Taylor & Francis
Taylor & Francis Group

Determinants of Retail Sales in Large Metropolitan Areas, 1954 and 1963

Author(s): Ben-Chieh Liu

Source: *Journal of the American Statistical Association*, Dec., 1970, Vol. 65, No. 332 (Dec., 1970), pp. 1460-1473

Published by: Taylor & Francis, Ltd. on behalf of the American Statistical Association

Stable URL: https://www.jstor.org/stable/2284330

JSTOR is a not-for-profit service that helps scholars, researchers, and students discover, use, and build upon a wide range of content in a trusted digital archive. We use information technology and tools to increase productivity and facilitate new forms of scholarship. For more information about JSTOR, please contact support@jstor.org.

Your use of the JSTOR archive indicates your acceptance of the Terms & Conditions of Use, available at https://about.jstor.org/terms

Taylor & Francis, Ltd. and *American Statistical Association* are collaborating with JSTOR to digitize, preserve and extend access to *Journal of the American Statistical Association*

JSTOR

This content downloaded from 96.50.241.211 on Thu, 11 May 2023 02:50:52 +00:00
All use subject to https://about.jstor.org/terms

© Journal of the American Statistical Association
December 1970, Volume 65, Number 332
Applications Section

Determinants of Retail Sales in Large Metropolitan Areas, 1954 and 1963

BEN-CHIEH LIU*

This article presents several specifications of a retail sales model to explain the differential sales of retail goods and services in large Standard Metropolitan Statistical Areas in 1954 and 1963. Stepwise regression is employed and some hitherto unused regressors such as local government expenditures and taxes are introduced to compare different forms of the model specification. Tests on the stability of the relevant elasticities obtained in the model over time and on the predictive accuracy of this model are performed. The results of these tests support the stable elasticities hypothesis and show very high predictive accuracy for the model.

1. INTRODUCTION

Several studies have analyzed the relationship between retail sales, either total or per capita in various U. S. cities, and some important variables which supposedly influence sales, such as population, income and distance between cities. Population and distance between cities are two determining factors used to study the variations in retail sales in Reilly's [12] "Law of Retail Gravitation," as one example. As another, Russell [13] showed that retail sales per capita were not correlated at all with median family income among 78 communities with populations of 25,000 to 49,999 in the urbanized U. S. in 1949. The correlation coefficient she obtained was insignificant and negative, i.e., $r=-0.06$.[1]

One year later, another article presented results different from those of Russell. Ferber [5] observed a high correlation between total retail sales and total income among the 51 Illinois cities with populations of 10,000 to 4,000,000 in 1954 ($r^2=0.91$).[2] In this multivariate regression model, population was the most dominant variable, while distance and income only played a secondary role in explaining total retail sales. But in explaining per capita sales, per capita income, stores per 10,000 residents and distance are principal determinants.

The relationship between income, population and retail sales was revisited one decade later by Tarpey and Bahl [16]. They support Ferber's findings that population, distance and number of stores have strong effects on variations in total sales in 136 cities with populations ranging from 5,000 to 25,000 in the four Southern states in 1958. When per capita sales are considered, they again find that income is an important determinant. On the contrary, Schwartzman [14] indicated that "(median family) income appears to have a small influence

---

* Ben-chieh Liu is project director of economic research, St. Louis Regional Industrial Development Corporation. The author wishes to thank Professors Edward Greenberg, Richard Muth and Fredric Raines of Washington University and Professor Maw Lin Lee of the University of Missouri (Columbia) for their comments on this article. He is also grateful to the associate editor and the referees for helpful suggestions which improved the article. The author assumes sole responsibility for any shortcomings.

[1] See [13, p. 330].

[2] See [5, Table 1, p. 297].

This content downloaded from
96.50.241.211 on Thu, 11 May 2023 02:50:52 +00:00
All use subject to https://about.jstor.org/terms

on sales per person engaged in retail trade as a whole,"[3] in 188 Standard Metropolitan Statistical Areas (SMSA) in 1958.

Although all studies mentioned are similar in the approach employed, namely, a cross-section study for a particular year, they are dissimilar in their findings on the relationship between income, population and retail sales. This dissimilarity is largely due to different model specifications and different groups of observations.

This article uses census data in the 1950's and 1960's and investigates the important determining factors which significantly affect the variations in retail sales in the 38 large SMSAs of the U. S.[4] The large SMSAs are used for several reasons. First, none of the studies cited except the last one took into consideration the observations of greater homogeneity in terms of economic activities. A SMSA is an economic entity in which economic activities relating to production, distribution and consumption may be physically integrated and meaningfully functional.[5] Second, and most important, the purpose of this article is to provide a hypothetical model for better prediction of metropolitan retail sales. For this reason it attempts to identify some important explanatory variables that jointly determine the variations in retail sales among the large SMSAs which Schwartzman's article did not attempt to identify.

Third, it is in the SMSAs that the service sector has grown very rapidly, following the urbanization process. In dealing with most urban problems, the SMSA can provide a better political and geographic boundary for efficient social and economic policies than can a single city. Particularly in this retail study, the geographical distance between cities can be omitted from the list of important determinants so that hitherto unused explanatory variables can be explored and studied. By so doing, our hypothetical model may be expected to achieve a higher accuracy of empirical prediction.

Besides constructing a useful model by which retail sales among large SMSAs may be predicted, another principal objective of this article is to estimate the partial impact of each determinant on the sales. Hopefully the implications of these estimates may be taken into account in efficient decision making relating to retail sales by businessmen in general and by local government administrators and planners in particular.

2. THE APPROACH AND THE MODEL

Multiple regression is employed to regress both total and per capita retail sales on selected variables. Since the objective of the article is not to derive a demand function, nor a supply function of retail goods and services, explanatory

---

[3] See [14, pp. 223–4].

[4] The SMSAs are: Akron, Ohio; Allentown, B., E., Pennsylvania; Atlanta, Georgia; Baltimore, Maryland; Birmingham, Alabama; Boston, Massachusetts; Buffalo, New York; Chicago, Illinois; Cincinnati, Ohio; Cleveland, Ohio; Columbus, Ohio; Dallas, Texas; Denver, Colorado; Detroit, Michigan; Grand Rapids, Michigan; Houston, Texas; Indianapolis, Indiana; Kansas City, Missouri-Kansas; Los Angeles, California; Louisville, Kentucky, Memphis, Tennessee; Miami, Florida; Milwaukee, Wisconsin; Minneapolis, Minnesota; New Orleans, Louisiana; New York, New York; Philadelphia, Pennsylvania; Phoenix, Arizona; Pittsburgh, Pennsylvania; Portland, Oregon; Rochester, New York; St. Louis, Missouri-Illinois; San Diego, California; San Francisco, California; Seattle, Washington; Toledo, Ohio; Washington, D. C.; Youngstown, Ohio. The relationship among retail sales, income and population in each of these SMSAs was also analyzed in a time-series study, see [10]. For data sources, see Footnote 10.

[5] See [3, pp. 35–8] and [10, pp. 71–3].

This content downloaded from
96.50.241.211 on Thu, 11 May 2023 02:50:52 +00:00
All use subject to https://about.jstor.org/terms

variables are selected regardless whether the variables are on the demand or supply side. However, stepwise regressions are performed according to the three types of explanatory variables selected; i.e., variables on the demand side, local government variables, and supply, educational and regional variables. The estimated partial impact may be used to indicate whether one type of variable is more important in affecting regional variations in retail sales. Although only the last regression that includes all explanatory variables is considered complete, results from each step are reported individually in order to make comparisons among model specifications, estimated coefficients and other relevant characteristics.

Observations on retail sales were collected from both 1954 and 1963 censuses, but are fitted to the model separately so that the consistency of estimated coefficients over a period of time may be studied.

Our model starts with the simple traditional demand function that quantity demanded for retail goods or services (Q) is defined as a function of price (p) of that good or service. Population (N) and per capita income (Y) are conventionally considered to be predominant variables which shift the demand schedule in the quantity-price plane. Demand for retail goods and services is positively related to population and income, but change in demand is understood to be proportionally less than that of population and income because the demands are generally inelastic with respect to income and population.

Different population density per square miles (D) may have different effects on total quantity sold. It is conceivable that quantity sold per capita in a small SMSA could be greater than in a larger SMSA. Therefore, population density is also included in order to test its relationship to retail sales.

Since we are interested only in estimating the differential dollar amounts of sales and since the price elasticity of demand for retail trade in 188 SMSAs was estimated by Schwartzman to be -1.0,[6] our first functional relationship may be specified as follows:

$$\begin{aligned} Q &= aP^{-1}N^bY^cD^du \quad \text{or} \\ Q\cdot P &= aN^bY^cD^du \quad \text{or} \\ S &= aN^bY^cD^du. \end{aligned} \tag{2.1}$$

S is the amount of retail goods and services sold which, of course, equals the product of the quantity sold and the price at which the transaction was completed: a, b, c and d are unknown constants and u is the disturbance term whose natural log is assumed to be distributed normally $(0, \sigma^2)$ and independently of all explanatory variables throughout this article.

Total retail goods and services sold in an area are also likely to be affected by local government expenditures and tax rates levied on retail business and goods and services. Local government current expenditures, especially expenditures on police and fire protection per capita may indicate the extent to which businessmen and consumers are protected. A high level of protection and hence, greater safety and security, undoubtedly encourages business expansion and consumer expenditure. On the other hand, a high sales or excise tax passed on

[6] See [14, p. 226].

This content downloaded from
96.50.241.211 on Thu, 11 May 2023 02:50:52 +00:00
All use subject to https://about.jstor.org/terms

to consumers conceivably would have an adverse effect on consumption as long as the price elasticity of demand is not zero. A high rate of business income tax may raise the cost of business operation, and, hence, discourage business expansion and new entries.

These local government effects have for some time been discussed in the literature[7] but have rarely, if at all, been tested empirically. For instance, none of the previously cited studies has considered these fiscal variables. Needless to say, the empirical effects of these fiscal variables would be of practical importance in local government decision making.

To capture these effects, either local government current expenditures per capita (G) or per capita expenditures on police and fire protection (GPF) and per capita local government revenues not derived from property taxes (T) will be employed to test the fiscal effects of local government on retail trade. The explanatory variables G and GPF are used alternately to test whether expenditures by function have different effects compared with total expenditures. Furthermore, since public expenditures have been typically cast as an exogenous variable, even in the study of public finance, and discussed in a simple descriptive fashion[8] the alternative inclusion of G and GPF together with T in this article emphasizes the functional economic effects of both taxes and expenditures simultaneously, not those of taxes alone.

With the introduction of these variables, (2.1) is enlarged to the following:

$$S = aN^bY^cD^dG^eT^fu \tag{2.2a}$$

$$S = aN^bY^cD^dGPF^eT^fu. \tag{2.2b}$$

Besides the demand and public variables which are hypothesized to be of dominant influence on the differential retail sale, this model may not be complete without being concerned with supply and other factors. Despite the fact that the supply of retail goods and services has been hypothesized to be perfectly[9] or highly elastic and, hence, has little effect on the differential sales between SMSAs, a supply variable, the number of retail establishments per 1,000 population (B), is also included simply to test the hypothesis.

An educational variable, the ratio of population with some university or higher education to that with less than eight years schooling (E) and a dependency variable, the ratio of population under 18 and over 64 years of age to total population (U), are also included. The educational variable is used to reflect different tastes and preferences in consumption and the dependency variable is used to represent different levels of dependency and conservative attitude. The former is probably also a quality index of productivity, while the latter is likely a quantity index of labor supply. In other words, we may expect, in general, a positive relationship between the sales and the educational variable and an inverse relationship between the sales and the dependency variable.

A regional dummy variable (M) is also included to indicate regional differ-

[7] For a summary of the results of earlier works on tax effects on industrial location, see [4, 8, and 17] for some general comments. See Advisory Commission on Intergovernmental Relations [1] for some synthetical analysis on both effects of taxation and expenditures.

[8] With exceptions, public finance has been heavily preoccupied with the revenue side and has virtually neglected the significance of expenditures. See [17, p. 256].

[9] For instance, Schwartzman [14, p. 222] made this assumption explicitly in his article.

This content downloaded from
96.50.241.211 on Thu, 11 May 2023 02:50:52 +00:00
All use subject to https://about.jstor.org/terms

ence in retail sales. Finally, the complete model takes the following forms for total retail sales.

$$S = aN^bY^cD^dG^eT^fB^gE^hU^iM^ju \tag{2.3a}$$

$$S = aN^bY^cD^dGPF^{\bar{e}}T^fB^gE^hU^iM^ju. \tag{2.3b}$$

For per capita retail sales, the population variable is eliminated from the right-hand side of (2.3a) and (2.3b).[10] All variables are expressed in natural logarithms except the regional dummy variable which equals one for SMSAs in the West Coast and zero elsewhere. The constants, a, b, c, d, e, f, g, h, i, and j are estimated by linear regression techniques. All constants except a and j are, in fact, partial elasticities of retail sales with respect to the respective variable in question, by holding other variables constant. These partial elasticities are assumed to be constant in all SMSAs.

3. THE RESULTS AND IMPLICATIONS

Tables 1 and 2 contain regression results for both total and per capita retail sales of 1954 and 1963, respectively. In determining the differential retail sales among all large SMSAa, all three specifications of the model were tested.

Population (N), per capita income (Y) and population density (D) are three extremely dominant variables which jointly explain 98 percent of the variation in total retail sales. The latter two variables are of substantial importance in determining per capita retail sales; jointly they explain more than 50 percent of the variations in per capita sales.

The contribution of the public variables, local government current expenditures and non-property tax revenues per capita (G and T), to the explanatory power in total retail sales is insignificant. However, the introduction of these two variables improves considerably the determination of per capita retail sales. The improvement can readily be seen from the substantially increased values of R^2's, even though it is well understood that their contribution cannot merely be measured by the difference in R^2's with and without them, nor can we, in general, allocate the joint explanation between the demand and public variables.[11]

Similarly, the addition to the model of the supply variable—number of retail

---

[10] Data sources and symbols:

S—total retail sales (millions of dollars) or per capita retail sales ($1,000), 1954 and 1963 (1950 and 1960 population were used) *Census of Business*. 1954 and 1963.

N—population, 1950 and 1960, *Census of Population* 1950 and 1960.

Y—per capita income, 1950 and 1959, *Survey of Current Business*, August 1968.

D—population density, persons per square mile, 1950 and 1960, same source as N.

G—per capita local government current expenditures ($1,000) 1957 and 1962 (1950 and 1960 population were used), *Census of Government*, 1957 and 1962.

GPF—per capita local government expenditures on police and fire protection, ($), 1957 and 1962, same source as G.

T—per capita local government tax revenues not derived from property taxes ($1,000), 1957 and 1962, same source as G.

B—number of retail establishments per 1,000 people, 1954 and 1963 (1950 and 1960 population were used), same source as S.

E—ratio of population with some university or higher education to population with eight years or less of schooling, 1950 and 1960, same source as N.

U—ratio of population under 18 and over 64 years old to total population, 1950 and 1960, same source as N.

M— regional dummy variable, 1 for West Coast SMSAs and zero elsewhere.

[11] For detailed analysis, see [6, pp. 29–49].

This content downloaded from 96.50.241.211 on Thu, 11 May 2023 02:50:52 +00:00
All use subject to https://about.jstor.org/terms

Table 1. REGRESSION RESULTS OF TOTAL AND PER CAPITA RETAIL SALES, 1954

Dependent variable	Constant term	Coefficients or partial elasticity										Coefficient of determination	$\bar{R}^2$
		Population	Per capita income	Population density	Current total expenditures	Police & fire expenditures	Non-property tax	Establishment	Education ratio	Dependency ratio	Regional dummy		
	(a)	(b)	(c)	(d)	(e)	(ē)	(f)	(g)	(h)	(i)	(j)	(R^2)	
	-12.2	0.958[a]	0.889[a]	-0.073[a]								0.978	
	-10.6	0.951[a]	0.738[a]	-0.069[a]	0.232[a]		-0.012					0.985	
Total sales	- 9.1	0.994[a]	0.580[a]	-0.109[a]	0.285[a]		-0.019	0.111	0.080[b]	0.117	-0.178[b]	0.989	0.986
	- 9.8	0.939[a]	0.477[a]	-0.082[a]		0.354[a]	-0.021					0.991	
	- 9.1	0.983[a]	0.235	-0.112[a]		0.352[a]	-0.029[b]	0.087	0.085[a]	-0.046	-0.106[b]	0.994	0.992
	- 4.9		0.766[a]	-0.082[a]								0.547	
	- 3.3		0.609[a]	-0.080[a]	0.222[a]		-0.018					0.677	
Per capita sales	- 2.2		0.564[a]	-0.111[a]	-.283[a]		-0.020	0.116[b]	0.083[b]	0.112	-0.182[a]	0.778	0.705
	- 2.5		0.332[b]	-0.094[a]		0.333[a]	-0.028					0.783	
	- 2.1		0.188	-0.119[a]		0.347[a]	-0.032[b]	0.100[b]	0.093[a]	-0.055	-0.118[a]	0.867	0.823

[a] Significant at the .01 level.
[b] Significant at the .05 level.
NOTE: $\bar{R}^2$ is the coefficient of determination adjusted for the degrees of freedom.

This content downloaded from
96.50.241.211 on Thu, 11 May 2023 02:50:52 +00:00
All use subject to https://about.jstor.org/terms

Table 2. REGRESSION RESULTS OF TOTAL AND PER CAPITA RETAIL SALES, 1963

Dependent variable	Constant term (a)	Population (b)	Per capita income (c)	Population density (d)	Current total expenditures (e)	Police & fire expenditures ($\bar{e}$)	Non-property tax (f)	Establishment (g)	Education ratio (h)	Dependency ratio (i)	Regional dummy (j)	Coefficient of determination (R^2)	$\bar{R}^2$
		Coefficients or partial elasticity											
Total sales	- 12.0	0.928[a]	0.897[a]	-0.075[a]								0.982	
	- 10.1	0.928[a]	0.697[a]	-0.073[a]	0.159[b]		-0.026					0.985	
	- 8.0	1.001[a]	0.212	-0.080[b]	0.239[a]		-0.015	0.277	0.088[b]	0.007	-0.134[b]	0.989	0.986
	- 10.9	0.899[a]	0.742[a]	-0.077[a]		0.210[b]	-0.023					0.986	
	- 9.4	0.970[a]	0.350	-0.082		0.192[b]	-0.016	0.258	0.083[b]	0.018	-0.087	0.989	0.986
Per capita sales	- 4.2		0.669[a]	-0.096[a]								0.504	
	- 2.5		0.495[b]	-0.093[a]	0.135		-0.032[b]					0.581	
	- 1.1		0.223	-0.080[b]	0.238[a]		-0.015	0.273	0.087[b]	0.007	-0.134[b]	0.740	0.665
	- 3.3		0.527[b]	-0.099[a]		0.113	-0.031					0.574	
	- 2.1		0.260	-0.091[b]		0.172[b]	-0.016	0.299[b]	0.095[b]	0.026	-0.101[b]	0.721	0.629

[a] Significant at the .01 level.
[b] Significant at the .05 level.
NOTE: $\bar{R}^2$ is the coefficient of determination adjusted for the degrees of freedom.

This content downloaded from
96.50.241.211 on Thu, 11 May 2023 02:50:52 +00:00
All use subject to https://about.jstor.org/terms

establishments per 1,000 people (B), the educational variable—the ratio of the population with some university education or better to those with eight years or less education (E), the dependency variable—the percentage of population over 64 and under 18 years of age (U), and the regional dummy variable—West Coast SMSAs (M), had a very limited result in raising the power of joint explanation in total retail rates. Again, these four variables had considerable effect on the determination in per capita retail sales as had the governmental variables, even viewed strictly from the increased values of R^2's alone.

The model is extremely satisfactory in explaining the variation in total retail sales; jointly the three sets of variables have explained almost all the variations in total sales in 1954 and 1963. For per capita sales among all selected SMSAs, the explanatory variables in this model accounted for 78 percent and 74 percent, respectively, of the differential per capita sales. The lower values of R^2's in per capita sales as compared to those of total sales are in large part attributable to the simultaneous elimination of population influence on total sales and the deletion of population among explanatory variables. As far as the coefficients of determination are concerned, this model including a variable of per capita gross local government expenditures yields better results than that developed by Ferber,[12] in both total and per capita sales.

Substituting per capita local government police and fire expenditures for per capita current government expenditures gives relatively better results for 1954. For 1963, the substitution does not make much difference in terms of total explanatory power of all independent variables.

All coefficients estimated for population are highly significant even at the one percent confidence level. As is expected, the relationship between population and total retail sales is positive, and the partial elasticity of sales with respect to population is less than unity. The population partial elasticities are not only relatively constant over time, i.e., 0.98 for 1954's and 0.97 for 1963's sales, they are also relatively constant for all equations in this stepwise regression test, ranging from 0.90 to 1.00. Total sales change about 0.9 percent for every one percent change in population, holding other things constant. This nearly unitary elasticity may also be seen from the simple correlation coefficient matrix (not presented here). Population and total sales are almost perfectly correlated with simple correlation coefficients of 0.97 and 0.98 for the two years under consideration. In addition, the partial population elasticities do not seem to be affected by other explanatory variables used in this study.

Per capita income here also has shown a uniquely positive effect on total as well as per capita retail sales. Most of the estimated coefficients or the partial elasticities are statistically significant at the five percent level of confidence or higher, except in the equations for 1963 when the third set of explanatory variables are included. The estimated partial income elasticities have tended to decline when more explanatory variables are added to each equation being studied, and their values vary only slightly over time. This is probably due in part to the collinearity existing between per capita income and per capita gov-

[12] The R^2's produced by Ferber's model are 0.93 and 0.64, respectively, for total sales and per capita sales, adjusted for degrees of freedom. See Robert Ferber [5, Tables 2 and 4]. The R^2's adjusted for degrees of freedom in our model are 0.99 and 0.71 for 1954 and 0.99 and 0.67 for 1963.

This content downloaded from
96.50.241.211 on Thu, 11 May 2023 02:50:52 +00:00
All use subject to https://about.jstor.org/terms

ernment expenditures and the educational variables. For instance, the simple correlation coefficients between income and education are around 0.30; between income and police and fire expenditures, 0.55. The simple correlation coefficient between income and current local government expenditures is low in 1954's sales equation (0.25) but high in 1963's sales equation (0.56).

Collinearity, nevertheless, should be recognized as a matter of the degree to which they are linearly related, rather than a problem of all or nothing.[13] If we agree with Klein [9, p. 101] that "multicollinearity is not necessarily a problem unless it is high relative to the over-all degree of multiple correlation . . .," we may feel much more comfortable about our results. Comparing the simple correlation coefficients to R^2's in Tables 1 and 2, we may conclude that the multicollinearity problem in our study is not harmful and, hence, our estimated partial elasticities are sustainable and usable in predicting retail sides.

A significant, negative coefficient between population density and retail sales has been obtained for all except one equation shown in Table 2. This indicates that for a given number of total population, the higher the population density per square mile, the smaller the amount of per capita sales and, hence, the smaller the amount of total sales. This negative influence of population density on retail sales has been rather stable and small. Every one percent change in the former would only result in about 0.1 percent change in the latter.

Per capita local government current expenditures, to a certain extent, represent the level of public goods and services provided for public consumption. They are typically the services of police and fire protection, more or less complementary to the demand for and supply of retail goods and services. Assuming that the safety, security and convenience in retail trade increases as the local government expenditures increase, then per capita retail sales and, hence, total retail sales should be higher where per capita local government current expenditures are higher. This hypothesis is accepted at the five percent significance level. After population and per capita income, this expenditure variable appears to have the greatest influence on both total and per capita retail sales. The partial elasticities of retail sales with respect to the current expenditures are estimated at around 0.2 to 0.3. By substituting the expenditures on police and fire protection for the total current expenditures, the estimated partial elasticities are slightly higher and statistically more significant in determining 1954's sales but are slightly lower and relatively less significant in determining 1963's sales.

Another public variable, per capita local government revenues derived from non-property taxes, has been estimated with correct signs throughout all equations, i.e., it is negatively related to retail sales as hypothesized in the model, due to its effect of raising prices of retail goods and services. Nevertheless the estimated partial elasticities of this tax variable are rather small and statistically indifferent from zero. Stated differently, the effect of sales and other taxes on retail goods' and services' prices does not significantly lower the demand for and supply of retail goods and services. This finding is of practical importance in local government decision making with respect to tax rate changes and revenue collections.

---

[13] See [7, pp. 192–4] and [6, pp. 79–83].

This content downloaded from
96.50.241.211 on Thu, 11 May 2023 02:50:52 +00:00
All use subject to https://about.jstor.org/terms

Also the supply variable, the number of retail trade establishments per 1,000 people in each SMSA, has some positive effect on retail sales, as previously specified in the model. Holding constant other variables, the greater the number of retail establishments, the greater are both per capita and total sales. This is probably due in part to the varieties of commodities and relatively low price levels, possibly resulting from a larger number of establishments. Again, the estimated partial elasticities of this supply variable, like that of the tax variable, are not statistically different from zero at the five percent level in most equations. This implies that the positive effect of the number of retail establishments on sales may be insignificant. This conclusion is different from the findings of Ferber [5] and Tarpey and Bahl [16]. In Ferber's study, this supply variable is significant even at the one percent level. The principal reason for this difference may be attributed to different observations included in the two studies. All of Ferber's observations were in Illinois, many of them from small cities, while this article analyzes large SMSAs all over the country. The observations used here are, of course, relatively homogeneous. Another possible reason for disagreement may be the difference in model specification.

The educational variable, the ratio of population with some university education or better to those with eight years or less of schooling, has a positive effect on retail sales, and the estimated partial elasticities are all significantly different from zero and constant over all equations. They imply that for the areas with more highly educated population, the greater are per capita as well as total retail sales. For every one percent increase in this educational ratio, retail sales increase by about 0.09 percent, ceteris paribus.

Statistically, the regional dummy variable indicates the existence of differential retail sales solely due to the geographical distribution of the SMSAs studied. This difference may simply be attributable to weather, climate and other physical conditions. In this article, the SMSAs in the West Coast (including Seattle, Portland, Los Angeles, San Diego and San Francisco) were shown by the negative coefficient to have sold less on the average than other SMSAs outside the Coast. It should be recognized that the significant regional dummy variable indicates only that the intercept in our model varies from region to region.[14]

The so-called partial determination coefficient may be used to measure the marginal contribution of an explanatory variable to the explanation of retail sales, given all the other explanatory variables. Although the disadvantage of this measure is that the sum of these partial determination coefficients is not in general equal to the R^2 of the original regression, the rank of the partial determination coefficient provides an objective measure of the importance of each individual explanatory variable toward the overall explanation.[15]

The partial determination coefficients for all the explanatory variables are shown in Table 3. In terms of individual contributions to the explanation of differential total retail sales, population contributed most; population density and local government expenditure ranked next. The educational ratio, number of establishments, per capita income and non-property tax revenues are rela-

---

[14] For the use of dummy variables in regression equations, see [15].
[15] See [7, pp. 197–200].

This content downloaded from
96.50.241.211 on Thu, 11 May 2023 02:50:52 +00:00
All use subject to https://about.jstor.org/terms

Table 3. PARTIAL DETERMINATION COEFFICIENTS[a]

Explanatory variable	Total retail sales				Per capita sales			
	1954		1963		1954		1963	
Population	.97	.98	.96	.94	---	---	---	---
Per capita income	.19	.06	.02	.05	.23	.04	.03	.04
Population density	.38	.53	.11	.11	.47	.62	.13	.14
Total current expenditures	.36	---	.21	---	.37	---	.21	---
Police and fire expense	---	.63	---	.16	---	.62	---	.15
Non-properly tax revenue	.04	.13	.03	.03	.04	.16	.03	.04
No. of establishments	.10	.10	.09	.08	.13	.15	.10	.11
Education ratio	.13	.23	.15	.12	.18	.32	.17	.19
Dependency ratio	.03	.01	.01	.01	.03	.01	.01	.02
Regional dummy	.21	.16	.16	.08	.24	.21	.18	.12

[a] There are two sets of partial determination coefficients under each heading in this table because total current government expenditures and expenditures on police and fire protection are alternately used in each regression equation. For per capita sales, population is eliminated from the regression equations.

tively less important. It is also interesting to note that similar rank orders are carried out in the explanation of the per capita retail sales in which the effect of population size was eliminated.

Since the principal objective of this article is to explain retail sales in large SMSAs, the stability of those coefficients obtained for the retail sales between 1954 and 1963 is of vitally important significance. To test the stability of those estimated coefficients over time, a Chow [2] test of equality between sets of coefficients in two linear regressions was conducted. The residual sum of squares under the null hypothesis that these two sets of coefficients are the same and the sum of squares of the deviations between the two sets of estimates of retail sales under the alternative hypothesis were computed. The difference between these two sums of squares was therefore obtained. The ratio of the difference to the latter sum of squares, adjusted for their individual degrees of freedom, follows an F distribution if the null hypothesis is true. The ratios for total and per capita retail sales equations in this study were found to be much smaller than the critical F values at the one percent level, regardless whether total local government current expenditures or only police and fire expenditures per capita (G and GPF respectively) were included as one of the explanatory variables in the four main equations. Thus, our null hypothesis is accepted, and consequently we conclude that coefficients estimated for the retail sales have not changed during the period being studied.

Since the stability of the coefficients is supported by Chow's test, the predictive accuracy of using the 1954 equation to predict the 1963 data was studied. Actual and predicted 1963 sales are shown in Table 4. The explanatory variables G and GPF were included alternately in the total as well as per capita sales equation. Consequently, two sets of predicted values are shown in the table,

This content downloaded from
96.50.241.211 on Thu, 11 May 2023 02:50:52 +00:00
All use subject to https://about.jstor.org/terms

Table 4. ACTUAL AND PREDICTED SALES BY S.M.S.A., 1963[a]

S.M.S.A.	Total sales			Per capita sales		
	Actual	Predicted		Actual	Predicted	
Akron, Ohio	6.7007	6.8068	6.5518	.4575	.5583	.3164
Allentown, Beth, Easton, Pa.	6.4998	6.6423	6.3983	.3010	.4362	.2047
Atlanta, Ga.	7.3896	7.4463	7.2715	.4648	.5177	.3610
Baltimore, Md.	7.7258	7.9024	7.8860	.2716	.4463	.4457
Birmingham, Ala.	6.6451	6.6908	6.5942	.1917	.2304	.1459
Boston, Mass.	8.2873	8.3619	8.2956	.4281	.5001	.4510
Buffalo, N. Y.	7.4236	7.7405	7.5406	.2481	.5595	.3744
Chicago, Ill.	9.1992	9.3144	9.1251	.4635	.5789	.4116
Cincinnati, Ohio	7.4091	7.6172	7.4542	.4322	.6368	.4915
Cleveland, Ohio	7.9069	8.0683	7.8938	.4077	.5647	.4063
Columbus, Ohio	7.0432	7.1074	6.9913	.5153	.5749	.4741
Dallas, Texas	7.5005	7.6607	7.5622	.5125	.6709	.5935
Denver, Colorado	7.3350	7.6566	7.3944	.5005	.8199	.5806
Detroit, Michigan	8.5929	8.5557	8.4362	.3601	.3189	.2157
Grand Rapids, Mich.	6.4968	6.5553	6.3776	.6019	.6562	.4974
Houston, Texas	7.5817	7.7493	7.5983	.4563	.6238	.4988
Indianapolis, Inciana	7.2449	7.3266	7.1467	.6973	.7780	.6223
Kansas City, Mo. - Kansas	7.4283	7.5836	7.4525	.4818	.6345	.5229
Los Angeles, California	9.2102	9.3416	9.1854	.3940	.5224	.3837
Louisville, Kentucky	6.8845	7.0204	6.8272	.2981	.4276	.2468
Memphis, Tennessee	6.7968	6.8723	6.7678	.3558	.4262	.3393
Miami, Florida	7.3889	7.3911	7.3328	.5483	.5458	.5027
Milwaukee, Wisconsin	7.4425	7.7579	7.6348	.3572	.6678	.5591
Minneapolic, Minnesota	7.6935	7.9794	7.6807	.4657	.7495	.4733
New Orleans, Louisiana	7.0238	7.1114	6.9974	.2616	.3426	.2416
New York, New York	9.2102	9.7205	9.5181	.0673	.4383	.2485
Philadelphia, Pa.	8.6547	8.7734	8.6239	.2784	.3970	.2695
Phoenix, Arizona	7.0326	7.0748	6.9100	.5351	.5722	.4278
Pittsburgh, Pa.	7.9649	8.1734	7.9687	.1794	.3857	.2013
Portland, Oregon	7.1538	7.3059	7.2449	.4422	.5894	.5447
Rochester, New York	7.0370	7.1439	6.8368	.6630	.7602	.4620
St. Louis, Missouri	7.9540	8.1463	8.0146	.3235	.5145	.4043
San Diego, California	7.2499	7.5216	7.3945	.3097	.5766	.4672
San Francisco, California	8.4143	8.4903	8.3495	.4829	.5514	.4280
Seattle, Washington	7.4662	7.7117	7.5001	.4566	.6961	.4999
Toledo, Ohio	6.7310	6.7348	6.5544	.6065	.6029	.4344
Washington, D. C.	8.1218	8.1835	8.0140	.5199	.5778	.4233
Youngstown-Warren, Ohio	6.4297	6.6488	6.4233	.1973	.4063	.1888

[a] There are two sets of predicted sales under each heading in this table. The first set is the result of employing G and the second set, the result of employing GPF as one of the explanatory variables in our model. Figures are in natural logs; $ million for total sales and $1,000 for per capita sales.

both for total and per capita sales, the first set employing G and the second set employing GPF as one of the explanatory variables.

The actual and predicted sales of 1963 do not show any statistically significant difference at the one percent level in a pairwise Chi-square test. Regressing predicted sales on actual sales, the coefficients of determination are 0.98 for total sales and 0.44 for per capita sales. The coefficients obtained in these simple regressions are all highly statistically different from zero; they are 0.932 and 0.934 in total sales equations and 0.661 and 0.734 in per capita sales equation. The difference in coefficients and in coefficients of determination between total sales and per capita sales is because population enters the total sales equation with a coefficient significantly different from one, as previously pointed out.

Although the numbers of over- and under-predicted sales among the 38 SMSAs are about even when using GPF and other variables to predict and the difference between actual and predicted sales is insignificant, it is interesting

This content downloaded from
96.50.241.211 on Thu, 11 May 2023 02:50:52 +00:00
All use subject to https://about.jstor.org/terms

to note from Table 4 that sales in some SMSAs tend to be under-predicted, such as in Buffalo, Cincinnati, Denver, Milwaukee, New York, Portland, San Diego, St. Louis and Seattle, regardless whether total or per capita sales are predicted. On the contrary, sales in Allentown, Atlanta, Columbus, Detroit, Indianapolis, Miami, Rochester, Toledo and Washington tend to be over-predicted. When equations containing G are used to predict, the predicted sales are greater than actual sales in almost all SMSAs, although the differences are not at all statistically significant.

4. SUMMARY AND CONCLUSIONS

This article has attempted to explain differential retail sales among large SMSAs in the U. S. in 1954 and 1963. It has explored some important though hitherto unused variables in explaining retail sales variations, such as per capita local government current expenditures and non-property tax revenues. The three sets of explanatory variables employed in this paper consistently accounted for about 99 percent of the total variations in total retail sales and about 70 percent of the variations in per capita retail sales when the population factor was eliminated. The empirical results obtained in this article are as follows:

1. The high coefficients of determination obtained in this cross-section study are probably due to a greater homogeneity among the sample observed, i.e., large SMSAs and better model specifications. As far as the sign, size and the statistical significance of coefficients estimated here are concerned, the model is quite satisfactory in explaining the variations in retail sales among large U. S. urban areas.
2. The estimated partial elasticities do not seem to vary much over time, as is supported by a Chow test of equality between sets of coefficients in the two linear regressions. Using the 1954 equations to predict the 1963 data, the predictive accuracy of these equations is found to be extremely satisfactory.
3. The responsiveness of total retail sales to changes in per capita income, local government expenditures and number of establishments is estimated to be relatively less than that to population changes. However, they are much higher than the responsiveness to other changes such as non-property tax, education and dependency ratios. Similarly, this pattern of differential responsiveness is observed for per capita sales.
4. In terms of the marginal contribution to explaining variations in retail sales, population, local government expenditures and population density are of greatest importance while education, number of establishments and income variables are next.
5. The coefficient of the regional dummy variable reveals the existence of geographical differences in terms of regression intercept, in both total and per capita retail sales, even if the parameters of demand, government, supply and other variables are held constant over regions.

REFERENCES

[1] Advisory Commission on Intergovernmental Relations, *State-Local Taxation and Industrial Location*, A Commission Report, A-30, April 1967.

This content downloaded from
96.50.241.211 on Thu, 11 May 2023 02:50:52 +00:00
All use subject to https://about.jstor.org/terms

[2] Chow, G. C., "Test of Equality Between Sets of Coefficients of Two Linear Regressions," *Econometrica*, 28 (July 1960), 592–3.

[3] Dorfman, T. "The Function of the City," A. H. Pascal (Ed.), *Contribution to the Analysis of Urban Problems*, The Rand Corporation, 1968, 35–38.

[4] Due, J. "Studies of State-Local Tax Influences on Location of Industry," *National Tax Journal* 14 (June 1961), 163–73.

[5] Ferber, R. "Variations in Retail Sales Between Cities," *The Journal of Marketing*, 22 (January 1958), 295–303.

[6] Goldberger, A. S., *Topics in Regression Analysis*, New York: The MacMillan Company, 1968.

[7] Goldberger, A. S., *Econometric Theory*, New York: John Wiley and Sons, Inc., 1964.

[8] Isard, W. et. al, *Methods of Regional Analysis: An Introduction to Regional Service*, New York: John Wiley & Sons, Inc., 1960.

[9] Klein, L. R., *An Introduction to Econometrics*, New York: Prentice-Hall, 1962, 101.

[10] Leven, C. L., *Theory & Method of Income and Product Account for Metropolitan Areas*, Center for Regional Economic Studies, University of Pittsburgh, Pittsburgh, Pa., 1963.

[11] Liu, B. C., "The Relationship Among Population, Income and Retail Sales in SMSAs, 1952–1966," *The Quarterly Review of Economics and Business*, 10 (Spring 1970), 25–40.

[12] Reilly, W. J., *The Law of Retail Gravitation*, New York: 1931.

[13] Russell, V. K., "The Relationship Between Income and Retail Sales in Local Areas," *The Journal of Marketing*, 21 (January 1957), 329–32.

[14] Schwartzman, D., "The Growth of Sales Per Man-Hour in Retail Trades, 1929–1963," in Victor Fuchs, Ed., *Production and Productivity in the Service Industries*, N.B.E.R., Inc., 1969, 201–29.

[15] Suits, D. B., "Use of Dummy Variables in Regression Equations," *Journal of the American Statistical Association*, 52 (December 1957), 548–51.

[16] Tarpey, L. X. and Roy W. Bahl, "Inter-City Variations in Retail Sales: Some Hypotheses Revisited," *Southern Journal of Business*, 3 (October 1968), 1–10.

[17] Thompson, W. R., *A Preface to Urban Economics*, Baltimore: Johns Hopkins Press, 1968.

This content downloaded from
96.50.241.211 on Thu, 11 May 2023 02:50:52 +00:00
All use subject to https://about.jstor.org/terms