SALES AND EARNINGS PATTERNS FOR RETAIL STORES IN METROPOLITAN AND NON-METROPOLITAN TRADE AREAS

By EMMANUEL C. KONDYLIS, Ph. D.
Deree-Pierce College

Abstract

A regression method for estimating sales and earning performance for retail stores in metropolitan and non-metropolitan trade areas is developed. It was found that, contrary to what is generally believed, metro and non metro stores have the same sales behavior while their earnings patterns are different.

* * *

The objective of this study was to develop models to estimate sales and earnings performance for retail stores. Extensive research into retail sales predictive models at the zip (postal area) district/zone, prestige mall (shopping centre) non-prestige mall, non-mall, metro (metropolitan area), non-metro and company-wide trade level was performed, with data supplied especially for this research by a major U.S. retail chain stores company which, wishing to remain anonymous, will be called subsequently Alpha Corporation.

Detailed data were obtained for more than 200 retail department stores opened between 1958 and 1973. After the examination of a fairly large variety of simple sales estimation models it was concluded that the most satisfactory, in terms of both intuitive form and statistical fit, was the following :

$$S = a.F^{b}.M^{c}$$

where : S = Alpha store's sales ($)

F = Store's selling space (sq. ft.)

M = Market sales within store's trading area.

The coefficients a, b, c are positive with $b + c$ near 1.

For the estimation of earnings (E) the model used was :

$$E = n.S - z.F$$

wherc the meaning of t. e parameters n and z is :

n = gross earnings rate per sales dollar

z = incremental cost per sq. ft.

It was found that the inclusion of the average inventory, (or the turnover) in the above models resulted in implications that were not intuitive. A result worth noticing is that the 23 «prestige» stores had a clearly distinctive earnings function reflecting their high breakeven point ($71.62/sq. ft.) and also high incremental profitability above this point. This was not true for the 48 non-prestige air-conditioned mall stores whose regression coefficients were between the ones for the metro and non-metro stores, and the 83 non-air conditioned mall stores, whose regression coefficients were similar to those of non-metro stores.

An exhaustive investigation of the possibility of estimating store sales by zip trade area was performed. The Detroit metropolitan district was chosen for (1) its size - 12 stores, (2) its relative freedom from neighborhood stores, and (3) its even geographic distribution of Alpha stores.

Before proceeding in the evaluation of a relatively complicated model, taking into consideration overlapping, trade areas, (i.e., zip sales generated from various Alpha stores) we tried a rather simple one which didn't allow trade area overlaps. It was felt that this was the proper action since it was necessary to pull manually a considerable amount of data even in the simple case with non-overlapping trade areas. As a rule of thumb it was assigned one zip to a specific store if more than 50 % of the Alpha credit sales for the zip were generated from this store. As a measure of market potential it was created a zip Market Sales figure by allocating the Market Sales of the store's trade area according to the number of households in the zip. It was also used as alternative measure of market potential the total zip population and total zip family income. The zip sales were estimated as a function of the market potential, as measured from one of the above measures, and the distance of the zip from the store. This was done for each one of the 12 stores as well as for all the stores grouped together. In the last case the selling space was included as a predictive variable. The results in all cases were unsatisfactory. The coefficients had great variability from store to store and they were mostly non-intuitive. Speculating that the problem may lie in the presence of strong multicolinearity among zip variables of a particular store, it was hypothesized the other extreme, i.e., perfect correlation among zip variables belonging to a particular store. The results were non-intuitive in this case too.

Convinced that one cannot work with the current analytical approach and available data at the zip level, it was tried the prediction of store sales for an entire metro. The Cleveland metro was chosen for the large number of stores in this metro (13). Once again the results were poor. Speculating again that the problem may be high correlation of the variables within a metro, the average values for the relevant variables were computed for each metro across the country and similar regressions were performed with this data. The results were still unsatisfactory.

Subsequently a higher level of aggregation was tried by regressing on metro and non-metro stores, with very satisfactory results this time. Speculating that stores have non-linear growth patterns during their early years, we have eliminated all stores opened less than two years from the year that the data was obtained (1973). The result was a drastic improvement in statistical fit for both, metro and non-metro stores. A close examination of the coefficients between metro and non-metro models revealed that the difference became narrower with the removal of the recently opened stores. We thought that this was probably not accidental and that the age factor was still very important in spite of the fact that we had already restricted the two groups by including only stores opened after 1958. Bringing the store age explicitly into the regressions, we obtained further improvement in statistical fit. Moreover, the coefficients of the regressions for metro and non-metro stores became very similar. Speculating further that, contrary to what was generally believed, metro and non-metro stores have the same sales behavior, we grouped all stores together. This resulted in the best fit that we had ever had for a sales predictive model (R = 91.11). However, the grouping of the corresponding earnings and marginal earnings models produced an inferior fit to the separate metro and non-metro earnings models.

The following models were derived with all variables but t (age of store in years) in thousands of their respective units :

For all Alpha stores opened between 1958 and 1971 inclusive :

$$S = (34.5480117) * F^{(.87942)} * M^{(.10720)} * e^{(.02117) * t}$$

For metro stores opened between 1958 and 1971 inclusive :

$$E = (.17521)*S - (10.82052) *F$$

For non-metro stores opened between 1958 and 1971 inclusive :

$$E = (.16023)*S - (8.90914) *F.$$

The sales model is simple and reasonable. In addition, it exhibits diminishing marginal returns to both selling space (F) and market sales (M).

The breakeven point for metro stores marginal earnings is :

$$BE = \frac{10.82052}{.17521} = \$ \ 61.76/\text{sq.ft.}$$

Where .17521 is the gross earnings rate per sales dollar, and 10.82052 is the incremental cost per sq. ft.

Similarly for non-metro stores the breakeven point is :

$$BE = \frac{8.90614}{.16023} = \$\ 55.58/\text{sq. ft}$$

A technical documentation of the above relations is included in tables 1 and 2.

TABLE I

REGRESSION ANALYSIS FOR SALES MODEL

$$\log S = \log a + b.\ \log F + c.\ \log M + d.t$$

Variable	Reg. Coef.	Std. Error Coef.	Computed T	Beta Coef.
log F	0.87942	0.06010	14.63283	0.72912
log M	0.10720	0.03042	3.52444	0.17970
t	0.02117	0.00522	4.05789	0.12160
Intercept		3.54235		
Multiple Correlation		0.91110 (Adjusted R =	0.91019)	
Std. Error of Estimate		0.27091 (Adjusted SE =	0.27222)	

ANALYSIS OF VARIANCE FOR THE REGRESSION

Source of Variation	D.F.	Sum of Sq.	Mean Sq.	F Value
Attributable to regression	3	73.509	24.503	333.870
Deviation from regression	205	15.045	0.073	
Total	208	88.554		

Table 1 presents the regression analysis for the sales model in linear form The ratio of the explained variation to the total variation is given by the squared multiple correlation coefficient (Coefficient of determination) $R^2=(0.91)^2=0.83$. This means that the model «explains» about 83 % of the variation in the store sales data. The measure of the precision of the prediction about the mean value of the logarithm of the store sales is given by the standard error of estimate : $\hat{\sigma} =$. 0.27.

TABLE 2

REGRESSION ANALYSIS FOR EARNINGS MODELS

$$(E/F) = m.(S/F) + n$$

A. METRO STORES

Variable	Reg. Coef.	Std. Error Coef.	Computed T	Beta coef.
S/F	0.17521	0.01140	15.37496	0.85497
Intercept		—10.82052		
Multiple Correlation		0.85497	(Adjusted R =	0.85497)
Std Error of Estimate		3.44287	(Adjusted SE =	3.44287)

ANALYSIS OF VARIANCE FOR THE REGRESSION

Source of variation	D.F.	Sum of Sq.	Mean Sq.	F Value
Attributable to regression	1	2802.005	2802.005	236.389
Deviation from regression	87	1031.241	11.853	
Total	88	3833.246		

B. NON-METRO STORES

Variable	Reg. Coef.	Std. Error Coef.	Computed T	Beta Coef.
S/F	0.16023	0.00848	18.90115	0.86701
Intercept		—8.00914		
Multiple Correlation		0.86701 (Adjusted R =	0.86701)	
Std. Error of Estimate		2.60317 (Adjusted SE =	2.60317)	

ANALYSIS OF VARIANCE FOR THE REGRESSION

Source of Variation	D.F.	Sum of Sq.	Mean Sq.	F Value
Attributable ato regression	1	2420.920	2420.920	357.253
Deviathon from regression	113	799.624	6.776	
Total	119	3220.544		

Table 2 presents the regression analysis for earnings models for both, metro and non-metro stores. For the metro stores the coefficient of determination is $R^2 = 0.73$ while the standard error of estimate is : $\hat{\sigma} = 3.44$.

For the non-metro stores the corresponding values are : $R^2 = 0.75$ and $\sigma = 2.60$.

It is worth noticing the values of the regression coefficients of the two store categories as well as the corresponding intercepts. It can be seen that the gross

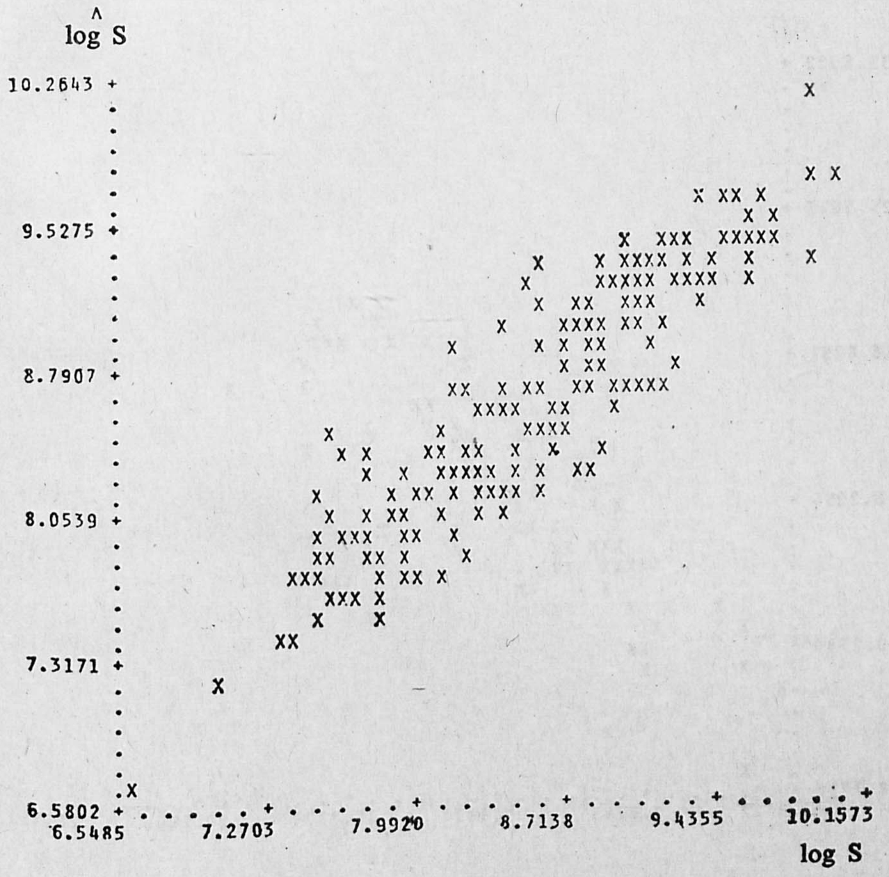

FIGURE 1 : Scatter diagram showing the relationship between the estimated logarithm of Sales (lôg s) and the actual one (log s) for all stores considered.

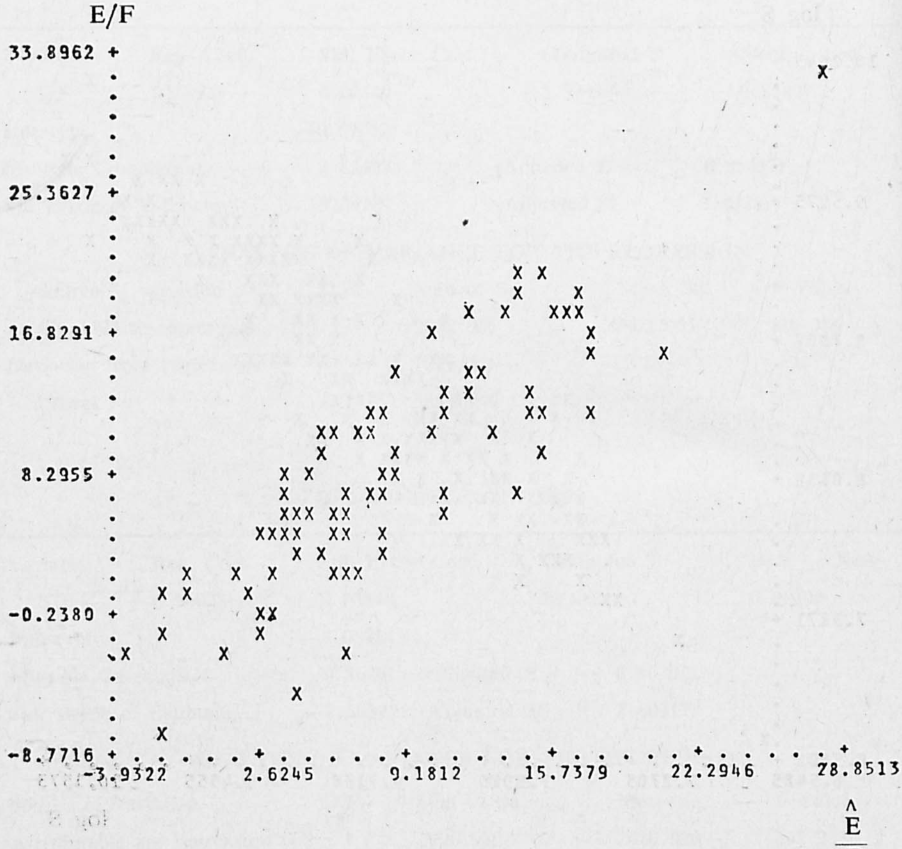

FIGURE 2 : Scatter diagram showing the relationship between the actual Earnings per sq. ft. (E/F) and the estimated ones (Ê/F) for metro stores.

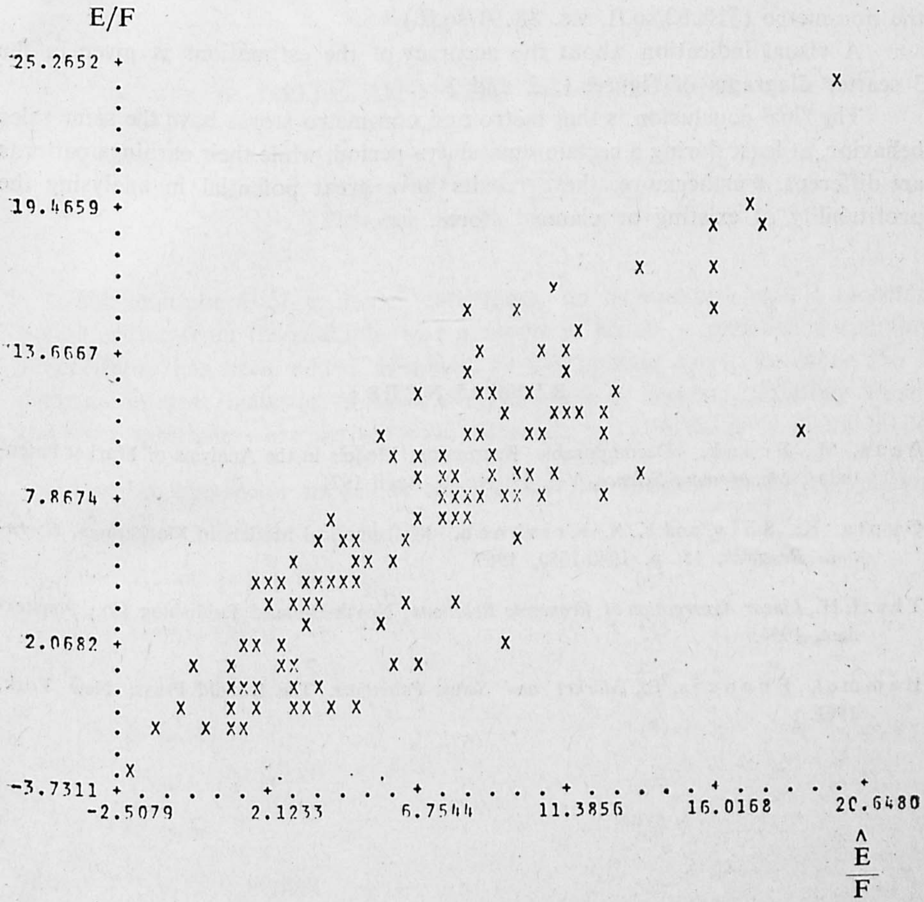

FIGURE 3 : Scatter diagram showing the relationship between actual Earnigs per sp. ft. (E/F) and the estimated ones (Ê/F) for non-metro stores.

earnings rate of the metro stores is 17.5 % while for the non-metro is 16 % per dollar sales, which implies that the stores in metropolitan areas are more productive than the stores in non-metropolitan ones. It is also true that the increamental cost per square foot for the metro stores is higher than the one for the non-metro ($10.82/sq.ft. v.s. $8.91/sq.ft.)

A visual indication about the accuracy of the estimations is given in the 3 scatter diagrams of figures 1, 2 and 3.

The final conclusion is that metro and non-metro stores have the same sales behavior, at least during a certain «maturity» period, while their earnings patterns are different. Furthermore, these results have great potential in analysing the profitability of existing or planned stores.

REFERENCES

Bass, M., Frank, «Decomposable Regression Models in the Analysis of Market Potentials», *Management Science*, Vol. 17, No. 8, April 1971.

Gupta, K., Shiv, and K. S. Krishnan, «Mathematical Models in Marketing», *Operations Research*, 15, p. 1040-1050, 1967

Theil, H., *Linear Aggregation of Economic Relations*, North-Holland Publishing Co., Amsterdam, 1954.

Hammel, Francis, E., *Market and Sales Potentials*, The Ronald Press, New York, 1961.