

**Abstract**

Graphing the relationship between market share and share of space in a set of urban areas is a useful way to describe issues of market saturation. This paper considers two versions of the relationship between market share and share of space. The two models are non-linear in nature: one is a logistic s-shaped curve, and the other is an exponential growth curve. Mahajan et al. (1988). Journal of Retailing 64(3) 315–333) have discussed these curves in the literature, as a tool for determining market expansion. In this paper the models are compared and discussed with regard to their fundamental assumptions and interpretation. The paper includes notes on formulation, implementation, estimation, and application. The concepts and techniques in the paper are illustrated with data from grocery chains collected from markets in the southeastern United States. © 2000 Elsevier Science Ltd. All rights reserved.

*Keywords:* Supermarkets; Location theory; Market share

## 1. Introduction: why study market share?

Market and space share data are worth studying from several different perspectives. First, there is a great deal of emphasis placed on the measurement and tracking of market share by retail chains, and many are engaged in tough battles to increase market share (Langston et al., 1997). Second, important clues about structural differences in retail operations are provided by size and share data. Third, market share data are useful in measuring the co-presence of chains before a potential merger. Excessive market concentration in overlapping markets may be cause for concern by regulators, while mergers of complementary (i.e. non-overlapping regions) are less likely to cause concern. In markets where combined market share is judged to be excessive, the merging chains can be made to divest of a number of properties to satisfy regulators. Fourth, market share data can give a direct measurement of the level of concentration of retail sales share among the top chains in a market. In grocery retailing for example, many Southeastern US markets are dominated by 3 or 4 major chains. There is intense competition in these markets where several chains have store locations within a few blocks of each other, making it difficult for new entrants considering the

market, and a hostile environment for smaller participants already in those markets. Examples of markets with both high concentrations and intense competition between chains include Columbia SC and Wilmington NC. (See Posey (1994) for remarks about the specific case of Columbia SC and also the work of Wrigley (1998,1999) for analysis and commentary of the impact of leveraged buyouts on concentration.)

Retail market share is dependent on many factors. This paper examines the geographical break down of a retail chain store's market share. A chain's share of market and of floorspace in each of a number of cities is examined. It is clear that the share in any market is a function of the size of that chain's presence in the market and the number and strength of competitive chains. The goal of this paper is to examine market share data from the point of view of a multi-store, multi-market retailer (Mahajan et al., 1988), with particular emphasis on saturation. While the idea of saturation is generally used with respect to all large firms competing in a market, (Langston et al., 1997) this paper looks at the individual retailer and attempts to analyze the likelihood of that firm gaining extra market share as a result of extra floor space in a particular city or region. This is an inherently dynamic question, and Mahajan et al. (1988) clearly suggest that an empirically fitted curve can be used to analyze and diagnose various market positions. This curve in turn might be estimated from cross-sectional data from similar operations by a firm in comparable markets (e.g. from the records of

*E-mail address:* okelly.1@osu.edu (M. O' Kelly).

many stores, clustered into cities in a homogeneous region). Such an analysis might give some basic idea about the way that a firm's market position varies with different levels of floor space share. This movement back and forth between cross-sectional data and dynamic interpretations certainly poses some difficulty, but this paper attempts to make these interpretative connections. Many complex issues are sidestepped here, but are recognized as fruitful avenues for further research: for example, it is assumed throughout that the aggregate size of the market in an area is held fixed, though in reality creative retailing strategies can create a larger total demand. Further while some discussion of differential locational strategies by firms is provided, the full extent of interdependence between the actions of multiple players in a market are not considered. We take the viewpoint of a single operator playing a zero sum game against a homogeneous opponent, though in reality retail competition and saturation undoubtedly create a much more complex set of outcomes.

This paper considers issues of retail market share in the context of two specific versions of the relationship between market share and share of space. The two models are non-linear in nature: one is a logistic s-shaped curve, and the other is an exponential growth curve. The sections of the paper are in the following order: model formulations (emphasizing basic assumptions); implementation; estimation; interpretation and applications. The paper works with empirical measures of sales and space share as aggregated from individual store sales estimates in selected Southeastern US markets. The ideas are illustrated with data from 53 SE US market areas using retail grocery data. Among the major participants in these markets are Winn Dixie, Harris Teeter, Food Lion, Bilo, Kroger, and Ingles. To preserve the confidentially of the data the specific chains and place names will not be detailed.

## 2. Models and notation

Suppose as a simple starting point that each town has one store representing each chain. Market share would simply be the percentage of the total business in the town accounted for by each store. For a small town with a rural catchment area, it is safe to assume that all the stores in the town directly compete with each other, as the relative differences in their locations in the town are minor compared to the overall separation between the consumers and the store. (For example, from 12 miles out, a 1.5 mile difference in store location in a town is inconsequential, so that they are essentially treated as if they co-exist at a point location.) This is somewhat similar to the assumption of central place theory, where the town is treated as a point location and there are no major agglomeration effects built into the analysis.

Matters become a bit more complex when towns have multiple stores, some of various kinds, and are large enough to have distinguishable sub-markets within their boundaries. How do we get the store share of market and space? Assume we have an estimate of the store floor-space and sales. Then aggregate by chain, so that the sum of sales and square footage is obtained and tabulated by chain for each town. This method would work very well for most purposes: i.e. if we wanted to know a general measure of "market share" for two chains in a specific city, this would be an appropriate way to do it. As a practical matter retail market share data are reported by many retail analytical services: these are subject to a wide margin of variability due to reporting differences, rapid developments in store opening and closing, and different data sources. There is always going to be difficulty accounting for every dollar of sales and every square foot of space, and the role of independent grocers in confounding these data should not be overlooked.[1]

The market and space share variables are related by the following formula for the logistic s-shaped curve:

$$M_{jt} = K/[1 + \exp(a - bS_{jt})] \tag{1}$$

where $M_{jt}$ is market share of chain $j$ in town $t$; indicating the percent of sales in the market area $t$, accounted for by chain $j$ and $S_{jt}$ is share of space of chain $j$ in town $t$; indicating the percent of floorspace in the market area $t$, accounted for by chain $j$.

In the notation above, $K$ is a saturation level beyond which it would be unrealistic to expect a chain to achieve a market position. For example, in all markets examined for this research no chain approached 65% of the market. It is reasonable to set $K = 65\%$, and this ceiling helps to anchor the upper portion of the s-curve. This arbitrary empirical constant should be subjected to sensitivity analysis in a real application. With more advanced statistical tools it would be possible for $K$ to be estimated endogenously at the cost of added complexity in the model, but $K$ in this paper is regarded as a constant numerical parameter.

In this model $a$ and $b$ are parameters or coefficients[2] that are estimated using the data in order to give a good fit between the model and the actual observations. (More on estimation later.) The $a$ and $b$ parameters might be estimated separately for each chain, in which case they should appear with a subscript "$j$". Thus, a firm with a steeper sloped curve would gain market share more rapidly with respect to changes in share of space than some other firm with a smaller slope.

Theoretically, the s-shaped curve has a few well-known properties (Mahajan et al., 1988). First, the curve has

---

[1] Chabot (1994) reports widely varying estimates of market share for chains as one of the concerns about using these kinds of data.

[2] Typically $b > 0$ and enters the model with a negative sign.