FACULTY WORKING PAPERS

College of Commerce and Business Administration

University of Illinois at Urbana-Champaign

August 28, 1973

CANONICAL CORRELATION ANALYSIS OF COMPETITIVE MARKET STRUCTURE

Johny K. Johansson and Jagdish N. Sheth

#128

CANONICAL CORRELATION ANALYSIS OF COMPETITIVE MARKET STRUCTURE

Johny K. Johansson, University of Illinois
Jagdish N. Sheth, University of Illinois

This paper reports the results of an empirical study of the margarine industry and its competitive structure. The emphasis in the paper is on the feasibility of utilizing canonical correlation analysis to identify dimensions of price competition. The empirical results indicate that major brands of margarine have only a limited degree of price competition.

INTRODUCTION

Identification of competitive market structure in monopolistic competition and in heterogeneous oligopoly situations is extremely difficult. The economic price theory is of limited usefulness in these market situations due to product differentiation and segmentation of the market. However, the manager and the public policy makers must understand the nature and structure of price competition in order to effectively compete and regulate it. To understand the competitive structure of an industry, we need to know the following: How are the different brands in the product class related to one another? Are they complimentary because they are demanded reciprocally or not at all related because there is no correlation between them? Second, who competes with whom? Does the price change in one brand have any effect on the demand of other brands? What is the extent of price elasticity and cross-elasticity across different brands? Finally, is it possible to identify any clusters of brands which are internally competing among themselves but there seems virtually no price competition between the clusters? In other words, how extensive is price competition in the market place?

The traditional econometric models tend to be less useful to provide insights into the above problems because they either require a priori presumptions about the market structure or insist on a highly formal mathematical models properly specified and identified even though it may be unrealistic. In short, the econometric models rely heavily on the theory of the firm to specify models of competition and the price theory has not addressed itself fully in situations other than perfect competition, oligopoly and monopoly all of which require undifferentiated products and markets.

A multivariate statistical model which seems appropriate to provide the empirical determination of dimensionality of market structure and the magnitude of price competition among various types of brands is canonical correlation analysis. The technique maximizes the structure of relationships between a set of criterion variables such as the demand for various brands and a set of predictor variables such as the average prices of those brands. The objective in canonical correlation analysis is to reduce the total complexity of relationships both within and between sets of variables to a common orthogonal multidimensional space [2] [5]. The development of the common underlying dimensionality which exists between and within several criterion and predictor variables enables the researcher to obtain a total perspective of the phenomenon which is otherwise not possible in the econometric techniques such as regression or even simultaneous equations.

Since it is often our objective to determine the dimensionality of competitive structure of a market it is logical to utilize canonical correlation analysis. Unfortunately, its application to econometric problems has been limited [3] due to the complexity of the technique itself [4] and due to problems in its interpretations [1]. However we hope to suggest some guidelines in the application of canonical correlation analysis in this paper which minimize these problems.

STUDY DESIGN AND RESULTS

The study was conducted to empirically estimate the structure of margarine industry with respect to price competition. The data for the demand of various brands of margarine are a part of a national probability panel of more than 10,000 households who report every week their purchase behavior with respect to grocery products. From this data base, we have calculated the average weekly price in cents for each of the ten largest selling brands of margarines. Similarly, we are in a position to estimate the average quantity demanded in ounces of each brand of margarine for each week. We have, therefore, 52 observations on the average price and quantity of each brand of margarine in one year history of continuous reporting by the panel members.

In order to examine the structure of price-quantity relationships among the ten brands, simple correlations are calculated and summarized in Table 1. The diagonal elements in the table represent the price elasticities of the brands of margarine. As can be seen from the table, the demand for brands 4 and 5 seems to be relatively more determined by their own prices than by others while the demand for brands 6, 3 and 7 seems to be determined by nonprice factors. Also, the average price is not the sole determinant of the average demand for any brand of margarine suggesting that other factors such as consumer preferences, habits, brand loyalties and product perceptions may be also partly responsible for a brand's demand.

The off-diagonal elements of the table represent the cross-elasticities of different brands of margarine. Surprisingly, most of the cross-elasticity correlations are not significant indicating relatively little interaction between the competing brands. Only in the case of brand 7, do we find that its cross-elasticity correlation with the price of brand 1 is stronger than its own price elasticity correlation. Second, brands whose demand is governed by their price are not equally sensitive to competitor's prices. For example, brand 4 has the highest price elasticity correlation but very small cross-

TABLE 1

Simple and Multiple Correlations Between Brand Demand and Price

		Brand Price										
		1	2	3	4	5	6	7	8	9	10	Multiple Correlation
Brand Quantity	1	-.54*	.13	.07	-.03	.30*	.03	.13	.06	.07	.09	.68*
	2	.31*	-.54*	.17	-.14	-.26	.29*	.19	-.03	-.28*	-.09	.67*
	3	-.24	.09	-.29*	.01	-.07	-.04	.18	.26	.07	.06	.52
	4	-.10	-.12	.15	-.70*	-.29*	.02	.02	-.23	-.09	-.14	.76*
	5	.31*	-.40*	.01	-.33*	-.63*	.13	.17	-.14	-.01	.01	.77*
	6	-.12	-.15	.12	-.06	-.01	-.11	.12	-.06	-.11	-.25	.36
	7	.44*	-.05	.15	.01	-.07	.09	-.31*	-.07	-.05	.00	.54
	8	.06	-.04	.10	-.26	-.22	.28*	-.02	-.51*	-.18	-.04	.61*
	9	-.03	-.11	.08	-.10	-.33*	.09	.32*	.08	-.51*	-.12	.62*
	10	-.04	.15	.32*	.36*	.12	-.04	.06	.07	.13	-.41*	.64*

*Significant at 0.05 level

elasticity correlations with other brands. Third, two brands seem to be able to bring about changes in the demand of other brands of margarine by changing their prices. These are brands 1 and 6 whose cross-elasticity correlations across other brands are larger than others. On the other hand, price changes in brands 8, 9 and 10 seem to have very little relationship with the demand of competing brands although their own demand is significantly affected by these price changes. We hypothesize this situation to be probable for regional or private label brands. Finally, there are several significant cross-elasticity correlations which are negative and, therefore, contrary to price theory. Fortunately, the magnitude of these correlations is not high enough to seriously examine for alternative explanations. The simple correlations do not fully portray the magnitude of price-quantity relationships for each of the brands. Therefore, a linear multiple regression of the demand for a brand on the prices of all the ten brands was performed on each brand of margarine. The results are also summarized in Table 1. Even though the multiple correlation is generally greater than the price-elasticity correlation of the brand, most of the inflated correlation arises by chance since the degrees of freedom are rapidly lost with small number of observations. A statistical test of significance clearly reveals that the total price-quantity relationship is not even significant in the case of brands 6, 3 and 7 whose demand is presumably determined by nonprice factors. On the other hand, the demand of brands 4 and 5 is fairly well determined by the interplay of their prices and competing brands' prices in the market place.

While we are able to examine the impact of price competition on each brand of margarine separately, it is not possible to determine the structure of price competition from the viewpoint of the total industry with the use of regression analysis. However, it is possible to examine the price-quantity relationship for the total industry with the use of canonical correlation analysis. We develop the canonical correlation model by treating the average demand of each of the ten brands as criterion variables and their average price as the predictor variables. The simultaneous regression of all the demand variables on all the prices represents the analysis of estimating competitive market structure for the total industry. Table 2 summarizes the results of canonical correlation analysis. The first five canonical variates were found to be significant at 0.05 level and are, therefore, retained for further analysis and inference.

As in other multivariate techniques, the canonical variates are estimated by arbitrarily maximizing the correlation between the criterion and the predictor sets of variables and deriving each additional canonical variate conditional upon the maximization by earlier variates. This computational convenience, however, presents serious problems of interpretation because of the choice of one set of canonical variates from an infinite combination of other sets all of which can maximize the total correlation between two sets of variables. In order to resolve the indeterminacy, it is advantageous to rotate the canonical axes similar to the rotation procedures commonly utilized in factor analysis. We have, therefore, performed an orthogonal varimax rotation of the significant canonical variates in order to assess the underlying market structure among the ten brands with respect to their price-quantity relationships.

The rotated canonical variates are summarized in Table 3. Each rotated canonical axis represents an independent (orthogonal) dimension or aspect of the total competitive market structure. Before we examine the substantive inferences of each dimension of the canonical space, it is worth nothing that the average variance explained in demand of all the ten brands of margarine by all the average prices is 39 per cent. Similarly, the redundancy measure proposed by Steward and Love [6] is only 33 per cent if we limit the calculations to the first five significant canonical variates. The difference of six per cent between the average R^2 and the redundancy measure is, therefore, due to the additional canonical variates and can be treated as error or unstable variance. However, it is obvious from the results that the economic price theory is only partially verified when we carry the analysis to the total industry level because the redundancy measure or the average R^2 represents the optimal amount of relationship between price and demand of the margarine market. It would, therefore, appear that the bulk of demand for specific brands of margarine is determined by nonprice factors including habit, loyalty, product perceptions and situational events impinging on micro level purchase activities.

If we examine the structure of each dimension of price competition, we can easily see that some brands are substitutes, others are complimentary and many others are isolated. The first dimension clearly reflects the substitute relationship between brand 1 and 8 because they exhibit both price elasticity and cross-elasticity relationship mutually for each other. The second dimension reflects the isolated demand for brand 10 as a function of its own price and the price of brand 3. We see the surprising empirical evidence where the cross-elasticity between two brands is not reciprocal: the price change in brand 3 has an effect on the demand of brand 10 but not vice versa. The third dimension brings out in bold relief the relatively dominant price-elasticity of brand 4. It is an isolated brand whose demand is primarily sensitive to changes in its own price.

The fourth and the fifth dimension reflect the complimentary relationship between brands. It seems that brands 5 and 9 are jointly demanded as a function of the price of brand 5. Strangely, the price changes in brand 7 also seem to determine somewhat this complimentary relationship. Finally, brands 2 and 7 also manifest complimentary relationship as a function of the prices of brands 1, 2 and 9. Once again, we see the lack of reciprocal relationship between the brands in that price effects brands 1 and 9 on brands 2 and 7 are not reciprocated.

It will be noted that none of the three brands whose individual price-quantity relationships are nonsignificant (brand 6, 3 and 7) exhibit any strong directionality on the five dimensional canonical space of the competitive market structure. This is quite congruent in view of the fact that the demand for these brands is more determined by nonprice factors than by price factors.

TABLE 2

Canonical Analysis of Price-Quantity Variables

FUNCTION	EIGENVALUE	CORRELATION	WILKS LAMBDA	CHI-SQUARE	DF
1	0.7900	0.8888	0.0049	220.4383	100
2	0.6524	0.8077	0.0235	155.6742	81
3	0.6120	0.7823	0.0676	111.8266	64
4	0.4084	0.6391	0.1741	72.5402	49
5	0.3407	0.5837	0.2944	53.7528	36

TABLE 2 (continued)

Canonical Analysis of Price-Quantity Variables

FUNCTION	EIGENVALUE	CORRELATION	WILKS LAMBDA	CHI-SQUARE	DF
6	0.3292	0.5737	0.4465	33.4640	25
7	0.2521	0.5021	0.6656	16.8953	16
8	0.0691	0.2628	0.8899	4.8418	9
9	0.0405	0.2012	0.9559	1.8724	4
10	0.0038	0.0614	09962	0.1566	1

Matrix* of Criterion Weights

	1	2	3	4	5
Quantity, Brand 1	-0.39759	0.17341	-0.03855	0.47071	0.37221
Quantity, Brand 2	0.20604	-0.32625	0.44315	0.59724	0.43645
Quantity, Brand 3	0.08795	0.20906	-0.42405	0.10744	-0.06724
Quantity, Brand 4	-0.25863	0.77067	0.46134	-0.01939	0.16902
Quantity, Brand 5	0.49386	0.47828	-0.19002	-0.12566	-0.27489
Quantity, Brand 6	-0.20884	0.08510	0.12558	0.39249	-0.11814
Quantity, Brand 7	0.17414	-0.40619	0.36654	0.09703	0.29456
Quantity, Brand 8	0.08239	-0.02196	0.18659	-0.51717	-0.18296
Quantity, Brand 9	0.27960	0.00778	-0.14166	0.50750	-0.44964
Quantity, Brand 10	-0.30981	0.06025	0.43321	0.03530	-0.59927

Matrix* of Predictor Weights

	1	2	3	4	5
Price, Brand 1	0.61124	-0.46582	0.32033	-0.24499	0.07676
Price, Brand 2	-0.17010	0.07430	-0.11964	-0.38689	-0.35252
Price, Brand 3	-0.37325	0.11772	0.56693	0.30895	0.01395
Price, Brand 4	-0.07633	-0.64315	-0.33481	0.02428	-0.68219
Price, Brand 5	-0.47317	-0.40639	0.25968	0.17731	0.71618
Price, Brand 6	0.12580	0.06679	-0.02275	0.16921	0.14934
Price, Brand 7	0.17004	0.23673	-0.28224	0.34104	-0.31225
Price, Brand 8	0.14327	-0.06134	-0.33832	0.51152	0.31017
Price, Brand 9	-0.14592	0.38864	-0.25229	-0.35827	0.02776
Price, Brand 10	0.18338	0.03435	-0.42572	-0.06690	0.42425

*Only significant canonical variates are included using 0.05 significance level.

TABLE 3

Rotated Canonical Space of Competitive Market Structure*

	Rotated Canonical Variate					R^2
Criterion Set (Quantity)	I	II	III	IV	V	
Brand 1	.68				.89	.46
Brand 2						.44

TABLE 2 (continued)

Canonical Analysis of Price-Quantity Variables

FUNCTION	EIGENVALUE	CORRELATION	WILKS LAMBDA	CHI-SQUARE	DF
6	0.3292	0.5737	0.4465	33.4640	25
7	0.2521	0.5021	0.6656	16.8953	16
8	0.0691	0.2628	0.8899	4.8418	9
9	0.0405	0.2012	0.9559	1.8724	4
10	0.0038	0.0614	09962	0.1566	1

Matrix* of Criterion Weights

	1	2	3	4	5
Quantity, Brand 1	-0.39759	0.17341	-0.03855	0.47071	0.37221
Quantity, Brand 2	0.20604	-0.32625	0.44315	0.59724	0.43645
Quantity, Brand 3	0.08795	0.20906	-0.42405	0.10744	-0.06724
Quantity, Brand 4	-0.35863	0.77067	0.46134	-0.01939	0.16902
Quantity, Brand 5	0.49386	0.47828	-0.19002	-0.12566	-0.27489
Quantity, Brand 6	-0.20884	0.08510	0.12558	0.39249	-0.11814
Quantity, Brand 7	0.17414	-0.40619	0.36654	0.09703	0.29456
Quantity, Brand 8	0.08239	-0.02196	0.18659	-0.51717	-0.18296
Quantity, Brand 9	0.27960	0.00778	-0.14166	0.50750	-0.44964
Quantity, Brand 10	-0.30981	0.06025	0.43321	0.03530	-0.59927

Matrix* of Predictor Weights

	1	2	3	4	5
Price, Brand 1	0.61124	-0.46582	0.32033	-0.24499	0.07676
Price, Brand 2	-0.17010	0.07430	-0.11964	-0.38689	-0.35252
Price, Brand 3	-0.37325	0.11772	0.56693	0.30895	0.01395
Price, Brand 4	-0.07633	-0.64315	-0.33481	0.02428	-0.68219
Price, Brand 5	-0.47317	-0.40639	0.25968	0.17731	0.71618
Price, Brand 6	0.12580	0.06679	-0.02275	0.16921	0.14934
Price, Brand 7	0.17004	0.23673	-0.28224	0.34104	-0.31225
Price, Brand 8	0.14327	-0.06134	-0.33832	0.51152	0.31017
Price, Brand 9	-0.14592	0.38864	-0.35229	-0.35827	0.02776
Price, Brand 10	0.18338	0.03435	-0.42572	-0.06690	0.42425

*Only significant canonical variates are included using 0.05 significance level.

TABLE 3

Rotated Canonical Space of Competitive Market Structure*

	Rotated Canonical Variate					R^2
Criterion Set (Quantity)	I	II	III	IV	V	
Brand 1	.68				.89	.46
Brand 2						.44

REFERENCES

1. Alpert, M. I. and Peterson, R. A., "On the Interpretation of Canonical Analysis," Journal of Marketing Research, Vol. 9 (May, 1972), pp. 187-92.

2. Anderson, T. W., Introduction to Multivariate Statistical Analysis (New York, John Wiley & Sons, 1958).

3. Carmone, Frank J., "The Use of Canonical Correlation Analysis in Marketing," unpublished paper presented at AMA Workshop on Multivariate Methods, Chicago, January, 1970.

4. Hannan, E. J., "Canonical Correlation and Multiple Equation Systems, in Economics Econometrika, Vol. 35 (January, 1967) pp. 123-142.

5. Hotelling, Harold, "Relations Between Two sets of Variates," Biometrika, Vol. 28 (December, 1936), pp. 321-77.

REFERENCES

1. Alpert, M. I. and Peterson, R. A., "On the Interpretation of Canonical Analysis," Journal of Marketing Research, Vol. 9 (May, 1972), pp. 187-92.

2. Anderson, T. W., Introduction to Multivariate Statistical Analysis (New York, John Wiley & Sons, 1958).

3. Carmone, Frank J., "The Use of Canonical Correlation Analysis in Marketing," unpublished paper presented at AMA Workshop on Multivariate Methods, Chicago, January, 1970.

4. Hannan, E. J., "Canonical Correlation and Multiple Equation Systems, in Economics Econometrika, Vol. 35 (January, 1967) pp. 123-142.

5. Hotelling, Harold, "Relations Between Two sets of Variates," Biometrika, Vol. 28 (December, 1936), pp. 321-77.

6. Stewart, Douglas and Love, William, "A General Canonical Correlation Index," Psychological Bulletin, Vol. 70 (September, 1968), pp. 160-63.