Econometric Diagnosis
of
Competitive Localization
by
Richard L. Schmalensee

WP# 1555-84 April 1984
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 a simple technique for assessing the nature and importance of such localization is presented. Applications of these methods to data on the U.S. ready-to-eat breakfast cereal industry suggests their potential value in other contexts.

Revised: April 1984
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 localized 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, 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.

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 the relevant econometric work in marketing simply assumes a symmetric model of the sort employed below. On the other hand, marketers commonly construct "perceptual maps" that depict individual brands' locations in product space. For a variety of reasons, including the difficulty of interpreting distances in such maps, these and related techniques are not generally usable 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. That specification is presented in Section 2, and a large-sample likelihood-ratio
test for departures from symmetry is derived. A simple technique for assessing the pattern and overall importance of localization is presented in Section 3. Section 4 describes the application of this approach to data on the U.S. ready-to-eat breakfast cereal industry.

2. Testing for Localization

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 the absence of localization (i.e., 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 ≠ i. Changes in relative sales would reflect substitution patterns derived from departures from symmetry in the pattern of brands' locations in product space, and the null hypotheses of no localization 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 function 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:

\[
M_i = \Phi_i(X_i, u_i) / \sum_{j=1}^{n} \Phi_j(X_j, u_j), \quad i=1,\ldots,N,
\]

where \( M_i \) is brand i's market share, \( \Phi_i \) 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 $i$, and $u_i$ is a disturbance term. If (1) is correctly specified, the $X_i$ must include variables that provide a complete description of sellers' marketing activity. 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 $\Phi_i$.

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) should provide a good approximation to aggregate behavior. (Note that such models have the important property that predicted shares are always between zero and one and always sum to one.\(^5\))
In order to go farther, it is necessary to make assumptions about the disturbance terms. It is convenient and not unusually restrictive to assume that the $u_i$ are jointly normally distributed with zero means and that the $\phi_i(X_i, u_i)$ 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:

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

If the $\phi_i$ have a finite number of unknown parameters, this system of equations can be jointly estimated by standard methods. 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 $\phi_i$.

If sufficient data are available, one can test the null hypothesis of no localization by using (2) to test the restrictions that only $\phi_i$ and $\phi_N$ affect $(M_i/M_N)$, $i = 1,...,N-1$. Though such a test is simple to describe, it is likely to involve a heavy computational burden and unlikely to have much power when $N$ is large.

An alternative, simpler test flows from the fact that the assumptions above imply that $E[u_i(t)u_j(t)]$ must be equal for all $i \neq j$. The argument is just an application of the definition of perfect symmetry given above to the unobservable determinants of brands' attractions. If there is no localization 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.

Let \( \Sigma \) be the \((N-1)x(N-1)\) covariance matrix (assumed constant over time) of the disturbances in (2) with typical element \( \sigma_{ij} \). Let \( V_i \) be the variance of \( u_i \), let \( Z \) be the covariance between \( u_i \) and \( u_j \) for all \( i \neq j \), and let \( v_i = V_i - Z \), for \( i = 1, \ldots, N \). Then the argument of the preceding paragraph implies that \( \sigma_{ii} = v_i + v_N \) and \( \sigma_{ij} = v_N \) for \( i \neq j \). By considering the covariance matrix of the \( u_i \), it is easy to see that at most one of the \( v_i \) can be negative if they are all distinct. This constraint is difficult to impose, however. We lose little generality by requiring that all the \( v_i \) be non-negative.

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 the usual large-sample properties. We lose none of those properties by treating the parameters of the \( \phi_i \) 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 \( \Sigma \) and zero mean vector.

Under this assumption, the \((N-1)x(N-1)\) matrix \( S \) has the Wishart distribution, and the log-likelihood function is the following:
(3) \[ L = C - \frac{T}{2}\ln |\Sigma| + [(T-N)/2]\ln |S| - \frac{T}{2}\text{tr}(\Sigma^{-1}S), \]
where \( C \) is a constant and \( T \) is the number of observations. The unrestricted maximum of \( L, L_u \), is obtained by setting \( \Sigma = S \).

It is straightforward (but extremely tedious) to show that the first-order conditions for a restricted maximum of \( L \) with respect to the estimates \( \hat{v}_i \) under the null hypothesis are the following:

\[
(4a) \quad \hat{v}_i = s_{ii} + \hat{v}_N[1 - 2(\hat{F}_i/F)], \quad i=1,...,N-1, \text{ and }
\]

\[
(4b) \quad \hat{v}_N = (\hat{F} - \hat{D})/\hat{D}^2, \quad \text{where}
\]

\[
(4c) \quad \hat{F}_i = \sum_{j=1}^{N-1} s_{ji}/\hat{v}_j, \quad i=1,...,N-1
\]

\[
(4d) \quad \hat{F} = \sum_{i=1}^{N-1} \hat{F}_i/\hat{v}_i, \text{ and }
\]

\[
(4e) \quad \hat{D} = \sum_{i=1}^{N-1} (1/\hat{v}_i).
\]

These equations may be solved iteratively. (A consistent starting point may be obtained by taking \( \hat{v}_N \) as the average of the off-diagonal elements of \( S \).) Substitution into (3) gives the restricted maximum of the log-likelihood function, \( L_r \).

The likelihood-ratio statistic for testing the null hypothesis of no localization is then just

\[
(5a) \quad \lambda = -2(L_r - L_u) = T \ln(|\Sigma_r|/|S|), \quad \text{where}
\]

\[
(5b) \quad |\Sigma_r| = \prod_{i=1}^{N-1} (1 + \hat{v}_N^2).
\]
The unrestricted maximum involves \( N(N-1)/2 \) parameters, while the restricted maximum involves \( N \) parameters, so that \( \lambda \) is asymptotically distributed as \( \chi^2 \) with \( N(N-3)/2 \) degrees of freedom.

In order to apply this localization test, one must complete the specification of (2) by making assumptions about the \( X_i \) and the \( \phi_i \). Since those assumptions are inevitably somewhat arbitrary, we encounter a classic problem, variants of which arise in most applied econometric work: rejection of the null hypothesis of symmetry may be signalling that the wrong symmetric model has been employed, not that the true model involves localization. The problem of specification uncertainty does not have a definitive solution, here in general, but several points seem worth noting in this context. 9

First, in selecting a symmetric specification for testing, one should be unusually reluctant to impose restrictions that lack strong \textit{a priori} support. Over-parameterization robs the localization test of some power, but misspecification is likely to bias it toward rejecting symmetry. Second, if localization \textit{is} present, \textit{all} symmetric models are misspecified. Failure to find a symmetric model that passes all (formal and informal) specification tests not involving localization may mean that the investigator lacks adequate data or ingenuity, or it may simply mean that localization is present. Third, if the residuals-based localization test described above rejects symmetry because localization is present, not because the wrong symmetric specification has been employed, the pattern of residuals ought to imply plausible patterns of localization. If such patterns are detected by the diagnostic technique described below, one's confidence in the presence of localization should be increased. Finally, the localization test provides a new specification test that should be applied to symmetric market share models before they are used for forecasting or other purposes. Rejection of symmetry always signals misspecification of one sort or another.
3. Assessing Patterns of Localization

If the localization test rejects the null hypothesis of symmetry in some market, it still seems reasonable to retain, at least provisionally, the assumption that market shares depend only on the \( \Phi_i(X_i,u_i) \) even though (1) does not hold. In such a generalized attraction model, the pattern of competitive effects is the same for all brand-specific marketing variables, like price and advertising. An asymmetric generalized attraction model seems a plausible representation of an "address model" with localization, in which the effects of changes in the \( \Phi_i \) 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 quantities can be used to analyze the nature and importance of localization:

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

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 the \( M_j \)-weighted averages of the \( K_{ij} \) to equal unity for all \( i \). 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. Moreover, Schmalensee (1982) presents a measure, \( G^* \), of the market-wide importance of localization that involves only the \( K_{ij} \) and the \( M_i \).

Under the generalized attraction assumption, the same pattern of localization applies to both observable and unobservable variables. We can
thus use prediction errors to describe localization as well to test for its presence. Given the difficulties of estimating more general models in small samples, we consider the use of statistics derived from estimation of equations (2). Since these equations are misspecified if localization is present, such an approach can yield diagnostic information, but not rigorously defensible estimates. (All this is in the spirit of the diagnostic techniques presented by Philips (1974, pp. 212-17)).

Let $\Omega$, with typical element $\omega_{ij}$, be the singular covariance matrix of the $\varepsilon_{it} = M_{it} - \hat{M}_{it}$, the differences between actual shares and those predicted by the symmetric model. I want to argue that for diagnostic purposes it is reasonable to substitute $(\omega_{ij}/\omega_{ii})$ for the ratio of derivatives in (6).

The intuitive rationale for this approach is that one would expect a larger than average negative covariance between the shares of brands that are closer than average substitutes, as buyers are more likely than average to substitute one for the other as market conditions vary. More formally, we estimate the ratio of structural derivatives in (6) by an expected ratio of changes. Suppose that i's share increases by an amount $\Delta$ due to factors that are either unobservable or excluded from (2) by virtue of the assumption of symmetry. The extent to which this is expected to be at the expense of brand j's share is given by $E[(\varepsilon_{jt} - \bar{\varepsilon}_j)/(\varepsilon_{it} - \bar{\varepsilon}_i) \mid (\varepsilon_{it} - \bar{\varepsilon}_i) = \Delta]$. Under normality, treating the mean prediction errors (the $\bar{\varepsilon}_i$) as known, this expression equals $(\omega_{ij}/\omega_{ii})$ for any $\Delta$.

Note also that since actual and predicted market shares always sum to one, the $\varepsilon_{it}$ always sum to zero across i, as do the $\bar{\varepsilon}_i$. This implies that the row and column sums of $\Omega$ are zero, and this in turn implies that the $M_j$-weighted averages of the estimated $K_{ij}$ equal one for all i.
4. An Application to Breakfast Cereals

In the course of antitrust proceeding involving the leading U.S. producers of ready-to-eat breakfast cereals, the U.S. Federal Trade Commission (FTC) obtained a good deal of monthly brand-specific data from the major producers. For the 40-month period from September, 1969 through December, 1972, the FTC 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. These six firms, which accounted for over 97% of U.S. 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 two percent 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 cereal revenues during this period.

Under the null hypothesis of perfect symmetry, if equations (1) and (2) hold in some market, one can estimate (2) and perform a valid test for localization using any subset of the brands on the market. On the other hand, if localization is present, the power of the test depends on which subset of brands is used. In a market with clusters of closer-than-average substitutes, power would be low if all brands analyzed were in the same cluster or all were in different clusters. On that reasoning, our test should have adequate power.
here; brands 7 and 9 are raisin brans, brands 3, 8, and 11 are pre-sweetened cereals consumed heavily by children, and brands 6 and 10 are aimed at buyers especially concerned with nutrition.

Following most of the relevant literature, we assume that \( \phi_i \) are Cobb-Douglas functions, so that equations (2) are linear:

\[
\ln[\phi_i(t)] = \alpha_i + \beta_i \ln[A_i(t)] + \gamma_i \ln[P_i(t)] + \delta_i \ln[M_i(t-1)], \quad i=1,\ldots,N,
\]

where \( A \) is advertising spending, \( P \) is 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.) Using three different base brands (as discussed below), F-tests decisively rejected the commonly-imposed restriction that the \( \beta \)'s, \( \gamma \)'s and \( \delta \)'s were the same for all brands.

Reported wholesale list prices were used for the \( P_i(t) \). 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 possible source of specification error, as such "deals" cannot be treated as exogenous, random events.

The reported advertising series exhibited large month-to-month fluctuations (10 of 11 coefficients of variation above 0.30) and contained two negative observations. As Aaker, Carman and Jacobsen (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 an attempt to 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 lagged variables in our specifications, reduced the number of observations to 38.

Given firms' information and decision-making lags, it seems plausible to treat both price and advertising as predetermined in monthly data of this sort. Accordingly, estimates of (2), using (7), were computed using iterated generalized least squares ("seemingly unrelated regressions") for three different choices of the base brand, brand \( N \) in equations (2): the largest (Kellogg's Corn Flakes), the smallest (Kellogg's Froot Loops), and a medium-sized non-Kellogg brand (General Mills' Wheaties). Coefficient estimates were not very sensitive to the choice of the base brand.

The model's explanatory power varied considerably across brands, though not across the three sets of estimates. R-square statistics for the differences between the logarithms of shares varied from 0.89 (Special K -- Froot Loops) to -0.07 (Rice Krispies -- Wheaties). Only two of the 30 \( R^2 \)'s computed were negative, while 20 exceeded 0.10.

As might be expected, given the large number of parameters estimated with short data series, slope coefficients were not generally both significant and of the expected sign. Eight of the 11 estimated \( \beta \)'s were negative with all three bases, and either five or six of the estimated \( \gamma \)'s were positive. Only one positive \( \beta \) exceeded twice its standard error, along with four or five negative estimates. No negative \( \gamma \)'s were significant; either one or three positive estimates had t-statistics above two. Seven of the 11 \( \delta \)'s were positive with all three bases, none ever exceeded unity, and between three and five positive \( \delta \)'s had t's above two. Overall, these were hardly superb results, but, in view of our sample size and model complexity, they did not seem to provide a decisive signal of misspecification.
Application of the localization test derived in Section 2 to the three residual covariance matrices does provide such a signal. The $\hat{v}_i$ are generally insensitive to choice of base brand, and the estimates for Kellogg brands are generally much smaller than for the others. The three values of $\lambda$ vary between 151.4 and 155.8. As $\lambda$ is distributed as $\chi^2(44)$ under the null hypothesis of perfect symmetry, with 0.5% critical value equal to 72, these are all highly significant. The evidence so far seems to indicate strongly that the assumption of no localization made in (2) is invalid in the U.S. ready-to-eat breakfast cereals industry.

Application of the diagnostic technique described in Section 3 to this model's prediction errors yields a surprisingly large number of negative estimated $K_{ij}$ with all three base brands. Negative estimates might arise 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 brands) in our sample 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 $j$.

Such exogenous, group-specific disturbances do not appear to be at work here, however. For all three 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, for which symmetry is always decisively rejected, are used to generate predicted shares. It is hard to see how this sign pattern could be reflecting localization, since both
Kellogg and non-Kellogg sets contain brands in most of the "segments" (raisin brans, pre-sweetened, nutritional, etc.) discussed by industry observers. It seems much more likely that we have either failed to observe important marketing variables (such as the departures from list price noted above) 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.  

Since the null hypothesis of symmetry is essentially the hypothesis that all $K_{ij}$ equal one, it is intuitively clear that that null hypothesis was rejected here mainly because of the negative estimated $K_{ij}$ (corresponding to positive residual covariances) within the Kellogg and non-Kellogg sets. If, as we have just argued, the estimated $K_{ij}$ likely reflect the absence of important variables or the presence of serious measurement problems, it follows that the size of our localization test statistic, $\lambda$, likely reflects specification problems unrelated to the presence of localization. In short, the techniques developed in Sections 2 and 3 have detected misspecification that would otherwise not have been apparent, but they have not done so in a fashion that makes it possible to assert that localization is definitely involved.

5. Conclusions

Section 2 presented an econometric test for the presence of localized competition that does not require unusually large data sets. Section 3 presented a complementary technique for assessing the nature and importance of such localization. Application of these methods to data on the U.S. ready-to-eat breakfast cereal industry in Section 4 produced good news and bad news.
The good news is that our techniques revealed their ability to signal clearly misspecification of commonly-employed symmetric market share models that would otherwise not likely be detected. The bad news is that in this particular application it was impossible to assert that the symmetric model examined was misspecified because of the presence of localization, since the technique of Section 3 signalled the presence of unrelated difficulties.
References


References (cont.)


Footnotes

* I am indebted to the National Science Foundation for financial support, to William Burnett for considerable help with data problems, to Jerry Hausman for many helpful conversations, and to Severin Borenstein, Ian Ayres, and Michael Salinger for superb research assistance. Earlier versions of this essay were presented at the Universities of Chicago and Laval and at the 1983 North American Winter Meeting of the Econometric Society. I am grateful to participants in those sessions and to Timothy Bresnahan, Grayham Mizon, Thomas Stoker, and anonymous referees for valuable comments. Only the author can be blamed for residual infirmities.

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

2. Alternative approaches to constructing such maps are discussed by Hauser and Koppelman (1979), McAlister and Lattin (1983), and the references they cite.

3. While this requirement has a strong formal resemblance to the choice-theoretic axiom of "independence of irrelevant alternatives," individual behavior here need not satisfy that axiom.

4. This is not strictly necessary for the absence of localization in competitive interaction. It would suffice to have the definition in the preceding paragraph apply only to the demand-determining variables under sellers' control. Under this weaker definition, market-wide variables like disposable income could enter the \( \Phi_i \) in (1), below. It is very difficult to imagine conditions that would produce perfect symmetry under this weaker definition but not under the definition in the text, however, and that more convenient definition is accordingly used in what follows.
Footnotes (cont.)

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

6. 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 equation: see Bultez and Naert (1975) and Naert and Weverbergh (1981). (Maximum likelihood estimates are invariant to the equation deleted.) While that approach appears more "balanced" than the one employed here, neither seems theoretically superior. Since the choice of a base in our approach is essentially arbitrary, marked variation in parameter estimates with the base brand employed signals problems with the demand specification. (See footnote 13, below).

7. Berndt and Savin (1975) derive the restrictions on autoregressive disturbance processes implied by the adding-up constraint.

8. 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 derived above, using a consistent estimator of \( \Sigma \) that is not maximum-likelihood. Use of such constrained estimates in a test procedure would complicate computation and yield no asymptotic benefits.

9. For general discussion of specification uncertainty and its implications, see Leamer (1978) and Feldstein (1982).

10. The extent of localization was an issue in that case; see Schmalensee (1978) and Berman (1981). The FTC data have also been employed by Aaker, Carman, and Jacobsen (1982); they provide additional descriptive information.

11. 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 and the case itself was finally dropped by the FTC. The other two national sellers were Nabisco and Ralston.
12. We were unable to obtain convergence for any (nonlinear) generalizations of this model involving quasi-differencing to eliminate serial correlation. A number of observations indicate that our inability to allow explicitly for serial correlation is not a serious problem here, however. First, despite the lagged dependent variable in (17), estimation of (2) does not involve the standard lagged dependent variable problems because because of the cross-equation constraints on $\delta_N$. Second, serial correlation was not generally estimated to be important in models that imposed equal slope coefficients across brands. Third, 26 of the 30 single-equation Durbin-Watson statistics were between 1.5 and 2.5, and 20 were between 1.7 and 2.3. This is reassuring, even though the DW statistic cannot be used for formal testing here. Fourth, one expects estimates of the $\delta$'s to be biased away from zero by serial correlation, but 10 of 11 estimates were between -0.24 and +0.58 for all three base brands. (The outlier, $\delta_{11}$, was always between 0.82 and 0.85.) Finally, our main interest here is in residual convariances, which need not be strongly biased by the presence of serial correlation.

13. As was observed in footnote 7, above, coefficient estimates should not vary markedly with the choice of base brand if the specification is correct. Comparing pairs of base brands, the three correlation coefficients for the 11 estimated $\beta$'s were between .96 and .99, for $\gamma$'s between .77 and .90, for $\delta$'s between .79 and .98, and for $\alpha$'s (setting $\alpha_N=0$) between .90 and .95. The weakest relation was always between estimates using Froot Loops and Wheaties bases. We also considered variation relative to estimation precision. Let $x_i$, $i=1$, 2, or 3, be the estimate of some coefficient obtained using the $i^{th}$ base brand, and let $h_i$ be the reciprocal of the corresponding squared standard error. Then if the specification is correct, one would expect the following statistic to be small relative to the
Footnotes (cont.)

13. (cont.) chi-square distribution with two degrees of freedom:

\[ C^2 = \sum h_i [x_i - (\sum h_i x_i / \sum h_i)]^2. \]

Only one of the 11 values of \( C^2 \) exceeded one for \( \beta, \gamma \) or \( \delta \). Only \( C^2 \) for \( \delta_7 \) exceeded two. (This statistic cannot be computed for the \( \alpha \)'s.) In contrast, the imposition of slope equality across all brands produced \( C^2 \)'s of 9.98 for \( \beta \) and 42.30 for \( \delta \).

14. Only \( R^2 \)'s involving pairs of base brands are directly comparable among the three sets of estimates, but these are quite stable. The \( R^2 \) for the differences between the logarithms of brand 1's and brand 11's shares was .755 when brand 1 was used as a base and .743 when brand 11 was the base. The corresponding two \( R^2 \)'s for brands 1 and 5 were .058 and .041, and the values for brands 5 and 11 were .269 and .305.

15. Pairwise correlations among the three sets of \( v_i \)'s range from .27 to .69. If the two base brand estimates (\( v_N \)'s) are excluded in each case, the range is from .64 to .85. The average \( \bar{v} \) over all three bases is .0138 for the six Kellogg brands and .0568 for the five non-Kellogg brands. This gap is widened if base brand estimates are excluded.

16. 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 Jacobsen (1982) for a set of Kellogg brands.

17. See the discussion of misspecification at the end of Section 2. It should be kept in mind that our data were compiled from the (presumably) uncoordinated responses of the leading cereal producers to a broadly-worded request for data from the FTC, which was in the process of bringing a major antitrust case against them. In retrospect, it is hard to imagine a procedure more likely to make it difficult or impossible to avoid specification error in market-wide models.
Table 1

Leading Ready-to-Eat Cereal Brands Used in Empirical Analysis<sup>a</sup>

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

<sup>a</sup> Dollar sales averages cover the years 1969-1972; pound sales averages cover the period October, 1969 through November, 1972.