

# Identifying the Determinants of Advertising Intensity in Iran's Industries Using the Multiple Discriminant Analysis Technique

Saeed Tajdini (Corresponding author)

Faculty of Management, University of Tehran, Tehran, Iran

Tel: 98-912-469-0755 E-mail: s.tajdini@ut.ac.ir

Mohsen Nazari

Faculty of Management, University of Tehran, Tehran, Iran

E-mail: Mohsen.Nazari@ut.ac.ir

Tahmoures Hasangholipour

Faculty of Management, University of Tehran, Tehran, Iran

E-mail: thyasory@ut.ac.ir

Received: May 10, 2012

Accepted: May 28, 2012

Published: July 1, 2012

doi:10.5296/jmr.v4i3.1789

URL: <http://dx.doi.org/10.5296/jmr.v4i3.1789>

## **Abstract**

This research is aimed at identifying the main variables in an industry based on which the advertising intensity in that industry can be estimated. Using prior knowledge, six variables, including investment, statism ratio, concentration, sales, added value and direct export were utilized by Multiple Discriminant Analysis, using the stepwise method. It turned out that among these variables, sales, direct export and concentration are variables which can better enhance the discrimination power of the model. Based on the means and standardized canonical coefficients of these three variables in different groups and also the group centroids in each function, we concluded that the first function is better at discriminating the first group from the other two and the second function is better at discriminating the third group from the other two. The correct classification percentage of 82% indicates the highly satisfactory performance of the model in classifying the data.

**Keywords:** Firms, Market Performance, Market Structure, Advertising Intensity, Multiple Discriminant Analysis

JEL Classification: L100

## 1. Introduction

In general, advertising is a kind of communication whose purpose is to encourage the audience to take a specific course of action or to go on doing what's being done at the moment. According to Belch (2001), advertising is any paid form of non-personal communication about an organization, product, service, or idea by an identified sponsor.

Advertising is, by no means, a modern phenomenon. The ancient Greek used papyrus as their advertising medium. Found-and-lost notices were common in the ancient Greece and Rome. Wall and rock postings are other types of ancient advertising media that can be found even today in Asia, Africa and Latin America. Their root can be traced back to 4000 years BC in the art of rock painting in India (Bhatia, 2000). Later, as human life developed and revolutionizing inventions such as the press machineries emerged, the integrated advertising industry formed and became an inseparable part of the business world. In the last decades, the spread of internet has also affected advertising and nowadays huge sums of money are spent on internet advertising.

In the current volatile and hypercompetitive business world, decision making about the firm's advertising can be a make-or-break issue for the firm survival or growth. In the USA alone, in 2001, 230 billion dollars were invested on advertising, a number doubled from 1980 (McCann-Erickson, 2001). Acknowledging the vital role of advertising, the remaining problem is that what discriminating features of an industry determine the ratio of its advertising expenditure to its sales.

Though much research has been conducted to examine the relationship between advertising intensity and some variables defining market structure, as far as the author is concerned, little scholarly effort has been put to identify the market structure variables which can be used to predict the behavior of firms and industries when it comes to budget their advertising. As far as the author is informed, no previous research has been done for the same goal and with the same statistical method. Furthermore, what still adds to the perceived need to conduct this research is the variety of independent variables used. Though variables like concentration and investment have been extensively used by previous researchers, variables like added value and statism ratio are novel.

## 2. Review of Literature

The SCP paradigm, developed in 1930s and 1950s by Edward Mason and his student, Joe Bain, considers a directional relationship between structure (S), conduct (C) and performance (P). This framework, which was originally used by the US government to combat corporate trust, later entered many other managerial and economic disciplines such as strategy and strategic management. Strategists like Michael Porter of Harvard University have used this framework to analyze the competitive activities of firms, for example. This paradigm, basically, states that performance of an industry (its success in creating value for its customers) depends on the conduct of the firms in that industry, which is, itself, dependent on the structure of the market (Scherer and Ross, 1990).

Following this mentality, many researchers have tried to test this hypothesis by examining how this directional relationship works. That is whether or not the market structure really influences conduct and, subsequently, the performance of an industry. What varies in the previous researches is mainly the choice of variables to represent these three elements: structure, conduct and performance.

Advertising intensity has been widely used as the dependent variable for the independent variables representing the market structure. No less that commonly researchers have investigated to see how structural variables like concentration, price-cost margins and product features affect the intensity with which firms advertise in an industry.

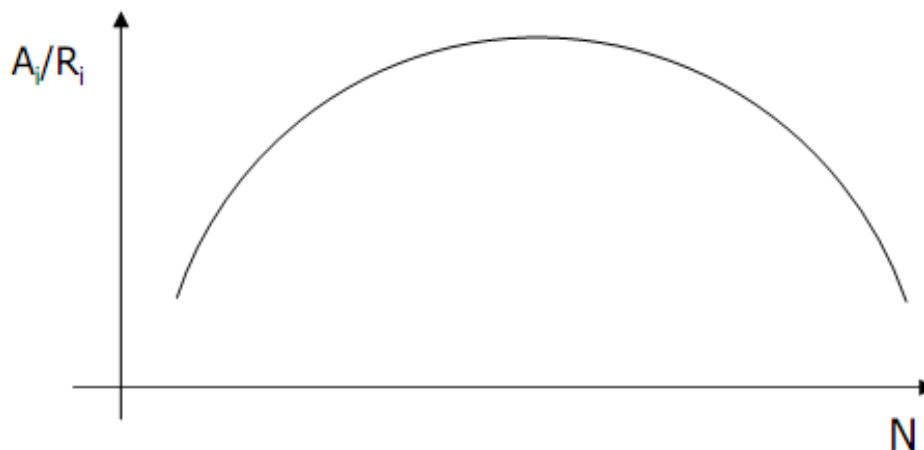
Market structure has been investigated using many variables. Each researcher has chosen a selection of variables in his/her research to represent market structure. Among these variables some are more common and used more frequently. Concentration, for example, seems to be a major market structure variable in the majority of researchers' minds. That what's the real relationship between this market variable and advertising intensity still remain a much-debated mystery in this research area. Many researchers claimed that there is an inverted U-shape relation between concentration and advertising intensity (Nazari and Tajdini, 2011; Lee, 2002<sup>1</sup>; Buxton and Davies, 1984<sup>2</sup>; Strickland and Weiss, 1976; Sutton, 1974; Cable, 1972, etc.). Many other researchers considered a linear and positive relation between the two variables (Comanor and Wilson, 1974; Albion, 1976; Brush, 1976). And also there are several other researchers who found no significant relationship between advertising intensity and concentration (Reekie, 1975; Ornstein, 1976; Rees, 1975).

It seems that the least compatible idea with both theoretic and practical knowledge is the no-relation claim. About the researches which found a positive and direct relationship, we can guess that there could have been some shortcomings with the sampling or other parts. For example, if the majority of the sample cases used in a research be composed of the producer goods firms, we can expect the maximum concentration (CR<sub>4</sub>, for example) to be noticeably higher compared to a sample with more focus on the consumer goods firms. As shape 1 depicts it, after the advertising intensity rises to maximum among the middle point on the horizontal axis, it will fall down again as N (the number of firms in an industry) increases, and, thus, concentration ratio decreases. In producer goods, since concentration is normally higher, one can expect that number of firms in such an industry will never rise enough for the advertising intensity graph to start its downward journey.

---

<sup>1</sup> Only within consumer products, not in producer products markets

<sup>2</sup> Only within consumer products, not in producer products markets



Shape 1. The inverted U-shaped relationship between advertising intensity and concentration

$A_i/R_i$  stands for Advertising/Revenue and  $N$  stands for number of firms in an industry (opposite of concentration notion).

Price-cost margin is one of the other variables frequently chosen by scholars to account for the market structure. Theoretically, we expect to find a positive relationship between advertising intensity and price-cost margin. The higher this margin is for an industry, the industry is expected to have more cash to spend on advertising. Of course, as with many other S-C-P relationships, we can expect this relationship to work the other way round, too. That is higher advertising intensity may lead to wider price-cost margins. However, the analysis of such bidirectional relationships is beyond the scope of this paper.

Among the common representative variables for product features, product durability has been used widely by researchers. We expect that the more durable a product is, the less it is advertised. So, if we divide our sample into industries producing durable and non-durable goods, we expect to see that industries producing non-durable goods generally have higher advertising intensities. That is durability has a negative relationship with advertising intensity. This stands to reason both from a common sense point of view and also relying on past scholarly investigations.

Nazari and Tajdini (2011) investigated the relationship between advertising intensity and six independent variables, namely concentration, concentration squared, added value, statism ratio, durability and investment. Statism ratio in this research was calculated by dividing the whole number of firms in that industry by the number of the state-run firms in an industry. They hypothesized that since state-run firms enjoy more of the governmental subsidies and other forms of aids, they feel less need for competition and thus advertising. Though they found a negative relationship between being state-run and advertising intensity as expected, the relationship is not sufficiently strong regarding the standard errors. Investment and concentration took the expected signs and both had statistically significant relationships with advertising intensity, a positive and inverted U-shaped relationship, respectively. According

to this research, added value has a very strong and very significant negative relationship with advertising intensity. This must be explainable if we consider a strong positive correlation between concentration and added value. The sample for this research consisted of 42 cases chosen randomly from 4-digit ISIC Iranian industries and OLS regression was used.

According to Misra (2010), who conducted his research using data from 59 Indian industries from 1999 to 2008, there is a positive and linear relationship between advertising intensity and concentration. Also, Misra claimed that presuming similar concentration degrees, those industries with lower price-cost margins advertise more. One other result of this paper is that the advertising intensity is higher in those industries which have lower market share or higher investments. In addition, Misra showed that advertising intensity is relatively higher in consumer goods industries compared to producer goods industries.

Resende (2006) used 2SLS regression technique because he believed that in case of simultaneous causal relationships between the variables, simple regression would fail to provide an accurate result. Of course, unlike Resende, several other researchers have disregarded this possibility as an accuracy threat (Willis, 1998; Strickland and Weiss, 1976, etc.). Resende's research can be considered one of the most comprehensive ones in this field regarding the various variables he has used. The researcher claimed that there is a positive relationship between advertising intensity and concentration and disregards the significant relationship between advertising intensity and concentration squared, rejecting the possibility of a non-linear relationship. According to this research, the higher the number of employees in a firm is, the higher advertising intensity is expected in that firm. Unlike the majority of previous researchers, Resende found a negative and strongly significant relationship between advertising intensity and price-cost margin.

Another research aimed at investigating the relationship between advertising intensity and market structure is the one that Lee (2002) performed on 426 of Korean 5-digit industries. This researcher divided the industries into two groups: dealing with consumer or producer goods. Lee believed in using a new variable called advertising competence. This variable shows the percentage of change in a firm's brand image as the result of 1 percentage change in the advertising expenditure. That is the capability of the firm to leverage the firm's brand using advertising. Lee defined advertising competence as the covariance of advertising intensity and market share (in sales). The empirical results confirmed his theorizing and showed that in consumer goods industries, where higher advertising competence is expected, the relationship between advertising intensity and concentration has an inverted-U shape while this relationship takes a direct positive or J-shaped form in producer goods industries, where lower advertising competence is expected.

Willis (1998) stated that merely taking into account the concentration index of an industry is not enough to debate about its practical concentration level. He argued that to overcome this problem, one should also take into account the standard deviation between market shares of the firms in the industry. Since his concentration index was limited to the first 4 largest firms in an industry, he also added a variable for the standard deviation between market shares of the first 4 largest firms in an industry and expected that the lower the market share dispersion

among these firms, the higher their tendency for non-price competition, including advertising will be. Market size, market growth rate and ratio of sales to end-user consumers were also added to his model as control variables. According to Willis, the relation between concentration ratio for the four largest firms ( $CR_4$ ) and advertising intensity follows a lazy J shape while the relation between Herfindahl index for the four largest firms ( $HI_4$ ) and advertising intensity follows a positive linear shape. What is exclusively novel about Willis's research, as far as the author is concerned, is that his novel variable, market share standard deviation, turns out to have a significant negative relationship with advertising intensity, as expected. He justified this finding by arguing that probably when there is little standard deviation among the largest firms in an industry, which is suggestive of only marginal superiority among them, they try to outdistance each other by leveraging non-price competitive activities such as advertising. Willis also found an expected positive relationship between advertising intensity and price-cost margin. All in all, this researcher concluded that the highest advertising intensity happens in a highly concentrated market, with high price-cost margins and where there is little dispersion among the rivals regarding their market share.

### **3. Methodology**

#### *3.1 Data*

The sample used in the research consists of 20 of the 2-digit Iranian industries according to ISIC. The data was gathered from 2001 to 2006, making a pooled data set of 120 cases. However, the sample size reduced to 101 after outliers and cases with incomplete data were removed.

#### *3.2 Variables*

Independent variables to be used were chosen based on the writer's own judgment and review of the previous literature. Sales, direct export, added value, concentration, statism ratio and investment were chosen because we believed that they can discriminate well between different advertising categories. Sales refer to the monetary volume of sales in an industry in one year. Direct export, investment and added value are also monetary volumes of the respective variables in one year. Concentration is calculated by dividing the total number of firms in an industry by number 1.

The dependent variable, advertising intensity was calculated by dividing the monetary volume of sales by the advertising budget in each industry. In order to change this variable to a categorical variable as the statistical technique demands, the original continuous variable was categorized using the 33.3% and 66.6% percentiles. The three categories were called high advertising intensity, medium advertising intensity and low advertising intensity.

#### *3.3 Procedure*

We used SPSS 16 and the Multiple Discriminant Analysis (MDA) with the backward stepwise algorithm for the analysis. At first, all the 6 independent variables were entered but regarding their power to lower the Wilk's lambda statistic and also the significance of F,

direct export, sales and concentration turned out to be the three variables which have the highest discriminating power to add to the model. The Wilk's lambda statistic varies between 0 and 1. The more it is inclined to 0, the higher the impact of inclusion of that specific variable is. To consider the significance of such impacts, MDA also takes account of the F-to-enter of the variables that can lower the Wilk's lambda in the model. As soon as an independent variable is found to be able to lower the Wilk's lambda statistic and also has a suitable F-to-enter ( $< .05$  in our analysis), it is taken into the model. This procedure goes on until no variable that can lower the Wilk's lambda statistic while keeping statistical significance is left out of the model.

The way MDA chooses the number of discriminant functions is to choose the lower of 1) The number of independent variables and 2) The number of dependent variable categories minus 1. So, since the latter is the smaller number in our case (we have 3 categories within our dependent variable), MDA offered two discriminant functions. The first one always has a better discriminating power than the second one but since the discriminating powers of each of the two discriminant functions are independent of each other, the second discriminant function can also be of major use if it has sufficient discriminating power.

MDA and Regression analysis are somehow similar in that both try to formulate a function that has the most enhanced explanatory/discriminating power through finding the most suitable weights/coefficients for the independent variables. The weights that MDA offers in the discriminant functions can be interpreted as the Betas in regression analysis. When we have discriminant functions and the coefficients, we can calculate the discriminant score for any given case by inserting the data about the independent variables of that case in the discriminant function. Having the discriminant scores available, the last thing to do is to compare it with the group centroids of the current data pool to see to what group centroid the calculated discriminant score is closer. To any group it is closer, we predict that that case will fall in that group. Moreover, SPSS provides us with some more visual outputs that make the classification prediction even easier. The last part of the statistical procedure is to find the classification success rate for the model. MDA will perform the classification for the current data pool using the derived model as if we didn't know their real classification. It is generally considered that a classification success degree that is about 20% more than the expected success degree by mere chance is satisfactory. Since we have three categories, the classification success degree by mere chance is expected to be 33.3%. So, the model will be considered successful if it yields a classification success degree of, at least, 53%.

## **4. Results**

### *4.1 The Preliminary Analyses*

Our valid data to enter the analysis reduced to 101 from 120 after the outliers and the cases missing data were removed. Table 1 shows the group statistics in this analysis.



Table 1. Group statistic

Adv. Int. category	Mean	Std. Deviation	Valid N (listwise)		
			Unweighted	Weighted	
Low advertising intensity	dir.export	5.9623	.82397	28	28.000
	investment	5.6573	.89523	28	28.000
	sales	7.3856	.60215	28	28.000
	concentration	-1.8286	.59530	28	28.000
	added.value	7.0194	.57252	28	28.000
	statism.ratio	-.8363	.40377	28	28.000
Medium advertising intensity	dir.export	6.2214	.69692	38	38.000
	investment	5.7082	.74227	38	38.000
	sales	7.0287	.59222	38	38.000
	concentration	-2.3283	.40353	38	38.000
	added.value	6.6946	.57232	38	38.000
	statism ratio	-1.0588	.37625	38	38.000
High advertising intensity	dir.export	4.3540	.99658	35	35.000
	investment	4.7904	.62411	35	35.000
	sales	6.1930	.59415	35	35.000
	concentration	-1.7106	.41342	35	35.000
	added.value	5.8898	.56092	35	35.000
	statism ratio	-.9299	.28866	35	35.000
Total	dir.export	5.5025	1.19080	101	101.000
	investment	5.3761	.85837	101	101.000
	sales	6.8381	.76883	101	101.000
	concentration	-1.9757	.54032	101	101.000
	added.value	6.5057	.73272	101	101.000
	statism ratio	-.9525	.36466	101	101.000

A preliminary judgment is that those variables whose means are noticeably different among different groups must be the best to discriminate between the groups. Though this idea stands to reason, and is actually the basis for MDA, judgment cannot be made so easily without

taking into account the standard deviations and the statistical significance of such inter-group differences.

Table 2 shows the steps MDA follows in the stepwise algorithm to find the variables that raise the discriminating power of the model better than the others.

Table 2. The steps in the stepwise algorithm (variables NOT in the model)

Step		Tolerance	Min. Tolerance	Sig. of F to Enter	Wilks' Lambda
0	dir.export	1.000	1.000	.000	.494
	investment	1.000	1.000	.000	.750
	sales	1.000	1.000	.000	.588
	concentration	1.000	1.000	.000	.733
	added.value	1.000	1.000	.000	.590
	statism ratio	1.000	1.000	.043	.938
1	investment	.806	.806	.541	.488
	sales	.690	.690	.000	.413
	concentration	.910	.910	.000	.421
	added.value	.678	.678	.000	.417
	statism ratio	.994	.994	.033	.461
2	investment	.261	.224	.000	.339
	concentration	.389	.295	.000	.206
	added.value	.040	.040	.995	.413
	statism ratio	.906	.629	.002	.364
3	investment	.206	.204	.794	.205
	added.value	.039	.034	.610	.204
	statism ratio	.714	.288	.569	.204

We see that in the step 0, no variables are in the model. Actually, MDA uses a backward stepwise algorithm. Dir. export (Direct export) is the first variable that enters the model because it fulfills the two entrance criteria: 1) it is the variable that reduces the Wilk's lambda statistic the most and 2) the test is significant based on the  $F < .05$  criterion. This procedure continues until, in the third stage, no variable is there to fulfill both criteria. Added value and statism ratio are there and actually reduce Wilk's lambda statistic to .204 but we see that the Wilk's lambda test is not at all significant ( $p=.61$  and  $p=.57$ ).

MDA shows the explanatory power of each function through canonical correlations. Table 3 shows one the most important, if not the most important, outputs from MDA.

Table 3. Eigenvalues and Canonical Correlations

Function	Eigenvalue	% of Variance	Cumulative %	Canonical Correlation
1	1.378 <sup>a</sup>	57.0	57.0	.761
2	1.038 <sup>a</sup>	43.0	100.0	.714

a. First 2 canonical discriminant functions were used in the analysis.

Simply put, the canonical correlation shows the Pearson Product Moment correlation between the discriminant scores and the dependent variable categories as numbered 1, 2 and 3 (Landau et al, 2003). In other words, for example, using the first discriminant function, 57% of variances in the discriminant scores is explained by the group differences and using the second discriminant function, 43% of variances in the discriminant scores is explained by the group differences

Eigen values are what we like to see maximized. They are the ratios between inter-group variances and in-group variances. This is the ANOVA concept in MDA.

MDA's counterparts for Betas in regression analyses are called standardized canonical coefficients. Table 4 shows the standardized canonical coefficients in our model.

Table 4. standardized canonical coefficients

	Function	
	1	2
dir.export	-.208	-.983
sales	1.680	.425
concentration	1.381	.595

We see that sales is the most important variable in the first discriminant function and direct export is the most important variable in the second discriminant function. Now the question is that how these functions and their coefficients can help us categorize a case into one of high advertising intensity, medium advertising intensity or low advertising intensity industries. This can be done, in its most basic form, using both table 1 and table 4 at the same time. Having in mind that sales is the most important variable in the first discriminant function, If we take a look at table 1 and compare variable sales among different groups, we'll see that sales has the largest mean in group 1 (low advertising intensity). So, regarding the positive sign of this variable in the first discriminant function, the larger the sales is, the larger the first discriminant function ( $DF_1$ ) will be. In addition, regarding the negative sign of direct export, we expect the  $DF_1$  to rise more as direct export decreases. Comparing group 1 and 2 in this regard, again we can conclude that higher  $DF_1$  is more indicative of a case belonging to the first category. Though there is a kind of competition between groups 1 and 3 regarding

concentration and direct export, it seems that the larger coefficient of sales and the larger mean of this variable in group 1, makes the classification of a case into group 1 because of its higher  $DF_1$  very probable. The same analysis can be made for the second discriminant function.

One other simpler way to predict group membership for a new case is to calculate its discriminant score and compare it to the mean of discriminant scores for each group, also called group centroid. The new case will belong to each group whose centroid is closer to the case's discriminant score. Table 5 shows the group centroids in our analysis.

Table 5. Group Centroids

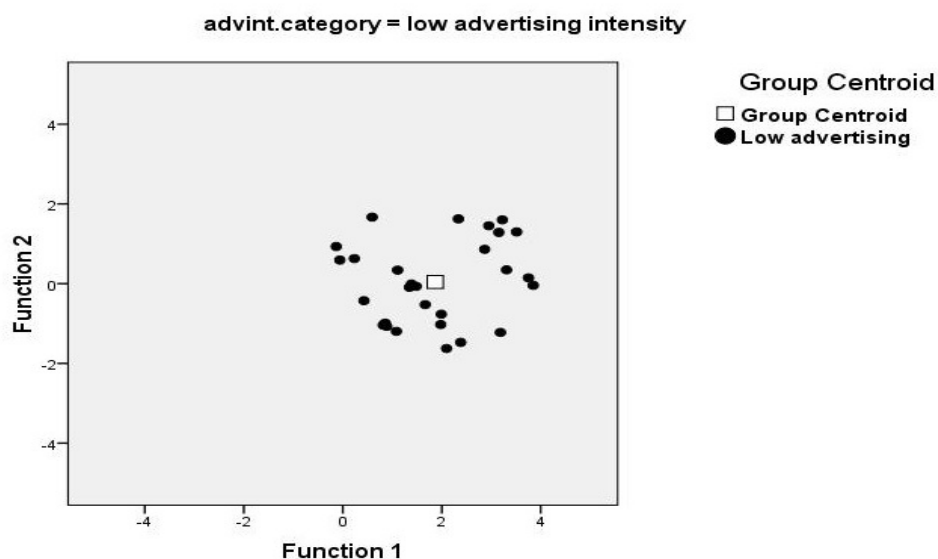
advint.category	Function	
	1	2
Low advertising intensity	1.866	.043
Medium advertising intensity	-.681	-1.149
High advertising intensity	-.754	1.213

Unstandardized canonical discriminant functions evaluated at group means

Using table 5, for example if the  $DF_1$  score for a new case turns out to be 1.9, we predict that the case will belong to the first group because 1.9 is closest to 1.866 in  $DF_1$ , which is the centroid for the group 1.

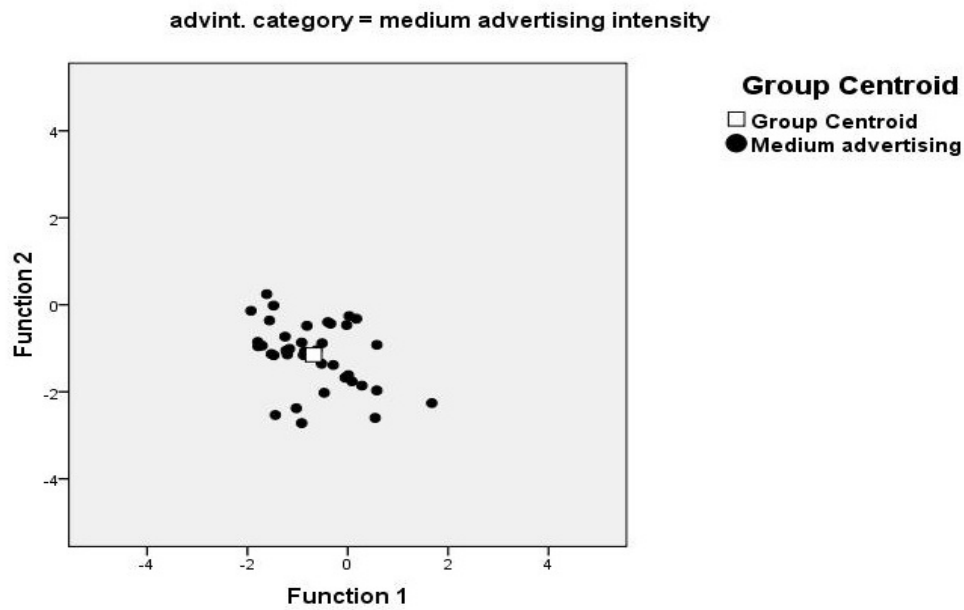
Shapes 2 through 4 show the concept of group centroids visually.

**Canonical Discriminant Functions**



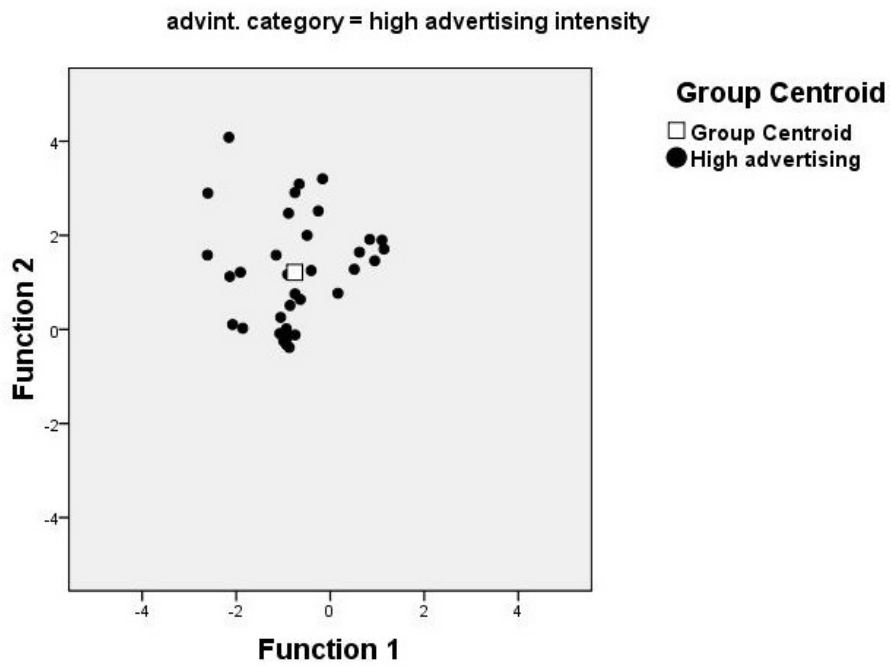
Shape 2. Group centroid for low-advertising group

**Canonical Discriminant Functions**



Shape 3. Group centroid for medium-advertising group

**Canonical Discriminant Functions**

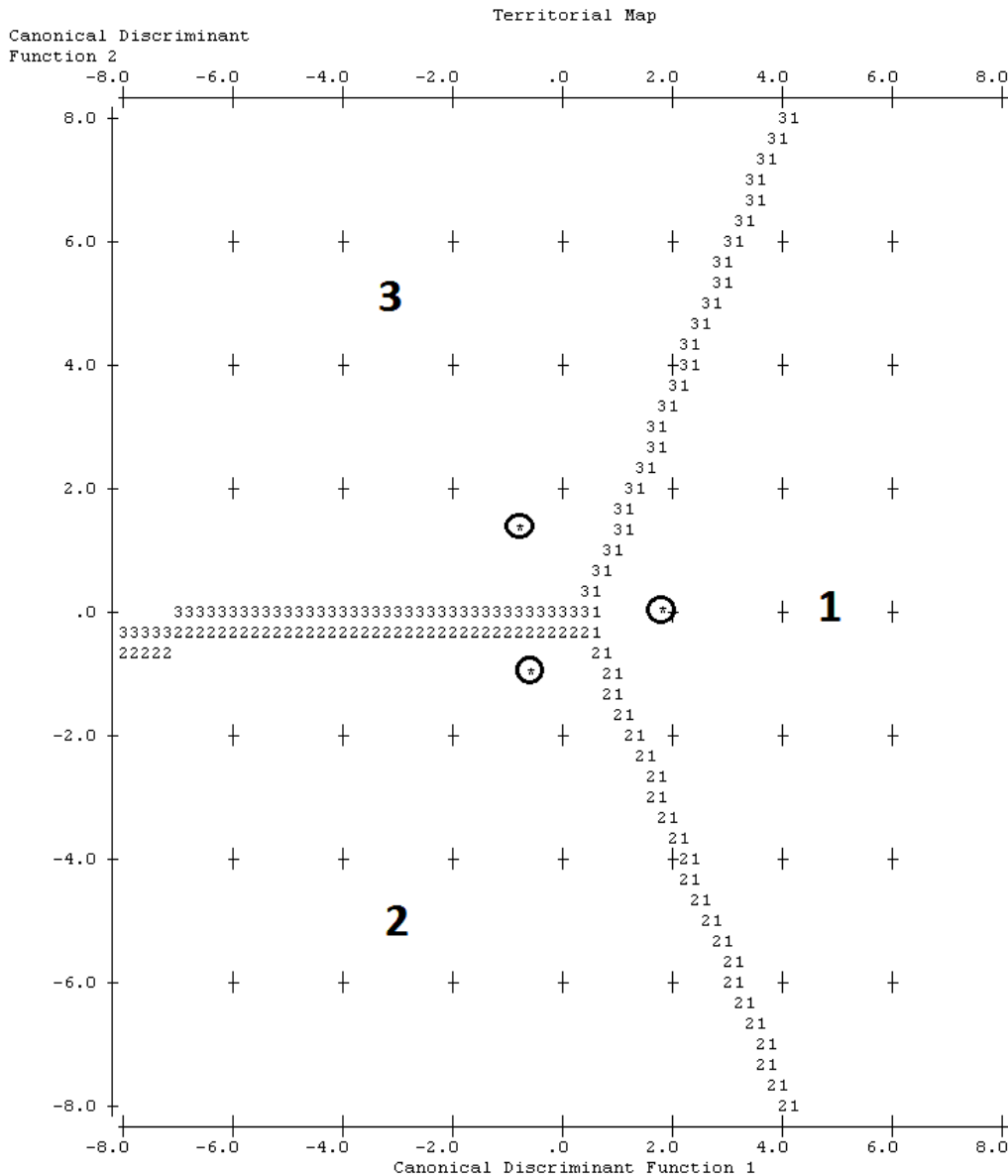


Shape 4. Group centroid for high-advertising group

From shape 2, we clearly see that a positive  $DF_1$  and a positive  $DF_2$  score suggest that the case will most probably belong to the first category, the low advertising industries. It would be more accurate to say that a case whose  $DF_1$  score is positive and whose  $DF_2$  score lies somewhere between -2 and +2 is most probably expected to belong to the first advertising intensity category.

We see that concluding group 3 and 4 memberships is not as clear-cut as it is with shape 2 and group 1 membership. The point is that  $DF_1$  group centroids for groups 2 and 3 are almost the same (somewhere like -0.5 on Function 1 axis). That's why it's said that  $DF_1$  cannot perform well in discriminating between groups 2 and 3. In this condition, the only way to distinguish memberships between these two groups is to use centroids in  $DF_2$ . We know that the second function does not have as much discriminating power as function 1 does. So, that's why prediction of group memberships between groups 2 and 3 is more difficult and more prone to error.

Perhaps one far simpler method of prediction in MDA is using the territorial map. Shape 5 shows this map in our analysis.

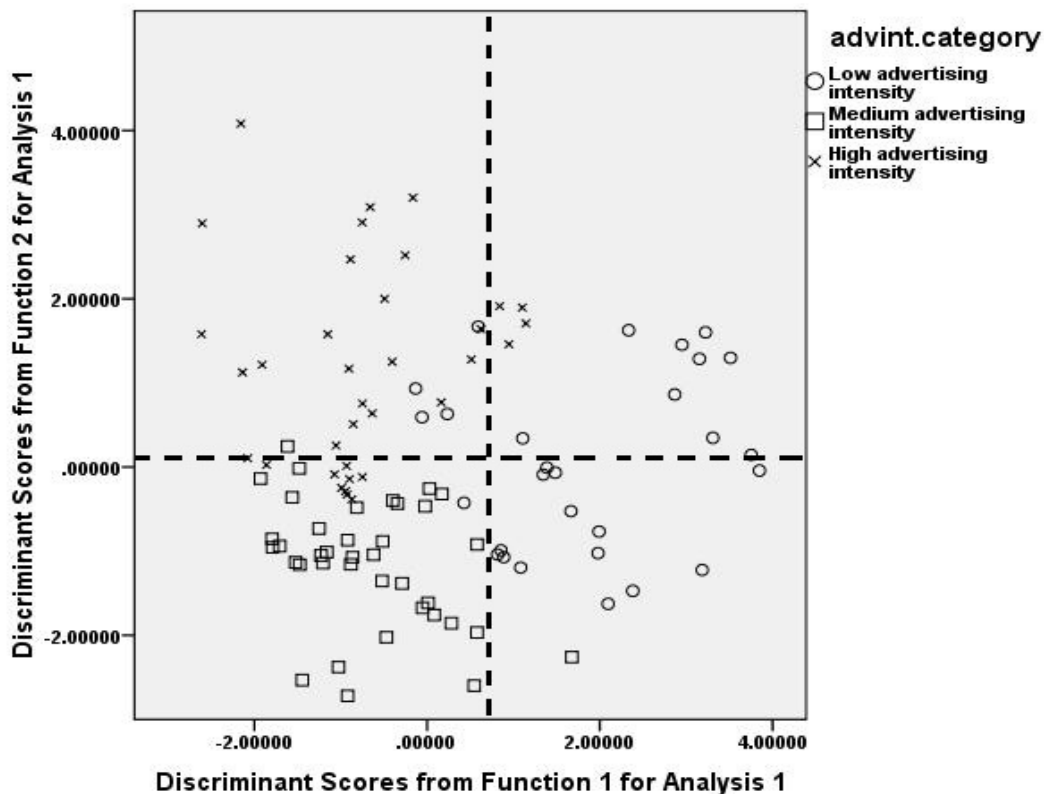


Shape 5. The Territorial Map

The map has the  $DF_1$  scores on the horizontal axis and  $DF_2$  scores on the vertical axis. The three circles show the position of the centroids and the numbers how the territory of each group is expanded. This figure also imply the superiority of  $DF_1$  regarding its discriminating power. For example, if the  $DF_1$  score for a new case is calculated to be 6, we can predict that it will belong to the first group, low advertising industries. We can do this prediction needless of the  $DF_2$  score since the discriminating power of  $DF_1$  is high enough. But we can't do the same with  $DF_2$ . For example, if the  $DF_2$  score for a new case turns out to be -4, we might be wrong if we predicted it would belong to the second category of advertising intensity. For the same case, if the  $DF_1$  score is -2, for example, the case will most probably belong to the second category but if the  $DF_1$  score is, say, +6.5, the case will most probably belong to the

first category. That’s why it is easier to predict group memberships with DF<sub>1</sub> score alone but more difficult with DF<sub>2</sub> score alone.

Shape 6 is still another way of showing the same thing. In this figure, each case is plotted in the space whose axes are the discriminant function scores. We see that taking into account the first function (the horizontal axis), cases belonging to group 1 are successfully discriminated from cases belonging to the other two groups with little, acceptable overlap (prediction error).



Shape 6. The plot of each case in the space with discriminant function scores on axes

We see that the highest overlap regarding group 1, is between this group and the third group. One possible justification may be the little difference between the concentration means of these two groups (only .11). The larger overlap is seen between the second and third group, which is due to lower discriminating power of the second discriminant function, which is responsible to act between groups 2 and 3 since discriminant function 1 is unable to do so.

#### 4.2 The Final Model

Having the standardized canonical coefficients available, the task of writing the discriminant functions to form the final model is as easy as doing so with regression models when we have the Betas available.

$$DF (1) = -.208 DX + 1.68 S + 1.381C \tag{1}$$



$$DF (2) = -.983 DX + .425 S + .595 C \quad (2)$$

In the formulas above, DF stands for discriminant function, DX for direct export, S for sales and C for concentration. In the first function, one unit increase in DX, S and C will lead to 20.8% of a unit decrease, 168% of a unit increase and 138.1 % of a unit increase in the DF<sub>1</sub> score, respectively.

In the second function, one unit increase in DX, S and C will lead to 98.3% of a unit decrease, 42.5% of a unit increase and 59.5 % of a unit increase in the DF<sub>2</sub> score, respectively.

#### *4.3 Test of Classification Accuracy*

SPSS can perform two tests to evaluate the prediction accuracy of the model: resubstitution estimate and, also, the cross-validation technique. The former tries to determine group membership of what percentage of the original data can be predicted correctly using the model. The latter takes out one single case, constructs the model without it and then tries to predict the group membership of that left-out case using the new model. This will be done for all cases repeatedly until the accuracy rate is calculated. The cross-validation technique provides a more honest image of the model since predicting the group membership of a case at the presence of which the model has been created, as with the resubstitution estimate technique, leaves much room to worry about the accuracy of the results. This problem is tackled in the cross-validation technique. Table 6 shows the results from the classification test.

Table 6. Classification Results

				Predicted Group Membership			Total
				Low advertising intensity	Medium advertising intensity	High advertising intensity	
advint.categ							
Original	Count	Low intensity advertising	22	2	4	28	
		Medium intensity advertising	1	36	1	38	
		High intensity advertising	2	8	25	35	
	%	Low intensity advertising	78.6	7.1	14.3	100.0	
		Medium intensity advertising	2.6	94.7	2.6	100.0	
		High intensity advertising	5.7	22.9	71.4	100.0	
Cross-validated <sup>a</sup>	Count	Low intensity advertising	22	2	4	28	
		Medium intensity advertising	1	35	2	38	
		High intensity advertising	3	10	22	35	
	%	Low intensity advertising	78.6	7.1	14.3	100.0	
		Medium intensity advertising	2.6	92.1	5.3	100.0	
		High intensity advertising	8.6	28.6	62.9	100.0	

b. 82.2% of original grouped cases correctly classified.

c. 78.2% of cross-validated grouped cases correctly classified.

The numbers on the diagonal show the classification results, both in absolute values and percentages. One notable point is that we can see in the cross-validation section of the table (the lower half) that cases belonging to group 1 were predicted as group one 79% of times, were predicted as group two 7% of the times and were predicted as group three 14% of the

times. So, we see that the main error of the model in predicting group 1 membership is when it tries to discriminate between this group and group 3, not group 2. This is in line with what we said about the overlapping severity when talking about shape 5 in the previous paragraphs.

The approximate success degrees using the resubstitution estimate and cross-validation are 82% and 78%, respectively. It is very natural for the success degree to fall when using the cross-validation method since this method exposes the model to a more realistic and challenging classification test. Nonetheless, 78% is still satisfactorily higher than the predicted success degree when classifying cases by chance (33.3%). So, we consider this model sufficiently helpful and reliable in predicting the group membership of an industry with regards to its advertising intensity.

## **5. Discussion**

The purpose of this paper is to identify the variables using which, one can predict the group membership of an industry regarding its advertising intensity: high advertising intensity, medium advertising intensity or low advertising intensity industries. Using MDA, we devised a model with which we can predict this group membership with an acceptable accuracy degree of 78% or 82%. We conclude that the higher the sales are in an industry, the higher the probabilities that it belongs to those industries with lower advertising intensity are. This is in line with results from some previous studies that advertising intensity has a negative relationship with market size. Willis (1998) considers a negative relationship between advertising intensity and market size and raises the economy of scale of advertising as one possible reason for this observed relationship. As another example, Pagoulatos and Sorensen (1981) found a negative relationship between concentration and market size. One other justification can be that as market size grows, more firms are tempted to enter the market and thus the market approaches total competition form. Following the classical theory expressed in shape 1 of the second section of this article, we expect advertising intensity to fall in such markets. Thus, one can use these two notions to believe that as market size (which can be expressed in sales) grows, advertising intensity falls. Another conclusion we can make from the analysis of our data is that the lower the direct export is in an industry, the higher the  $DF_2$  score will be for a case from that industry, which, in turn will make the model more prone to predict it as belonging to the low advertising intensity industries. It seems that those industries that engage in direct export more vigorously tend to spend a higher portion of their sales on advertising. One possible justification can be the fact that as firms increase their direct export and broaden their market, they may feel more need to invest on brand awareness and product awareness.

## **6. Implications and Applications**

One main source of application of this model is within advertising industry. For sure, advertising agencies and other bodies involved in the advertising industry not always have sufficient information about the advertising figures in different firms and industries. Using this model, we can, at least, predict in what category of advertising intensity a firm falls based on the data about its sales, direct export and concentration.

This classification can also act as a benchmark for firms. Knowing what ranges of advertising intensity each advertising intensity category is defined for, a firm can see if its advertising intensity is in line with the industry's norm, is behind it or is ahead of it. For example, a firm which finds that it will fall in the medium advertising intensity category (which is defined as an advertising intensity between 8% and 20%, for example) and its current advertising intensity is below 7% may feel urged to boost its advertising relative to its sale.

## **7. Limitations and Recommendations for Further Research**

We encourage other researchers to conduct similar studies using Artificial Neural Networks (ANN) technique instead of MDA or, preferably, besides it. Some researchers, including Manel and others (1999) claim that ANN performs better than MDA and Regression Logistics methods and offers a model with higher accuracy rate. Nevertheless, as far as the choice of data and the dataset, in general, differ from research to research, a definite conclusion cannot be made as which technique yields better results. Thus, performing both techniques on a single dataset can be very beneficial in this regard.

We couldn't use the Herfindahl index as the measure for market concentration due to lack of data about market shares. Instead, the simple concentration index was used. However, the Herfindahl index seems more robust a measure for concentration since it can be defined accurately and takes into account the market shares of the largest firms in an industry and that's why a large number of researchers have used this variable (For example, Greer, 1971; Cable, 1972; Sutton, 1974; Strickland and Weiss, 1976; Martin, 1979; Buxton et al., 1984; Willis and Rogers, 1998). We recommend other researchers to carry out similar researches using the Herfindahl index.

## **References**

- Bhatia, T. K. (2000). *Advertising in Rural India: Language, Marketing Communication, and Consumerism*. Tokyo Press, Tokyo, Japan.
- Buxton A. J.; Davies, S. W.; Lyons, B. R. (1984). Concentration and Advertising in Consumer and Producer Markets. *The Journal of Industrial Economics*, 32(4), 451-464. <http://dx.doi.org/10.2307/2098229>
- Cable, J. (1972). *Market Structure, Advertising Policy, and Inter-Market Differences in Advertising Intensity*. In *Market Structure and Corporate Behaviour*, edited by K. Cowling. London: Grays Mills.
- Greer, Douglas F. (1971). Advertising and Market Concentration. *Southern Economic Journal*, 38(1), 19-32. <http://dx.doi.org/10.2307/1056065>
- Landau, S. and Everitt, B. S. (2003). *A handbook of statistical analyses using SPSS*. Chapman & Hall/CRC Press LLC
- Lee, C. (2002). Advertising, its determinants, and market structure. *Review of Industrial Organization*, 21, 89-101. <http://dx.doi.org/10.1023/A:1016020026207>
- Manel, S., Dias, J.-M., & Ormerod, S. J. (1999). Comparing discriminant analysis, neural

networks and logistic regression for predicting species distributions: a case study with a Himalayan river bird. *Ecological Modelling*, 120(2-3), 337-347. doi:10.1016/S0304-3800(99)00113-1

Martin, S. (1979). Advertising, Concentration, and Profitability: The Simultaneity Problem. *The Bell Journal of Economics*, 10(2), 639-647. <http://dx.doi.org/10.2307/3003355>

McCann-Erickson U.S. Advertising Volume Reports and Bob Coen's Insider's Report for December 2001

Misra, S. D. (2010). Advertising and Market Structure: A Study of the Indian Consumer Goods and Services Sector. *International Research Journal of Finance and Economics*, 42.

Nazari, M., & Tajdini, S. (2011). An Empirical Investigation of the Relationship between Market Structure and Advertising Intensity. *Interdisciplinary Journal of Contemporary Research in Business*, 3(3), 1417-1427.

Pagoulatos, E., & Sorensen, R. (1981). A simultaneous equation analysis of advertising, concentration and profitability. *Southern Economic Journal*, 47(3), 728-741. <http://dx.doi.org/10.2307/1057367>

Resende, M. (2006). The determinants of advertising intensity in the Brazilian manufacturing industry: an econometric study. *Nova Economia*, 16(3), 407-422. <http://dx.doi.org/10.1590/S0103-63512006000300002>

Scherer, Frederic M. & Ross, David 1990. *Industrial Market Structure and Economic Performance*, 3rd ed. Houghton Mifflin Company

Strickland, A. D., & Weiss, L. W. (1976). Advertising, concentration, and price-cost margins. *The Journal of Political Economy*, 84(5), 1109-1121. <http://dx.doi.org/10.1086/260499>

Sutton, C. J. (1974). Advertising, Concentration and Competition. *The Economic Journal*, 84(333), 56-69. <http://dx.doi.org/10.2307/2230483>

Willis, M. (1998). Market share dispersion among leading firms as a determinant of advertising intensity. *Review of Industrial Organization*, 13, 495-508.

Willis, Michael S.; Rogers, Richard T., Market Share Dispersion among Leading Firms as a Determinant of Advertising Intensity. *Review of Industrial Organization*, 13(5), 495-508.