# Estimating Sales for Retail Centers: An Application of the Poisson Gravity Model

*A. Ason Okoruwa**
*Hugh O. Nourse***
*Joseph V. Terza****

*Abstract.* **The projection of total retail sales for a shopping center development is of critical importance in its valuation, in the making of investment decisions by investors, and to the retail merchants who must make location decisions. In this study, we apply the Poisson Gravity Model to forecast the number of shopping trips attracted to each of the major retail centers in the Atlanta metropolitan area. In the second stage, the estimated total retail sales for all the shopping centers covered in the study, are allocated to the individual centers, based on their estimated shopping trip shares.**

## Introduction

For any shopping center development, accurate sales projections are critically important. It is a common practice with large shopping centers to utilize a straight percentage lease or a guaranteed minimum rent combined with a percentage of sales volume above a threshold. Furthermore, rental income is the most important input in making cash flow projections to determine the financial feasibility and/or value of a shopping center.

In estimating sales for retail centers, gravity-type models are commonly used. The principle underlying retail gravity models is that the attractiveness of a retail center is directly related to its size and inversely related to the distance shoppers must travel to patronize it. The gravity concept was first applied to retail marketing by Reilly (1931). Subsequently, Huff (1968) reformulated Reilly's gravity model in terms of probability. It is this variation that is commonly applied in retail market potential analysis. Some of the studies using gravity models include LaLonde (1962), Dent (1978), Ellwood (1954), Lakshmanan and Hansen (1965), Pankhurst and Roe (1978), and Turner and Cole (1980).

The major criticism with the traditional gravity model is that it uses only size and distance to estimate sales potential of a retail center. To improve on the accuracy of the basic gravity model, many researchers included additional variables. The studies by Stanley and Sewall (1976) and McDougall (1978) represent major improvements in the estimation of the gravity model. Stanley and Sewall (1976) included store image variables such as quality, cleanliness, location, prices, friendliness, variety, and corpo-

*Finance Department, College of Business Administration, University of Northern Iowa, Cedar Falls, Iowa 50614.
**Department of Insurance, Real Estate and Legal Studies, College of Business Administration, University of Georgia, Athens, Georgia 30602.
***Department of Economics, Pennsylvania State University, University Park, Pennsylvania 16802.
Date Revised—July 1993; Accepted—August 1993.

rate image as additional variables. McDougall (1978) estimated a weighted gravity model. To accomplish this, he adjusted the size variable of a retail center depending on its rating on the following attributes: quality of neighborhood, layout of center, accessibility, appearance, whether center is enclosed or opened, parking adequacy, and parking cost.

Another approach to sales forecasting begins with the estimation of shopping trips taken by individuals or households to retail centers. Studies by Adler and Ben-Akiva (1976), Cleveland and Mueller (1961), Domencich and McFadden (1995) and Lakshmanan and Hansen (1965) have shown that the number of shopping trips taken to a retail location is highly correlated with sales generated at the location.

In this study, we provide additional improvements to the gravity model. We include the economic and demographic attributes of the shoppers in addition to size, distance, and retail center-specific characteristics. Our model is able to account for differences in shoppers' retail center preferences that arise due to unobservable and/or unmeasurable retail center attributes, given individuals' demographic and economic characteristics. We used the enhanced gravity model with additional variables to estimate the number of shopping trips taken to the different retail centers covered in the study. In a subsequent step, we estimated retail sales for the shopping centers based on their previously estimated shopping trip shares.

We specify shopping behavior as a Poisson process. The Poisson regression provides us the estimates of the number of shopping trips taken to the different centers included in the study, during a one-month period. Based on retail centers' estimated number of shopping trips, we then estimate their annual sales volume. This specification follows the example by Flowerdew and Aitkin (1982) in their study of migration flows between pairs of Standard Metropolitan Labor Areas (SMLAs) in Great Britain.

Some other researchers who have applied Poisson regression analysis in different contexts include Davis (1952), analysis of failure data; Gart (1964), virology; Hausman et al. (1984), research and development relationship; Kumar and Shih (1978), murder rate; Paull (1978), consumer purchase panel data; and Weber (1971), accident rate potential. The development and application of Poisson regression analysis to estimating shopping trips can be found in the studies by Okoruwa, Terza and Nourse (1988), Okoruwa (1985), Terza (1985), and Terza and Okoruwa (1985). In this present paper, we show how estimates of shopping trip shares for the alternative major retail centers, estimated using the Poisson Gravity Model (PGM), could be translated into their sales forecast.

In section two, the development of the PGM is presented. The data used for estimating the model is discussed in section three. Section four is devoted to estimating the model and allocating shopping trips to the retail centers. In the fifth section, the total volume of general merchandise sales in the SMSA is estimated and then allocated among the retail centers included in the study, based on their shares of shopping trips. In section six, we present conclusions.

## The Poisson Gravity Model (PGM)

The Poisson distribution is useful in describing the random occurrence of a given discrete event in a specified time interval, distance, volume, or space. A Poisson random variable assumes nonnegative integer values from zero to infinity. Some examples of Poisson-distributed events are the number of telephone calls to a switchboard, the number of airplanes arriving at an airport, and the number of customers purchasing a given product in a given time period. The main assumptions of the distribution are (1) the events are independent, (2) the probability of the occurrence of each event per unit of measurement is small, therefore, the probability of two or more occurrences approaches zero, and (3) the probability of the occurrence of an event per unit of measurement, is approximately proportional to the size of the unit against which it is measured. Shopping trip behavior satisfies the first assumption. The number of shopping trips made by an individual in a given time period (one-month interval) are assumed to be independent given (1) that shopping goods are the category of goods under consideration, and (2) the one-month interval is long enough so that the number of shopping trips undertaken in one interval does not directly affect those made in a future interval. Ehrenberg (1959) noted that for the purchases made by an individual in successive periods to be independent, two conditions are necessary, namely: (1) the successive periods of time must be of equal length and (2) the periods are long enough so that purchases made in one period do not directly affect those made in the next. Clearly, shopping behavior is consistent with the second and third assumptions.

Flowerdew and Aitken (1982), in applying the gravity concept in a migration study, characterized migration flows between pairs of Standard Metropolitan Labor Areas (Great Britain) as a Poisson process. Following their basic approach, we included socioeconomic characteristics of individual shoppers and retail center characterisics, in addition to the traditional retail center size and distance variables, in our PGM development. (See Okoruwa, Terza and Nourse 1988, for a more detailed derivation.)

Individual respondents are denoted by the subscript $t$, where $t=1, \ldots, T$. The alternative shopping destinations are denoted by the subscript $j$, and $j=1, \ldots, J$. And $Y_{jt}$ is the Poisson random variable representing the number of shopping trips made to the $j^{th}$ center by the $t^{th}$ individual. The likelihood that $Y_{jt}=y_{jt}$ will be the observed value of the dependent variable for the $t^{th}$ member of the sample and the $j^{th}$ center can be represented by the Poisson distribution expressed as

$$Pr\{Y_{jt}=y_{jt}\}=f_{jt}(y_{jt})=e^{-m_{jt}}(m_{jt})^{y_{jt}}/y_{jt}! \,, \tag{1}$$

and $m_{jt}$ is the Poisson parameter we want to estimate. We assume that the mean monthly patronization rate, $m_{jt}$, is a function of shopping center-specific variables, $X_{2j}$, for the $j^{th}$ center and a vector of socioeconomic and demographic characteristics of the $t^{th}$ individual, $x_{1t}$, and written as

$$m_{jt}=\exp\{X_t B_j\} \,, \tag{2}$$

where

$X_t=[X_{1t} \mid X_{2j}]$,
$X_{1t}=[1, x_{Djt} \mid X_t^*]$,
$x_{2j}$ = a row vector of center-specific characteristics of the $j^{th}$ center,
$x_{Djt}$ = the driving time, in minutes, from the $t^{th}$ individual's residence to the $j^{th}$ shopping center,

$X_t^*$ = a row vector of socioeconomic and demographic characteristics of the $t^{th}$ individual,

$B_j' = [B_{1j} \mid B_M]$,

$B_{1j}'$ = a vector of coefficients indicating how individual characteristics affect trips to the $j^{th}$ center,

$B_M'$ = a vector of coefficients of the center-specific variables.

For a sample of size $T$, the joint Poisson process $Y_t = [Y_{1t}, \ldots, Y_{Jt}]$, assuming independently distributed, its log-likelihood function can be represented as

$$L(B) = \sum_{t=}^{T} \sum_{j=1}^{J} \{y_{jt}\log(m_{jt} - m_{jt}) - \log(y_{jt}!)\} . \quad (3)$$

The dependent variable, the number of shopping trips taken in a one-month interval, gathered from the survey data used for estimating the parameters of the model is censored. Specifically, respondents were asked to indicate the number of shopping trips they took to the alternative major shopping centers in the past one-month period. The only possible responses were zero, one, two, or three or more trips. To eliminate possible estimation bias, censoring must be corrected for (Terza, 1985). Correcting for censoring, the log-likelihood function of the observed outcomes for a sample of size $T$ can then be written as

$$L(B) = \sum_{t=1}^{T} \sum_{j=1}^{J} \{d_{jt}\log(f_{jt}(y_{jt})) + (1 - d_{jt})\log(F_{jt})] , \quad (4)$$

where

$$f_{jt}(y_{jt}) = e^{-mjt(mjt)^{yjt}}/y_{jt}!,$$

$$F_{jt} = 1 - \sum_{i=0}^{2} f_{jt}(i).$$

$d_{jt} = 1$ if $y_{jt} < 3$, 0 otherwise.

The likelihood of the observed outcome for the $t^{th}$ observation of the sample is $f_{jt(yjt)}$ if $d_{jt} = 1$ and $F_{jt}$ otherwise. Given that structural variation in the regression coefficients across alternative shopping destinations is permitted and the censoring problem discussed, the maximum likelihood equation is estimated using the Partitioned Newton-Raphson Algorithm (PNRA) developed by Terza and Okoruwa (1985). Because of the large number of shopping centers and individual shoppers and the need to estimate the model simultaneously, the PNRA essentially partitions the resulting large matrix to make inversion manageable and improve computational efficiency.

## Data for Estimating Shopping Trip Shares

For this part of the study, three categories of data were used: (1) socioeconomic and demographic characteristics of shoppers, (2) shopping center-specific variables, and (3)

**Exhibit 1**
**Retail Centers and Definitions of Variables**

| | |
|---|---|
| | Retail Centers |
| BELV | Belvedere Plaza |
| BROA | Broadview Plaza |
| COBB | Cobb Center |
| COLU | Columbia Mall |
| CUMB | Cumberland Mall |
| GREE | Greenbriar Mall |
| LENO | Lenox Square |
| NDEK | North Dekalb Mall |
| NLAK | Northlake Mall |
| PERI | Perimeter Center |
| PHIP | Phipps Plaza |
| SDEK | South Dekalb Mall |
| SLAK | Southlake Mall |
| STEW | Stewart-Lakewood Center |
| WEND | West End Mall |
| | Variables Included in $X_{jt}$ |
| *MART* | 1 if respondent is married, 0 otherwise. |
| *AGE* | Respondent age in years. |
| *SEX* | 1 if male, 0 female. |
| *SCHC* | Last grade of school completed by the respondent. |
| *RACE* | 1 if respondent is White, 0 otherwise. |
| *EMPL* | 1 if respondent is employed, 0 otherwise. |
| *LADD* | Length of time residing at current address. |
| *ROWN* | 1 if residence owned, 0 otherwise. |
| *NPER* | Total number of persons in the household. |
| *TINC* | Total annual household income. |
| *DTIM* | Driving time from the residence of a shopper to a mall location. |
| | Variables Included in $X_{2j}$ |
| *MAGE* | Age of the mall in years. |
| *COVER* | 1 if the mall is totally enclosed, 0 otherwise. |
| *MSIZE* | Size of center in gross leaseable square footage. |
| *NDEPT* | No. of department stores in the mall. |
| *CONT* | Constant term. |

impedance or friction factor. These variables are similar to the variables typically found in retail sales estimation studies. Exhibit 1 shows the retail centers covered by the study and the definitions of the included variables, and Exhibit 2 presents descriptive statistics of the data. The data on socioeconomic and demographic characteristics and on the shopping center choices of households in the Atlanta Metropolitan Statistical Area (MSA) were obtained from a survey conducted in 1977 for the Atlanta Journal and Constitution Newspapers by Market Opinion Research, Inc. The number of usable observations was 828 out of the original sample size of 1572 after the observations with incomplete variables were excluded. As part of the survey data, the number of shopping trips, $Y_{jt}$, made by the $t^{th}$ individual to the $j^{th}$ alternative shopping destination in the metropolitan Atlanta retailing system, during a thirty-day interval, were recorded. The shopping centers covered in the survey included only class

**Exhibit 2**
**Descriptive Statistics of Data**

| | Retail Centers and Driving Time | | |
|---|---|---|---|
| Variable | Mean | Minimum | Maximum |
| *MAGE* | 11.30 | 1.00 | 23.00 |
| *MSIZE* | 69,5147.34 | 50,000.00 | 1,415,000.00 |
| *NDEPT* | 1.70 | 1.00 | 4.00 |
| *DTIM* | 25.82 | 2.00 | 88.00 |

| | Socioeconomic and Demographic | | |
|---|---|---|---|
| Variable | Mean | Minimum | Maximum |
| *AGE* | 43.61 | 21.00 | 70.00 |
| *SCHC* | 13.42 | 10.00 | 18.00 |
| *LADD* | 8.94 | .50 | 15.00 |
| *NPER* | 3.11 | 1.00 | 9.00 |
| *TINC* | 15,804.95 | 2,500.00 | 35,000.00 |

| | Variable | Frequency | Percent |
|---|---|---|---|
| *MART* | Married | 563 | 68.00 |
| | Not Married | 265 | 32.00 |
| *SEX* | Males | 370 | 44.69 |
| | Females | 458 | 55.31 |
| *RACE* | Whites | 571 | 68.96 |
| | Others | 257 | 31.04 |
| *EMPL* | Employed | 423 | 51.09 |
| | Unemployed | 405 | 48.91 |
| *ROWN* | Owners | 585 | 70.65 |
| | Renters | 243 | 29.35 |

1 and class 2 centers. Class 1 centers are defined as regional centers with one or more department stores and class 2 centers are centers with retail space greater than 100,000 square feet. The dependent variable—the number of shopping trips taken during a one-month period, is censored due to the design of the survey instrument. Each respondent was asked the question: "How many trips have you made to the $j^{th}$ shopping center in the last month?" Again, the only possible responses were: zero, one, two, or three or more trips. The summary of the shopping trips to the retail centers covered in the study is presented in Exhibit 3.

Driving time variable is used as the proxy for the friction or impedance factor between residential and shopping zones. The *1982 Atlanta Area Transportation Study* conducted by the Bureau of Plan Development, Georgia Department of Transportation, provided the driving times between residential and shopping zones. The driving times for the Atlanta MSA counties not covered by the transportation study were estimated by applying the estimated average speed of travel (45 m.p.h.) in the rural counties to measured straight-line distances.[1] For example, studies by Clapp (1980) and Daniels (1980) have shown that straight-line distances are good proxies for distances along roads. Exhibit 2 gives a summary of the driving time variables.

## Exhibit 3
## Shopping Trips Summary

| Retail Center | Trips | Frequency | Percent |
|---|---|---|---|
| Belvedere Plaza | 0 | 768 | 92.754 |
| | 1 | 18 | 2.174 |
| | 2 | 6 | 0.725 |
| | ≥3 | 36 | 4.348 |
| Broadview Plaza | 0 | 777 | 93.841 |
| | 1 | 25 | 3.019 |
| | 2 | 9 | 1.087 |
| | ≥3 | 17 | 2.053 |
| Cobb Center | 0 | 767 | 92.633 |
| | 1 | 26 | 3.140 |
| | 2 | 14 | 1.691 |
| | ≥3 | 21 | 2.536 |
| Columbia Mall | 0 | 749 | 90.459 |
| | 1 | 26 | 3.140 |
| | 2 | 18 | 2.174 |
| | ≥3 | 35 | 4.227 |
| Cumberland Mall | 0 | 684 | 82.609 |
| | 1 | 60 | 7.246 |
| | 2 | 33 | 3.986 |
| | ≥3 | 51 | 6.159 |
| Greenbriar Mall | 0 | 731 | 88.285 |
| | 1 | 40 | 4.831 |
| | 2 | 25 | 3.019 |
| | ≥3 | 32 | 3.865 |
| Lenox Square | 0 | 682 | 82.367 |
| | 1 | 47 | 5.676 |
| | 2 | 34 | 4.106 |
| | ≥3 | 65 | 7.850 |
| North Dekalb Mall | 0 | 742 | 89.614 |
| | 1 | 29 | 3.502 |
| | 2 | 19 | 2.295 |
| | ≥3 | 38 | 4.589 |
| Northlake Mall | 0 | 683 | 82.488 |
| | 1 | 54 | 6.522 |
| | 2 | 41 | 4.952 |
| | ≥3 | 50 | 6.039 |
| Perimeter Center | 0 | 738 | 89.130 |
| | 1 | 42 | 5.072 |
| | 2 | 25 | 3.019 |
| | ≥3 | 23 | 2.778 |
| Phipps Plaza | 0 | 770 | 92.995 |
| | 1 | 33 | 3.986 |
| | 2 | 15 | 1.812 |
| | ≥3 | 10 | 1.208 |
| South Dekalb Mall | 0 | 749 | 90.459 |
| | 1 | 29 | 3.502 |
| | 2 | 10 | 1.208 |
| | ≥3 | 40 | 4.831 |
| Southlake Mall | 0 | 744 | 89.855 |
| | 1 | 26 | 3.140 |
| | 2 | 24 | 2.899 |
| | ≥3 | 34 | 4.106 |
| Stewart-Lakewood Center | 0 | 759 | 91.667 |
| | 1 | 20 | 2.415 |
| | 2 | 17 | 2.053 |
| | ≥3 | 32 | 3.865 |
| West End Mall | 0 | 744 | 89.855 |
| | 1 | 25 | 3.019 |
| | 2 | 19 | 2.295 |
| | ≥3 | 40 | 4.831 |

Finally, the data on the physical characteristics of the shopping centers covered in the study were gathered from Atlanta Journal and Constitution Newspapers publications (1977, 1970). Exhibit 2 shows the summary of the retail centers data.

## Estimation of the Poisson Gravity Model

The maximum likelihood estimation of the Poisson Gravity Model (corrected for censoring), equation (4), was performed using the Partitioned Newton-Raphson Algorithm (PNRA) (1985). The number of observations used for estimating the model was 600, leaving out a holdout sample of 228 observations. The estimation results can be found in Okoruwa et al. (1988).

### *Prediction of Shopping Trip Shares*

To predict the number of shopping trips made to the fifteen alternative shopping centers, using the calibrated model, the holdout sample of 228 observations was used. The predicted trip shares are then compared to the observed trip shares. The proportion of trips allocated to the $j^{th}$ center was based on the ratio of the total number of trips made to the $j^{th}$ center by all members of the sample to the total number of trips made to all the $J$ centers in the retailing system. This allocation procedure was based on a rather strong assumption that all shopping trips are made to only the centers covered in the study. The mean monthly number of trips to each shopping center is predicted using equation (2) and the estimation results. The total number of mean monthly trips by the $t^{th}$ individual to all the $J$ centers is determined by summing the number of mean trips to each center. Then, the grand total of the number of mean trips made by all the individuals to all the retail centers is determined by summing the number of mean trips made by each individual to the $J$ centers. In the final step, we allocated trips among the centers based on the proportion of the $j^{th}$ center predicted trips to the total predicted trips for all $J$ centers. Exhibit 4 shows the shopping trip share predictions, obtained using the holdout sample of 228 observations along with the actual trip shares observed in the holdout sample. To the test the null hypothesis that the difference between the observed and expected trip shares is too large to be attributable to chance alone, a *chi*-square statistic was computed. At a 1% level of significance and 14 degrees of freedom, the critical *chi*-square value is 29.17. Therefore, the null hypothesis was rejected with 99% level of confidence. The estimation results indicate that reasonably accurate aggregated retail center share predictions can be obtained by applying the Poisson Gravity Model to micro-level data. In the next section, we show how shopping trip shares could be used in estimating sales for retail centers.

## Total General Merchandise Sales (GMS) for the MSA

In this section, we first estimate the total shopping goods sales for the Atlanta MSA using equation (5), shown on the next page. In the second step, the estimated MSA

**Exhibit 4**
**Shopping Trip Shares of the Retail Centers**

| Retail Center | Observed % | Predicted % |
|---|---|---|
| Belvedere Plaza | .0604 | .0442 |
| Broadview Plaza | .0336 | .0475 |
| Cobb Center | .0362 | .0434 |
| Columbia Mall | .0738 | .0534 |
| Cumberland Mall | .1034 | .0986 |
| Greenbriar Mall | .0805 | .0785 |
| Lenox Square | .1262 | .1221 |
| North Dekalb Mall | .0631 | .0669 |
| Northlake Mall | .1007 | .1069 |
| Perimeter Center | .0631 | .0550 |
| Phipps Plaza | .0389 | .0310 |
| South Dekalb Mall | .0537 | .0607 |
| Southlake Mall | .0376 | .0478 |
| Stewart-Lakewood Center | .0617 | .0681 |
| West End Mall | .0671 | .0757 |

sales volume is allocated among the major shopping centers based on their shares of shopping trips estimated in the last section. As stated before, previous studies by Cleveland and Mueller (1961), Lakshmanan and Hansen (1965), Domencich and McFadden (1965), and Adler and Ben-Akiva (1976) have shown that shopping trips taken are highly correlated with sales generated in retail locations. Analysis of retail expenditures potential at the MSA level is justified by the following reasons: (1) MSAs are generally economically integrated geographic areas, (2) a proportionally high level of consumption expenditures are confined to retail establishments in an MSA, (3) there is very little level of cross-shopping between MSAs retailing systems, (4) advertisements using television, radio and newspaper are local in orientation and do not typically extend beyond the confines of an MSA, (5) secondary data at the MSA geographic unit of measurement are readily available. Many researchers have conducted studies estimating aggregate shopping goods sales at the MSA level. Studies by Ingene and Lusch (1980) and Ingene and Yu (1981) reported very high values of $R$-squared.

The study by Ingene and Yu (1981) is particularly noteworthy for its theoretical development of the behavior of consumers and producers in spatially large markets. In this present research, we follow their approach in specifying aggregate retail sales by using effective buying income per capita, population per square mile, unemployment rate, average household size, and total population. Exhibit 5 shows the definitions of these variables.

The model for estimating the total general merchandise sales is specified as

$$GMS = \beta_0 + \beta_1 EBINC + \beta_2 UNEM + \beta_3 HSIZE + \beta_4 PDEN + \beta_5 POP + e \,. \qquad (5)$$

Data on general merchandise sales total (*GMS*), effective buying income per capita (*EBINC*), average size of household (*HSIZE*), population per square mile (*PDEN*), and population of metropolitan statistical area (*POP*) were gathered from *Sales & Marketing Management Magazine* (1987). The unemployment rate variable (*UNEM*)

**Exhibit 5**
**Metropolitan Statistical Area Variables**

| | |
|---|---|
| *GMS* | General merchandise sales total |
| *EBINC* | Effective buying income per capita |
| *UNEM* | Rate of unemployment |
| *HSIZE* | Average size of households |
| *PDEN* | Population per square mile |
| *POP* | Population total for the MSA |

**Exhibit 6**
**Estimation Results**

| Variable | Coefficient Estimate | *t*-Statistics |
|---|---|---|
| CONSTANT | −522994171.00 | −3.47 |
| *EBINC* | 33031.24 | 6.92 |
| *HSIZE* | 76839260.97 | 1.78 |
| *UNEM* | −150469539.00 | −.47 |
| *PDEN* | −135216.00 | −6.79 |
| *POP* | 757.01 | 66.62 |

data was collected from U.S. Department of Commerce, Bureau of the Census (1982).

Ordinary least squares regression was used to estimate equation (5), using a sample of 300 MSAs. The adjusted-$R$-squared for the model is 97%. All the variables except unemployment rate (*UNEM*) are statistically significant at the 5% level and have the expected signs. Estimation results are given in Exhibit 6.

Using the calibrated model shown in Exhibit 6, we predicted the 1978 general merchandise sales total for the Atlanta Metropolitan Statistical Area to be $1,289,075,904.

### *Allocation of GMS among the Major Shopping Centers in the MSA*

To estimate the amount of retail sales generated at the fifteen shopping centers covered in the study, we utilize the predicted total merchandise sales for the MSA estimated in the last section. The total predicted *GMS* is allocated among the retail centers based on their shopping trips shares estimated previously. We corrected for the fact that shopping trips were made to other retail centers in the MSA by multiplying the *GMS* by the proportion (approximately 56.42%) of total square footage of retail space represented by the shopping centers covered in the study.[2] Additionally, the effects of macro-marketing variables on sales volume were not considered. Essentially, demand for general merchandise goods is assumed fixed. Exhibit 7 shows the estimated retail sales for the shopping centers.

One way to test the predictive reliability of the PGM is to compare the model's predicted sales to the actual sales of the shopping centers. We were unable to obtain actual sales volume data from the managements of the retail centers. A proxy for

**Exhibit 7**
**Prediction of Shopping Center Sales**

| Shopping Center | Proxy for Actual Sales: Gross Leaseable Area × Median Sales Per Sq. Ft. | PGM Estimated Sales | Sales Residuals |
|---|---|---|---|
| | $ | $ | $ |
| Belvedere | 29,480,400 | 32,146,511 | −2,666,111 |
| Broadview Plaza | 29,024,400 | 34,546,589 | −5,522,189 |
| Cobb Center | 32,756,000 | 31,564,674 | 1,191,326 |
| Columbia Mall | 32,756,000 | 38,837,640 | −6,081,640 |
| Cumberland Mall | 92,945,150 | 71,711,447 | 21,233,703 |
| Greenbriar Mall | 53,064,720 | 57,092,785 | −4,028,065 |
| Lenox Square | 98,268,000 | 88,802,918 | 9,465,082 |
| North Dekalb Mall | 36,850,500 | 48,656,144 | −11,805,644 |
| Northlake Mall | 80,661,650 | 77,748,009 | 2,913,641 |
| Perimeter Mall | 61,826,950 | 40,001,314 | 21,825,636 |
| Phipps Plaza | 35,622,150 | 22,546,195 | 13,075,955 |
| South Dekalb Mall | 61,417,500 | 44,146,905 | 17,270,595 |
| Southlake Mall | 90,079,000 | 34,764,778 | 55,314,222 |
| Stewart-Lakewood Center | 18,050,000 | 49,528,900 | −31,478,900 |
| West End Mall | 28,415,830 | 55,056,354 | −26,640,524 |

the actual sales for a retail center was derived using the average median sales per square foot for a sample of regional shopping centers in the Southeast.[3] The last column of Exhibit 7 shows the residuals or the differences between "actual" and predicted sales. In analyzing the residuals, one must keep in mind that the 'actual' sales are in fact naive estimated sales based only on average median sale per square foot for regional shopping centers in the southeastern United States and the square footage of a retail center. Therefore, we were unable to make definite statements about the residuals. In addition, Nourse (1986) reported that one very successful fast food corporation indicated that after seventeen years of refining forecasts and doing postmortems, the level of accuracy achieved, half the time, was only in the order of plus or minus 20% or 25%.

## Summary and Conclusion

In this present study, we developed a technique to estimate sales for the major shopping centers in the Atlanta Metropolitan Statistical Area. First, the Poisson Gravity Model was used to estimate the market shares for the shopping centers. It was determined from a test performed using a holdout sample that it accurately determined the shopping trip shares of the centers covered in the study. In the second stage, multiple regression analysis was employed to estimate the total general merchandise sales for the whole MSA. The model had a very high adjusted $R$-squared value (97%). Finally, the total GMS is allocated among the shopping centers based on their shopping trips shares.

We have demonstrated how sales for retail centers could be estimated without the use of the traditional gravity model. Our technique utilized micro-level socioeconomic and demographic characteristics of shoppers, retail center-specific variables to model individual shopping behavior and to estimate the level of retail sales in major retail centers in a metropolitan statistical area. We hope that in future, further research on the reliability of the model for allocating sales among retail centers would be performed with actual retail sales data.

## Notes

[1]In the transportation study, *Connecticut Interregional Planning Program: Technical Summary No. 3* (Hartford, CT: Connecticut Highway Department 1963), it was estimated that speeds on rural expressways and arterial roads were 55 m.p.h. and 35 m.p.h., respectively. The average speed for the two road types, 45 m.p.h., is used in determining the driving time from the centroid of a zipcode area in the counties surrounding the Atlanta Regional Commission counties to the nearest traffic zone covered in the Georgia Department of Transportation study.

[2]In the traditional gravity model, shopping center size and distance are the basic variables used in estimating retail center sales. Constructed indexes by the authors to take into consideration the missing components in the model and the sales of retail centers not covered in the study produced only degraded results.

[3]The median sales per square foot was determined by taking the average of the median sales per square foot for department stores and for other mall tenants in regional shopping centers in the southeastern United States. Data was gathered from The Urban Land Institute's publication, *Dollars and Cents of Shopping Centers 1978*. For each shopping center, the proxy for actual sales is the average median sales per square foot multiplied by the center's gross leasable area.

## References

Adler, T. J. and M. Ben-Akiva, Joint-Choice Model for Frequency, Destination, and Travel Mode for Shopping Trips, *Transportation Research Record*, 1976, No. 569, 136–50.

Atlanta Journal and Constitution Newspapers, *Atlanta Regional Centers: 15-County Metro Atlanta*, Atlanta: Atlanta Newspapers, Inc., 1977.

———, *Shopping Centers in 5-County Metropolitan Atlanta*, Atlanta: Atlanta Newspapers, Inc., 1970.

Clapp, J. M., The Intrametropolitan Location of Office Activities, *Journal of Regional Science*, 1980, 20:3, 387–99.

Cleveland, D. E. and E. A. Mueller, *Traffic Characteristics at Regional Shopping Centers*, New Haven, Conn.: Yale University, Bureau of Traffic, 1961.

Daniels, P. W., *Office Location and the Journey to Work: A Comparative Study of Five Urban Areas*, Westmead, Farnborough, Hants, England: Gower Publishing Co., 1980.

Davies, D. J., An Analysis of Some Failure Data, *Journal of the American Statistical Association*, June 1952, 47, 113–50.

Dent, B. D., Trade Area Analysis of Atlanta's Regional Shopping Centers, Atlanta: Georgia State University, Department of Geography, 1978.

Domencich T. A. and D. McFadden, *Urban Travel Demand*, New York: American Elsevier Publishing Co., 1975.

Ehrenberg, A. S. C., The Pattern of Consumer Purchases, *Applied Statistics*, 1959, 8, 26–41.

Ellwood, L. W., Estimating Potential Volume of Proposed Shopping Centers, *Appraisal Journal*, 1954, 22, 581–89.

Flowerdew, R. and M. Aitkin, A Method of Fitting the Gravity Model Based on the Poisson

Distribution, *Journal of Regional Science*, 1982, 22, 191–202.

Gart, J. J., The Analysis of Poisson Regression with an Application in Virology, *Biometrika*, December 1964, 51, 517–21.

Hausman, J., B. H. Hall and Z. Griliches, Econometric Models for Count Data with an Application to the Patients–R&D Relationship, *Econometrica*, July 1984, 52, 909–38.

Huff, D. L., A Probabilistic Analysis of Shopping Center Trade Area, *Land Economics*, 1963, 53, 81–90.

Ingene, C. A. and R. F. Lusch, Market Selection Decisions for Department Stores, *Journal of Retailing*, 1980, 56:3, 21–40.

——— and E. S. H. Yu, Determinants of Retail Sales in SMSAs, *Regional Science and Urban Economics*, 1981, 11, 529–47.

Koppelman, F. R. and J. R. Hauser, *Consumer Travel Choice Behavior: An Empirical Analysis of Destination for Non-Grocery Shopping Trips*, Evanston, Ill: Northwestern University, Transportation Center, 1977.

Kumar, T. K. and W.-F. P. Shih, An Application of a Multiple Regression Model of the Poisson Process to the Murder Supply Equation, *American Statistical Association: Proceedings of the Business and Economic Statistics Section*, 1978, 715–20.

Lakshmanan, T. R. and W. G. Hansen, A Retail Market Potential Model, *Journal of the American Institute of Planners*, May 1965, 134–76.

Lalonde, B. J., Differential in Supermarket Drawing Power, Economics and Business Research, East Lansing, Mich: Michigan State University, 1962.

McDougall, E. J., A Consumer Utility Model for Allocation of Sales Among Major Retail Centers, Ph.D. dissertation, University of Florida, 1978.

Nourse, H. O., Improve Investment Decisions by Breaking the DCF Habit, *Real Estate Review*, 1986, 75–80.

Okoruwa, A. A., J. V. Terza and H. O. Nourse, Estimating Patronization Shares for Urban Retail Centers: An Extension of the Poisson Gravity Model, *Journal of Urban Economics*, 1988, 24, 241–59.

———, Predicting Patronization Rates for Urban Retail Centers: An Extension of the Poisson Gravity Model, Ph.D. dissertation, University of Georgia, 1985.

Pankhurst, I. C. and P. E. Roe, An Empirical Study of Two Shopping Models, *Regional Studies*, 1978, 12, 727–48.

Paull, A. E., A Generalized Compound Poisson Model for Consumer Purchase Panel Data, *Journal of the American Statistical Association*, December 1978, 73, 706–13.

Reilly, W. J., The Law of Retail Gravitation, New York: Putnam Co., 1931.

*Sales and Marketing Management Magazine*, The Survey of Buying Power: Data Service 1987, New York: Sales & Marketing Management Magazine, 1987.

Stanley, T. J. and M. A. Sewall, Image Inputs to a Probabilistic Model: Predicting Retail Potential, *Journal of Marketing*, July 1976, 40, 48–53.

Terza, J. V., A Tobit-Type Estimator for the Censored Poisson Regression Model, *Economics Letters*, 1985, 18, 361–65.

——— and A. A. Okoruwa, An Algorithm for the Estimation of Poisson Regressions Involving Structural Change, *Communications in Statistics*, 1985, 14, 853–66.

Turner, R. and H. S. D. Cole, An Investigation into the Estimation and Reliability of Urban Shopping Models, *Urban Studies*, 1980, 17, 139–57.

The Urban Land Institute, *Dollars and Cents of Shopping Centers 1978*, Washington D.C.: The Urban Land Institute, 1978.

U.S. Department of Commerce, Bureau of the Census, *State and Metropolitan Area Data Book, 1982.* Washington, D.C.: 1982.

Weber, D. C., Accident Rate Potential: An Application of Multiple Regression Analysis of a Poisson Process, *Journal of the American Statistical Association*, June 1971, 66, 285–88.