WORKING PAPER
ALFRED P. SLOAN SCHOOL OF MANAGEMENT

Econometric Diagnosis of Competitive Localization

Richard Schmalensee

WP# 1390-83   January 1983

MASSACHUSETTS
INSTITUTE OF TECHNOLOGY
50 MEMORIAL DRIVE
CAMBRIDGE, MASSACHUSETTS 02139

Econometric Diagnosis of Competitive Localization

Richard Schmalensee

WP# 1390-83   January 1983

Econometric Diagnosis of Competitive Localization

Richard Schmalensee*

Sloan School of Management
Massachusetts Institute of Technology

ABSTRACT

Competitive localization is present in an extreme form in one-dimensional spatial models; each brand faces only two rivals no matter how many brands are on the market. A test that does not require unrealistically long data series is proposed for detecting the presence of localized competition in markets with differentiated products, and techniques for assessing the nature and importance of such localization are presented. Application of these methods to data on the U.S. ready-to-eat breakfast cereal industry suggests their potential value in other contexts but, because of apparent data problems, yields no definite conclusions about that industry.

January 1983

Alfred ...

Sloan School of Management
Massachusetts Institute of Technology

ABSTRACT

Competitive localization is present in an environment in which additional central ...

1. Introduction

Two polar-case models of market demand and seller interaction emerged roughly simultaneously as economists began to study product differentiation systematically: the spatial model of Hotelling (1929) and the symmetric model usually associated with Chamberlin (1933). The spatial framework stresses buyer heterogeneity; additional brands make it more likely that any particular buyer will find a single one well-suited to his or her tastes. Symmetric models usually involve a representative buyer who is imagined to buy all brands and to benefit directly from increased variety. Both polar cases appear frequently in the modern literature: compare the spatial analyses of Salop (1979) and Schmalensee (1978) with the symmetric models of Spence (1976) and Dixit and Stiglitz (1977).

In the original Hotelling model, with sellers arrayed along a line in geographic or product space, rivalry is <u>localized</u> in the sense that each brand faces only a small number of direct competitors, no matter how many sellers are present in the market as a whole. This sort of localization can convert an apparently large-numbers situation into one of overlapping oligopolies, so that overall market concentration may be a seriously misleading indicator of the likelihood of non-competitive behavior. In symmetric models, on the other hand, each firm affects and is affected by all others in a symmetric fashion, so that the fact of product differentiation has no effect on the pattern of rivalry, and concentration may be measured on a market-wide basis.

While one-dimensional models of the Hotelling variety are useful for illustrating the implications of localization [as in Schmalensee (1978)], they build in an extreme form of this effect. As Archibald and Rosenbluth (1975) have shown, localization need not be important or even present in spatial models of higher dimensionality. Thus, even if one accepts the persuasive

arguments of Archibald, Eaton, and Lipsey (1982) that "address models," in which brands are identified by their locations in the space of possible products, generally provide the right framework for the study of differentiated markets, one is led to the conclusion that the importance of localization in any particular differentiated market is an empirical question. This essay is concerned with developing techniques for answering such questions. It is thus intended as a contribution to the growing literature on econometric analysis of individual markets.[1]

With enough data, of course, one could assess the importance of localization by examining the coefficients of estimated unrestricted brand-specific demand functions. In such functions, each brand's sales would depend on such things as the price, advertising, and lagged sales of all brands on the market, along with variables such as income that would influence total market demand. But this approach encounters a serious problem of dimensionality: if there are a substantial number of brands, each brand's equations will have many unknown parameters, and unrealistically long data series would be required to obtain reliable estimates. The existing literature in economics provides no alternative approaches that are noticeably less data-intensive.

Most of the relevant econometric work in marketing simply assumes a symmetric model.[2] On the other hand, marketers frequently construct "perceptual maps," based on questionnaire and other non-market data, that depict individual brands' locations in product space.[3] For a variety of reasons, including the difficulty of interpreting distances in such maps, these and related techniques are not generally useable by an economist interested in the overall importance of localization in any particular market.

The approach taken here sidesteps the dimensionality problem that plagues unrestricted brand-specific demand systems by focusing on the residual

covariance matrix obtained under a fairly general symmetric specification.[4]
That specification is presented in Section 2, and a large-sample likelihood-
ratio test for departures from symmetry is derived. Section 3 reports the
results of applying that test to data on the U.S. ready-to-eat breakfast
cereal industry, which is referred to simply at "the RTE cereal industry" in
what follows. The presence of localization was an important element of the
government's case in a recent antitrust proceeding involving that industry;
see Schmalensee (1978) for a summary of the argument. While the initial
decision in that proceeding found that localization "exists to a sufficient
degree to have the type of impact theorized by complaint counsel," the nature
of the available evidence compelled the finding that "the extent to which
localization prevails in the RTE cereal industry is an unknown" (Berman 1981,
p. 171). Techniques for assessing the pattern and importance of localization
are presented in Section 4, and their application to the RTE cereals data is
discussed. Our conclusions are briefly summarized in Section 5.

2. Testing for Localization

The objective of this section is to derive refutable implications of the
null hypothesis of no localization, that is, of perfect symmetry in
competitive interactions. Operationally, I interpret this hypothesis as
follows. Consider a market in which N brands of some product are sold.
Consider a change in an observable or unobservable variable that has the
direct effect of making brand i more attractive to buyers. Examples would be
an increase in i's advertising, a reduction in its price, a shift in buyer
preferences toward i, and (for normal goods) an increase in consumer income.
I interpret perfect symmetry to require that such changes not affect the ratio
of brand j's sales to brand k's sales, for all j, k $\neq$ i.[5] Changes in
relative sales would reflect substitution patterns derived from departures

from symmetry in the pattern of brands' locations in product space, and our null hypothesis rules out such departures.

Under perfect symmetry, changes in "market-wide" variables (such as consumer income) that affect the level of total market demand must leave all market shares unchanged. The implied multiplicative separability of brand-specific demand functions then allows us to neglect such variables when working with market shares and leads naturally to the use of what has been called an "attraction" model to determine those shares:

$$(1) \qquad M_i = \phi_i(X_i, u_i) / \sum_{j=1}^{N} \phi_j(X_j, u_j), \qquad i=1,\ldots,N,$$

where M_i is brand i's market share, ϕ_i is its non-negative "attraction", (the part of its demand function not involving "market-wide" variables), X_i is a vector of observable predetermined variables that directly affect the demand for brand i (such as its own price and advertising), and u_i is a disturbance term. The disturbances are most naturally interpreted as summarizing the effects of unobservable variables that directly affect each brand's attractiveness and that change during the sample period. This interpretation is central to our approach, as we restrict the u_i to affect market shares only through changes in the ϕ_i.

Functions like (1) relate market shares to prices in the symmetric representative- individual models of Spence (1976, Sect. 5) and Dixit and Stiglitz (1977) that employ generalized CES utility functions. (Not all symmetric utility functions are consistent with (1), of course.) Since consumers' tastes appear to differ, and all consumers do not buy all brands in most markets, a more plausible interpretation of (1) and of the null hypothesis of perfect symmetry is the following. As was noted above, in "address models" of high dimensionality, the pattern of buyer and brand

"addresses" need not imply the existence or importance of competitive localization. Precise conditions on the patterns of buyers' tastes and brands' attributes sufficient to rule out localization in such models are not known, however. Moreover, direct verification of any such conditions would almost certainly require both large quantities of data on individual brands and buyers and the use of strong maintained hypotheses. The argument here is that if the patterns of tastes and attributes in any particular market are such as to render localization absent or unimportant, an equation like (1) must provide a good approximation to aggregate behavior.

Bell, Keeney, and Little (1975) and Barnett (1976) provide alternative axiomatic justifications for attraction models, and the survey of Parsons and Schultz (1976, ch. 7) attests to their popularity in applied work on market shares. In addition to perfect symmetry, such models have the important property that predicted shares are always between zero and one and always sum to one.[6]

In order to go farther, it is necessary to make assumptions about the disturbance terms. It is convenient, standard, and not unusually restrictive to assume that the u_i are jointly normally distributed with zero means and that the ϕ_i may be written as $\phi_i(X_i)\exp(u_i)$. Under this last assumption, if we use brand N as the base brand, take logarithms of (1) and subtract, we obtain

$$(2) \quad \ln[M_i/M_N] = \ln[\phi_i(X_i)] - \ln[\phi_N(X_N)] + (u_i - u_N), \qquad i=1,\ldots,N-1.$$

If the ϕ_i have a finite number of unknown parameters, this system of equations can be jointly estimated by standard methods.[7] If the u_i follow autoregressive processes with the same coefficients, quasi-differencing can be used to eliminate serial correlation, and the coefficients of the

autoregressive process can be estimated along with the parameters of the ϕ_i.[8]
We make the usual assumption that all disturbance variances and covariances
are constant over time.

Suppose that the null hypothesis is correct and that a system like (2) has
been properly specified, so that the X_i include (at least) variables that
provide a good description of sellers' marketing activity. I now want to
argue that under those assumptions, $E[u_i(t)u_j(t)]$ must equal some constant,
z, for all $i \neq j$. The argument is just an application of the definition of
perfect symmetry given above to the <u>unobservable</u> determinants of brands'
attractions. If the null hypothesis is correct and the demand system is
properly specified, the unobservable determinants must reflect either industry-
wide or brand-specific factors. (If the system is properly specified, all
important marketing variables are included in the X_i, so that significant
unobservable firm-wide marketing changes by multi-brand firms are ruled out.
Taste changes that affect subsets of the set of brands are ruled out by the
null hypothesis of no localization.) Under our definition of perfect symmetry,
an increase in u_i caused by a change in either type of factor cannot be
associated with a change in $(u_j - u_k)$ in (2) for any $j \neq i \neq k$, since
otherwise the relative sales of brands j and k would be affected. But this is
easily seen to require that the covariance of u_i and u_j be identical for
all i and j, as asserted.

The preceeding argument implies testable restrictions on the covariance
matrix of the reduced-form disturbances in (2). If Σ is that matrix, with
typical element σ_{ij}, and v_i is the variance of u_i, the restrictions
are the following:

$$
(3) \qquad \sigma_{ij} = \begin{cases} (v_i - z) + (v_N - z), & i=j \\ \\ (v_N - z), & i \neq j \end{cases}, \qquad i,j=1,\ldots,N-1.
$$

By considering the covariance matrix of the u_i, it is easy to see that z can exceed at most one of the v_i if they are all distinct. This constraint is difficult to impose, however. We lose little generality by requiring that all the v_i exceed z, so that we can simplify notation in what follows by setting z to zero.

We can test for departures from perfect symmetry by estimating a system like (2) with an unconstrained contemporaneous disturbance covariance matrix and then using the residual covariance matrix to test for the validity of the restrictions given by (3). We employ the asymptotic distribution of a likelihood ratio statistic, so that the test has only the usual large-sample properties. We lose none of those properties by treating the parameters of the ϕ_i as known in deriving the test. The residual covariance matrix, S, with typical element s_{ij}, is thus treated in what follows as having been generated by observations from a multivariate normal distribution with covariance matrix Σ and zero mean vector.[9]

Under this assumption, the $(N - 1) \times (N - 1)$ matrix S has the Wishart distribution, and the log-likelihood function is the following:

$$(4) \quad L = C - (T/2)\ln|\Sigma| + [(T - N)/2]\ln|S| - (T/2)\mathrm{tr}(\Sigma^{-1}S),$$

where C is a constant and T is the number of observations. (See, for instance, Rao (1973, p. 398). Rao uses S to denote the dispersion matrix, which equals T times the sample covariance matrix.) Given S, the unrestricted maximum of L is obtained by setting $\hat{\Sigma} = S$, and the maximum value is

$$(5) \quad L_u = C - (N/2)\ln|S| - T(N - 1)/2.$$

In order to maximize L subject to the restrictions given by (3) with z = 0, we make use of the fact that the following are valid if those restrictions

hold (Rao 1973, pp. 32-3):

$$(6a) \qquad |\Sigma| = (\prod_{i=1}^{N-1} v_i)(1 + v_N D), \text{ where}$$

$$(6b) \qquad D = \sum_{i=1}^{N-1} (1/v_i); \text{ and}$$

$$(6c) \qquad [\Sigma^{-1}]_{ij} = \delta_{ij}/v_i - v_N/[v_i v_j(1 + v_N D)],$$

where $\delta_{ij} = 1$ if $i = j$ and is zero otherwise. Using (6), it is straight-forward but tedious to show that the first-order conditions for maximizing L with respect to the estimates $\hat{v}_i$ are the following:

$$(7a) \quad \hat{v}_i = s_{ii} + \hat{v}_N[1 - 2(\hat{D}\hat{F}_i/\hat{F})], \qquad\qquad i=1,\ldots,N-1, \text{ and}$$

$$(7b) \quad \hat{v}_N = (\hat{F} - \hat{D})/\hat{D}^2, \text{ where}$$

$$(7c) \quad \hat{F}_i = \sum_{j=1}^{N-1} s_{ji}/\hat{v}_j, \qquad\qquad i=1,\ldots,N-1,$$

$$(7d) \quad \hat{F} = \sum_{i=1}^{N-1} \hat{F}_i/\hat{v}_i,$$

and $\hat{D}$ is given by (6b) as a function of the $\hat{v}_i$. These equations may be solved iteratively for the maximum likelihood estimates. The constrained maximum of L can be shown to be equal to

$$(8) \qquad L_r = C - (t/2)\ln|\hat{\Sigma}| + [(T + N)/2]\ln|S| - T(N - 1)/2,$$

where $|\hat{\Sigma}|$ is given as a function of the $\hat{v}_i$ by (6a) and (6b).

The test statistic is then just

$$(9) \qquad \lambda = -2(L_r - L_u) = T \ln(|\hat{\Sigma}|/|S|).$$

The unrestricted maximum involves $N(N - 1)/2$ parameters, while the restricted maximum involves N parameters, so that λ is asymptotically distributed as χ^2 with $N(N - 3)/2$ degrees of freedom.

In order to apply this test, one must complete the specification of (2) by making assumptions about the X_i and the ϕ_i. This of course gives rise to a classic problem: if the test described above rejects the null hypothesis, it may be signalling mis-specification of the particular symmetric model employed, rather than a true violation of symmetry. Examination of alternative specifications and use of the diagnostic techniques presented in Section 4 should help one choose between these alternatives, as we illustrate below.

3. An Application to the RTE Cereals Industry

There are two reasons for attempting to test for localization in the RTE cereals industry. First, as was mentioned in the introduction, the existence and importance of localization in that industry were issues in a recent antitrust proceeding. Second, in the course of that proceeding, the U.S. Federal Trade Commission (FTC) obtained from the leading sellers monthly brand-specific data of the sort that would appear suitable for the application of the procedure described in Section 2. Those data are used here; they have also been employed by Aaker, Carman, and Jacobson (1982), who provide additional descriptive information.

Although some monthly series for some firms run from 1957 through 1973, only for the 40-month period from September, 1969 through December, 1972 do the data contain sales (factory shipments), wholesale list prices, and advertising outlays for the four largest producers, along with some information on the other two national sellers.[10] These six firms, which accounted for over 97 per cent of RTE cereal sales in 1970 (Berman 1981, p.

64), had over 70 brands in national distribution during this period. It would clearly not be possible to examine localization by estimating an unrestricted demand system involving all brands on the market. Even system estimation of our symmetric model (2) using all brands is computationally infeasible. Accordingly, we chose to work exclusively with the 11 leading brands listed in Table 1. These were the only brands that individually accounted for at least 2.0 per cent of industry dollar sales in each of the four years 1969-1972. As Table 1 indicates, these brands in aggregate accounted for just under half the six leading firms' total RTE cereal revenues during this period. Our sample includes at least one brand offered by each of the four largest producers.

If equation (1) holds in some market, and the additional assumptions that led to equations (2) are correct, one can employ (2) using any subset of the brands on the market. For purposes of testing for localization, we can thus treat these 11 brands as if they were the whole market. Doing this naturally limits our ability to make definite statements about patterns of localization if the null hypothesis of perfect symmetry is rejected, however. Our choice of 11 brands reflects a tradeoff between comprehensive coverage and computation cost.

Given the form of equations (2), it is no surprise that it is standard in the relevant literature to assume that the ϕ_i are Cobb-Douglas functions, and we adopt that assumption. As Case (1981) demonstrates, if shares are to be invariant to equiproportional changes in elements of the X_i across brands (a doubling of all prices, for instance), the ϕ_i must be Cobb-Douglas with equal exponents. This form emerges in the symmetric CES models of Spence (1976, Sect. 5) and Dixit-Stiglitz (1977). Naert and Weverbergh (1981) note that relaxing assumption of equal exponents often causes numerical problems akin to those caused by multicollinearity, and we encountered such problems

Table 1

Leading RTE Cereal Brands Used in Empirical Analysis

| | | | Average Percentage Share | |
| | | | Industry Dollar Sales | Leaders' Pound Sales |
Brand	Company	Brand Name		
1	Kellogg	Kellogg's Corn Flakes	6.0	18.2
2	General Mills	Cheerios	7.8	13.9
3	Kellogg	Super Frosted Flakes	5.9	13.2
4	Kellogg	Rice Krispies	7.2	12.3
5	General Mills	Wheaties	3.6	7.8
6	Kellogg	Special K	5.0	7.5
7	Kellogg	Kellogg's Raisin Bran	2.6	6.8
8	Quaker Oats	Cap'n Crunch	3.2	6.4
9	General Foods	Post's Raisin Bran	2.2	5.7
10	General Mills	Total	2.7	4.3
11	Kellogg	Froot Loops	2.4	4.0
		Totals:	48.6	100.0

Notes:

1. Dollar sales averages cover the years 1969 - 1972; pound sales averages cover the period October, 1969 through November, 1972.

with these data. (We were in fact unable to obtain convergence for specifications that relaxed this assumption.) Accordingly, we also make the assumption of equal exponents across brands and specify the ϕ_i as follows:

$$(10) \quad \ln[\phi_i(t)] = \alpha_i + \beta \ln[A_i(t)] + \gamma \ln[P_i(t)] + \delta \ln M_i[(t - 1)],$$

$$i=1,\ldots,N,$$

where A is advertising spending, P is wholesale price, and M is market share in pounds of cereal sold. The lagged share term represents an attempt to capture the effects of buyer inertia. (This is algebraically equivalent here to the use of lagged pound sales.) In addition, we assume that the $u_i(t)$ in (2) follow a first-order autoregressive process with autocorrelation coefficient ρ.[11] Finally, given firms' information and decision-making lags, it seems plausible to treat both price and advertising as predetermined in monthly data. (See Schmalensee (1972) on firms' advertising decisions.)

Differences in the α_i across brands may be interpreted as reflecting differences in unobservable determinants of brands' attractions that do not vary during the sample period. As we are interested in changes in quantities demanded, it is appropriate to work with shares of sales measured in pounds. Changes in shares then reflect movements in quantities sold, even though one might argue that levels of shares so measured have doubtful significance. (Pound sales for General Mills were constructed using price and dollar sales series; sales in pounds were reported for the other firms.) We used the reported wholesale list prices for the $P_i(t)$. (A pound-weighted average of list prices for several sub-brands was used for Cap'n Crunch.) Although these firms generally did not depart from list price during most of the 1960's, "trade deals," short-term discounts below list price, were used with increasing frequency during our sample period (Berman 1981, pp. 106-13). This

is a potentially important source of specification error, as such "deals" cannot realistically be treated as exogenous, random events, but data to construct a series of transactions prices do not exist.

The reported advertising series exhibited large month-to-month variations. (The median coefficient of variation in the 11 advertising series was 0.41; these coefficients were above 0.20 for all brands and above 0.30 for all but one. In contrast, only one of the 11 coefficients of variation for pound sales exceeded 0.30.) More surprisingly, there were two negative advertising numbers in our data. (These were for brands 9 and 10.) As Aaker, Carman, and Jacobson (1982) discuss, these features of the advertising series undoubtedly reflect billing and accounting practices that cause short-term timing differences between the appearance of ads and the manufacturer's payment for them. In order to smooth the series and, it was hoped, eliminate the effects of random timing differences, we applied a moving average filter, with weights of 0.2 on the next month and the last month and 0.6 on the current month. These weights were chosen in part so as to eliminate the negative observations. Use of this filter, along with the presence of serial correlation and lagged dependent variables in our specification, reduced the number of observations to 37.

As we noted above (footnote 6), the choice of a base brand [brand N in equations (2)] is arbitrary. Table 2 presents coefficients and standard errors obtained by joint nonlinear generalized least squares using the largest and smallest brands (in terms of average pound sales) as the base. (We employed the ZELLNER option in the GREMLIN package on the TROLL system at MIT.) The coefficients using the largest brand (Kellogg's Corn Flakes) as a base seem reasonable, but five of the 10 individual R^2 statistics are negative. Only one R^2 is negative when the smallest brand (Kellogg's Froot

Loops) is used as the base, but the coefficient estimates are much less satisfactory. (As the dependent variables differ, one cannot compare R^2 values between the two sets of estimates.) The sensitivity of coefficient estimates to the choice of the base brand is troubling. It strongly suggests the possibility of mis-specification. (See footnote 7, above.)

Table 2 also reports the results of the likelihood ratio test for localization presented in Section 2.[12] It is interesting to note that the estimated disturbance variances do not differ much between the two sets of estimates. Also, the estimated variances for the six Kellogg's brands (numbers 1, 3, 4, 6, 7, and 11) are much smaller than for the other five brands. The test statistic, λ, defined by equation (9) is distributed as χ^2 with 44 degrees of freedom under the null hypothesis. The values shown in Table 2 are thus both highly significant. (The 0.5 per cent critical level is 72.)

Taking these test results at face value, one would be compelled to reject the hypothesis of perfect symmetry and accept the presence of localization. But the quality of our estimates and the data problems discussed above strongly suggest the possibility of mis-specification, so that caution is called for at this stage.[13]

4. Assessing Patterns and Importance of Localization

As was noted in the introduction, the least restrictive approach to assessing departures from symmetry involves simply estimating unrestricted brand-specific demand functions. But this usually requires unrealistically long data series, and it is not feasible with the particular data set used here. We can make some headway if we are willing to retain the assumption that market shares depend only on the $\Phi_i(X_i, u_i)$ even if (1) does not hold. In such a generalized attraction model, the pattern of competitive

Table 2

Estimation and Test Results for RTE Cereals Data

| Estimate of | Base Brand Employed | |
	1 (= Corn Flakes)	11 (= Froot Loops)
β	0.062 (0.013)	−0.003 (0.006)
γ	−0.253 (0.161)	−1.335 (0.250)
δ	0.367 (0.053)	−0.046 (0.021)
ρ	0.113 (0.037)	0.648 (0.033)
v_1	0.017	0.023
v_2	0.115	0.134
v_3	0.007	0.007
v_4	0.005	0.005
v_5	0.112	0.124
v_6	0.002	0.003
v_7	0.003	0.004
v_8	0.133	0.126
v_9	0.081	0.081
v_{10}	0.106	0.115
v_{11}	0.012	0.007
λ	204.1	203.8

Notes:

1. Figures in parentheses are standard errors.

effects is assumed to be the same for all brand-specific marketing variables, like price and advertising. While this is of course somewhat restrictive, it does not seem generally implausible. An asymmetric generalized attraction model seems a reasonable reduced-form representation of an "address model" with localization, in which the effects of changes in the Φ_i on the relative sales of rival brands reflect and summarize the patterns of brand and buyer locations in product space.

In such a model, the development in Schmalensee (1982, Sect. 2) suggests that the following can be used to analyze the nature and importance of localization:

(11) $K_{ij} = -[(1 - M_i)(\partial M_j/\partial \Phi_i)]/[M_i(\partial M_i/\partial \Phi_i)], \quad i=j; \; i,j=1,\ldots,N.$

In the symmetric model, equation (1), it is easy to see that the K_{ij} all equal unity. In general, the adding-up restriction on the M_i requires share-weighted averages of the K_{ij} to equal unity;

(12) $\sum_{j \neq i} K_{ij} M_j / \sum_{j \neq i} M_j = 1, \quad\quad\quad i=1,\ldots,N.$

If one of the K_{ij} exceeds unity, it implies that increases in brand i's share come more at the expense of brand j than a symmetric model would predict. Direct examination of the K_{ij} can thus yield information on patterns of localization.

To summarize the importance of localization across the market as a whole, Schmalensee (1982, Sect. 2) proposes using the following generalization of the standard H-index of concentration:

(13a) $H^* = \sum_{i=1}^{N} (M_i)^2/G_i,$ \quad\quad\quad where

(13b) $G_i = M_i + (1 - M_i)^2/[\sum_{j \neq i} (K_{ij})^2 M_j], \quad i=1,\ldots,N.$

It is easy to see that H* is bounded between the H-index and unity and that it attains its lower bound only in the symmetric case in which the K_{ij} all equal one. The ratio of H to H* thus provides a measure of the importance of localization that is bounded between zero and one:

$$(14) \qquad\qquad G^* = (1/H^*) \sum_{i=1}^{N} (M_i)^2.$$

Low values of G* signal important localization of competitive interaction; G* is equal to one only in the perfectly symmetric case.

Let us now consider three approaches to estimating the K_{ij} and G* that economize on data. All three involve using statistics derived from estimation of equations (2), which are mis-specified if localization is present. These approaches can thus yield diagnostic information but not rigorously defensible estimates. (All this is in the spirit of the diagnostic techniques presented by Phlips (1974, pp. 212-17).) The first two techniques employ residual covariances from a symmetric model to provide information on the localization of the effects of the unobservable determinants of brands' attractions. Under the generalized attraction assumption, the same pattern of localization applies to both observable and unobservable variables.

The first and simplest technique works with the (singular) covariance matrix of the $e_{it} = M_{it} - \hat{M}_{it}$, the differences between the actual shares and those predicted by the symmetric model. Let this matrix be S', with elements s'_{ij}. Then use $\{E[(e_{jt} - \bar{e}_j)|(e_{it} - \bar{e}_i) = c]/c\}$, where c is any constant, as an estimate of the ratio of derivatives in (11). The rationale is that this estimator provides a rough answer to the following question: Given an increase in brand i's share caused by unobservables, how much of that increase is expected to come at the expense of brand j's share? A problem with this non-structural approach is that we cannot associate increases

in e_{it} directly with increases in brand i's own attraction. Under

normality, and treating the $\bar{e}_i$ as known, our estimator is just

(s_{ij}'/s_{ii}'), and the estimated K_{ij} are obtained by substitution in (11).

We applied this technique to the RTE cereals data, using the estimates

reported in Section 3 and the sample mean values of the M_i. For both sets

of estimates, we obtained a surprisingly large number of negative estimated

K_{ij}. One might expect to obtain negative estimates of these parameters if

exogenous, unmeasured shocks affected the attractiveness of groups of close

substitute brands as against all others. For example, large shifts in buyers'

preferences for raisins might be expected to cause the sales of brands 7 and 9

(the two raisin brans) in our sample to move together. If one could identify

such influences during the sample period, and if the pattern of negative K_{ij}

made sense in light of those influences, one might be led to conclude that

$K_{ij} < 0$ implies an unusually close substitute relation between brands i

and j. This would be informative, even though one could not sensibly employ

the overall localization measure $G*$ in the presence of negative estimates of

the K_{ij}.

Unfortunately, no such exogenous group-specific disturbances are apparent

here. Moreover, the pattern of negative values actually observed seems to

point clearly to specification and data problems. For both sets of estimates,

if brands i and j are either both Kellogg brands or both non-Kellogg brands,

K_{ij} is always negative, while if i is a Kellogg brand and j is not, K_{ij} is

always positive. (This same pattern generally holds when simpler

specifications are used to generate predicted shares.) It does not make much

sense to think of this as reflecting localization, given that both Kellogg and

non-Kellogg sets contain brands in most of the "segments" discussed by

industry observers. (Brands 3, 8, and 11 are pre-sweetened, brands 7 and 9

are raisin brans, brands 6 and 10 are sold on the basis of nutrition, and so on.) It seems much more likely that we have either failed to observe important marketing variables (such as the departures from list price noted in Section 3) that affect all Kellogg brands together or that there are differences in accounting or reporting practices between Kellogg and the other firms that give rise to spurious movements in measured relative sales.[14]

Not only do these results indicate that we cannot use our data to say anything reliable about the patterns or importance of localization in the RTE cereal industry, they also indicate that the test results reported in Section 3 cannot be taken at face value as reflecting the presence of localization. Examination of the K_{ij} makes it clear that the null hypothesis of symmetry was rejected mainly because of positive residual covariances within the Kellogg and non-Kellogg sets. If, as we have just argued, those covariances reflect either important omitted variables or serious measurement problems, the necessary maintained hypothesis that the symmetric model was properly specified must be set aside.

The second and third techniques for the assessment of localization are based on the following generalized attraction model:

(15a) $$M_i = \Psi_i / \sum_{j=1}^{N} \Psi_j, \qquad i=1,\ldots,N, \text{ where}$$

(15b) $$\ln(\Psi_i) = \sum_{k=1}^{N} \theta_{ik} \ln(\Phi_k), \qquad i=1,\ldots,N.$$

Differentiation of (15) and substitution into (11) yield

(16a) $$K_{ij} = - [(1 - M_i)O_{ji}]/[M_i\Theta_{ii}], \qquad i=j; \ i,j=1,\ldots,N, \text{ where}$$

(16b) $$O_{ij} = \theta_{ij} - \sum_{k=1}^{N} M_k\theta_{kj}, \qquad i,j=1,\ldots,N.$$

Let ϵ_{it} be the difference in period t between the actual and predicted values of $\ln(M_{it})$, using the symmetric model. (A simple variant of this technique uses prediction errors in the M_{it} themselves.) The generalized model (15) then suggests the following log-linear specification:

$$(17) \qquad \epsilon_{it} = \sum_{j=1}^{N} r_{ij} u_{jt}, \qquad\qquad i=1,\dots,N; \; t=1,\dots,T.$$

where u_{it} summarizes the effects of unobservable determinants of brand i's attraction in period t, as before. Following the discussion in Section 2, if the symmetric model is properly specified, it is reasonable to assume that the u_{it} are uncorrelated across brands. Because of the homogeneity of this expression, and because we are only concerned with ratios of the form (r_{ji}/r_{ki}), we can set the variances of the u_{it} to unity with no loss of generality. Let R be the matrix of the r_{ij}. It is natural to choose the signs of the elements of R so that positive values of u_i are associated with positive values of ϵ_i. This is most easily done by requiring R to be non-negative definite. Under these assumptions, the covariance matrix of the ϵ_i is just RR', and one can estimate R as the unique non-negative definite square root of the sample covariance matrix of the ϵ_i. Equations (16), with r_{ij} replacing θ_{ij} everywhere then yield estimates of the K_{ij}. The results reported just above seem so strong as to make systematic application of this approach to the RTE cereals data rather pointless, however.[15]

Finally, if estimation of (2) yields plausible estimates of the parameters of the ϕ_i, it may be instructive to construct and employ a series of estimated attractions, $\bar{\phi}_{it}$, for each brand. Applying the transformation discussed by Bultez and Naert (1975) to (15), one obtains

$$(18) \quad \ln(M_i) - \sum_{k=1}^{N} M_k \ln(M_k) = \sum_{j=1}^{N} O_{ij} \ln(\phi_i), \qquad i=1,\dots,N,$$

where the O_{ij} are defined by (16b). These last quantities are not in
general constant over time, of course. But if the variation in shares is not
great, substituting ϕ_i for Φ_i in (18), adding an intercept term, and
running a set of N linear regressions should yield servicable estimates of the
average values of the O_{ij} that can be used in (16a).[16] The quality and
instability of the estimates reported in Section 3 would also seem to rule out
serious use of this technique with the available RTE cereals data.

5. Conclusions

Section 2 presented an approach to testing for the presence of localized
competition that can be implemented without enormous quantities of data. In
order to produce such an approach, it was necessary to make assumptions both
about functional form and about the relative unimportance of individual
unobserved determinants of brand sales. Application of this approach to data
on the RTE cereals industry in Section 3 strongly rejected the null hypothesis
of perfect symmetry, but the estimates on which that rejection was based were
unsatisfactory in several respects. Section 4 presented techniques for
assessing the nature and importance of localization that also economized on
data requirements. Application of the simplest of those techniques to the RTE
cereals data clearly demonstrated the importance of specification or
measurement problems that apparently make it inappropriate to interpret
Section 3's test results as providing a refutation of the null hypothesis of
no localization. Those results seem rather to be driven by inadequacies of
the available data.

Our techniques for diagnosis of localization in market analysis thus fail
to produce positive results when applied to the FTC's RTE cereals data. But
there is every reason to suspect that they or refinements of them will be
useful in the diagnosis of localization in other contexts. (After all, the

RTE cereal firms responded in a presumably uncoordinated fashion to a broadly-worded request for data from the FTC, which was in the process of bringing a major antitrust case against them. Other approaches to data collection should yield systematically better results.) It is also encouraging that a straightforward application of the diagnostic techniques presented in Section 4 was able to detect serious underlying weaknesses in the publicly-available data on this interesting industry, weaknesses that might otherwise have gone unnoticed. These techniques should be of value even if they are only used to assess the validity of symmetric specifications of the sort that are widely used in market share analysis.

References

Aaker, D.A., Carman, J.M. and Jacobson, R. (1982), "Modeling Advertising-Sales Relationships Involving Feedback: A Time Series Analysis of Six Cereal Brands," Journal of Marketing Research, 19 (February), 116-25.

Archibald, G.C. and Rosenbluth, G. (1975), "The 'New' Theory of Consumer Demand and Monopolistic Competition," Quarterly Journal of Economics, 89 (February), 569-90.

Archibald, G.C., Eaton, B.C., and Lipsey, R.G. (1982), "Address Models of Value Theory," Presented at the International Economics Association Conference on New Developments in Market Structure, Ottawa, May 1982.

Barnett, A.I. (1976), "More on a Market Share Theorem," Journal of Marketing Research, 13 (February), 104-09.

Bell, D.E., Keeney, R.E. and Little, J.D.C. (1975), "A Market Share Theorem," Journal of Marketing Research, 12 (May), 136-41.

Berman, A.L. (1981), "Initial Decision in the Matter of Kellogg, et al., Docket No. 8883," Washington: U.S. Federal Trade Commission, mimeographed, September 1.

Berndt, E.R. and Savin, N.E. (1975), "Estimation and Hypothesis Testing in Singular Equation Systems with Autoregressive Disturbances," Econometrica, 43 (September-November), 937-57.

Bresnahan, T.F. (1981), "Identification of Market Power," Stanford University, Workshop on Factor Markets, Research Paper No. 15, October.

Bultez, A.V. and Naert, P.A. (1975), "Consistent Sum-Constrained Models," Journal of the American Statistical Association, 70 (September), 529-35.

Case, J. (1982), "Ad-Response Models and Early Entry," in Keon, J.W. (ed.) Proceedings of the Third ORSA/TIMS Conference on Marketing Measurement and Analysis (Providence: The Institute of Management Science).

Chamberlin, E.H. (1933) The Theory of Monopolistic Competition, 1st Ed.
     (Cambridge: Harvard).

Dixit, A.K. and Stiglitz, J.E. (1977), "Monopolistic Competition and Optimum
     Product Diversity," American Economic Review, 67 (June), 297-308.

Hauser, J.R. and Koppelman, F.S. (1979), "Alternative Perceptual Mapping
     Techniques: Relative Accuracy and Usefulness," Journal of Marketing
     Research, 16 (November), 495-506.

Hotelling, H. (1929), "Stability in Competition," Economic Journal, 39
     (March), 41-57.

Huber, J. and Holbrook, M.B. (1979), "Using Attribute Ratings for Product
     Positioning: Some Distinctions Among Compositional Approaches," Journal
     of Marketing Research, 16 (November), 507-16.

Jeuland, A.P. (1980), "Empirical Investigation of Price and Advertising
     Competition Using a Market Share Model," in Montgomery, D.B. and Wittink,
     D.R. (eds.) Market Measurement and Analysis (Cambridge: Marketing Science
     Institute).

Naert, P. and Weverbergh, M. (1981), "On the Prediction Power of Market Share
     Attraction Models," Journal of Marketing Research, 18 (May), 146-53.

Parsons, L.J. and Schultz, R.L. (1976) Marketing Models and Econometric
     Research (Amsterdam: North-Holland).

Phlips, L. (1974) Applied Consumption Analysis (New York: Elsevier).

Rao, C.R. (1973) Linear Statistical Inference and Its Applications, 2nd Ed.
     (New York: Wiley).

Salop, S.C. (1979), "Monopolistic Competition with Outside Goods," Bell
     Journal of Economics, 10 (Spring), 141-56.

Schmalensee, R. (1972) The Economics of Advertising (Amsterdam:
     North-Holland).

Schmalensee, R. (1978), "Entry Deterrence in the Ready-to-Eat Breakfast Cereal Industry," Bell Journal of Economics, 9 (Autumn), 305-27.

Schmalensee, R. (1982), "The New Industrial Organization and the Economic Analysis of Modern Markets," pp. 253-85 in Hildenbrand, W. (ed.) Advances in Economic Theory (Cambridge: Cambridge University Press).

Spence, A.M. (1976), "Product Selection, Fixed Costs, and Monopolistic Competition," Review of Economic Studies, 43 (June), 217-35.

Theil, H. (1975) Theory and Measurement of Consumer Demand, Vol. 1 (New York: Elsevier).

Footnotes

1. See Bresnahan (1981) for a valuable survey and bibliography.

2. Parsons and Schultz (1976, ch. 7) provide a survey; see Naert and Weverbergh (1981) for a recent example.

3. For recent overviews of this work, see Hauser and Koppelman (1979) and Huber and Holbrook (1979).

4. Our approach is in the spirit of the "second moment model" of Theil (1975), especially as regards testing for deviations from symmetry, and of the diagnostic techniques presented by Phlips (1974, ch. 8).

5. While this requirement has a strong formal resemblance to the "independence of irrelevant alternatives" assumption in choice theory, it should become clear below that I am not assuming that individual behavior conforms to this assumption.

6. On the implications of these restrictions, see Berndt and Savin (1975), Bultez and Naert (1975), Naert and Weverbergh (1981), and the references they cite.

7. It seems somewhat more common in applied work to subtract the average of all log-shares from each of the N logged equations and then delete one of the resulting equations: see Bultez and Naert (1975), Naert and Weverbergh (1982), and equation (18), below. (Maximum likelihood estimates are invariant to the equation deleted.) While this approach

appears more "balanced" than that employed here, neither approach seems theoretically superior. Since the choice of a base in such transformations is essentially arbitrary, variation in parameter estimates with the base brand employed provides an interesting, if informal, check on the demand specification. (This is explored further in Section 3.)

8. Berndt and Savin (1975) provide a general treatment of the restrictions on vector autoregressive disturbance processes implied by the adding-up constraint.

9. Theil (1975, pp. 318-20) follows this same approach in a very similar context. Jueland (1980) deals with estimation of (2) subject to the restrictions in (3), with $z = 0$. (He uses a consistent estimator of Σ that is not maximum-likelihood.) One could employ such constrained estimates in a test procedure, but this would complicate computation and yield no asymptotic benefits. The small sample gain, if any, from such an approach is of course unknown.

10. The four largest firms, who were initially charged in the FTC's antitrust action, are those appearing in Table 1. Quaker was later dropped from the case. The other two national sellers were Nabisco and Ralston.

11. Because our data series were so short and seasonal patterns in RTE cereal consumption did not appear to be important, we did not experiment with 12th-order (seasonal) autoregressive processes.

12. The first-order conditions were solved using the Gaus-Seidel algorithm in the SIMULATE package on the TROLL system at MIT. Consistent estimates were used as starting values, and no evidence of multiple solutions was encountered.

13. The diagnostic results reported below strongly suggested the futility of experimentation with variants of the specification reported here. (We did

perform localization tests using raw covariance matrices of differences in log-shares and simple autoregressive models; perfect symmetry was always decisively rejected.) Without those results, it would have been natural to explore the use of dollar shares (as in the trans-log literature) and to experiment with the alternative timing assumption used by Aaker, Carman, and Jacobson (1982). (They found surprising weak brand-specific advertising-sales relationships for a set of Kellogg brands.)

14. These results suggest that data problems may go a long way toward explaining the surprisingly weak bivariate associations between advertising and sales found by Aaker, Carman, and Jacobson (1982) for Kellogg brands.

15. Since we lacked confidence in the coefficient estimates reported in Section 3, we applied this technique with deviations from means replacing prediction errors. All but 3 of the 50 K_{ij} elements above the main diagonal conformed to the sign pattern described above. (The matrix of K_{ij} is sign-symmetric.)

16. In constructing the $\hat{\phi}_i$ under specifications like (10), an arbitrary positive value of α_N must be chosen. The addition of an intercept term to (18) serves to compensate for errors in this choice. Since all N equations have the same right-hand side variables, there is no efficiency gain from system estimation here. With better data, of course, one could attempt to estimate all of the parameters of models like (15) directly.