APPLICATION OF COUNT DATA PROCEDURES TO ESTIMATE THRESHOLDS FOR RURAL COMMERCIAL SECTORS

Thomas R. Harris, Kalyan Chakraborty, Lijuan Xiao, and Rangesan Narayanan*

Abstract-This paper extends previous research in the estimation of minimum demand thresholds for rural commercial sectors by employing count data procedures. Advantages of count data procedures are contrasted with the traditional double-log model. Also discussed is the incorporation of results from count data procedures into a rural commercial sector development strategy.

I. INTRODUCTION

Traditionally, rural economic development has concentrated on the recruitment and attraction of export-oriented, goods-producing industries. Industrial recruitment yields industries which are primarily export-oriented and provide a base for existing local economic sectors while generating input demands for further economic development. Important for local economic development professionals, attraction of a goods-producing industry such as manufacturing is highly visible. The direct employment and income effects of the relocated industry are measurable and the local community economic development team usually reaps abundant media coverage.

Nonetheless, industrial recruitment programs prove to be costly, risky and often yield little payoff. Rural communities are often unsuccessful at industrial recruitment because these communities have very limited resources (Hansen 1970). In order to attract goods-producing industries, rural communities with meager resources often grant tax concessions to new or relocating firms, thereby eliminating opportunities for fiscal gain (Kieschnick 1981, Shaffer 1989). Usually the outcome of this type of industrial recruitment is that the local tax burden of the resident populace in the local community increases because increased community

*Thomas R. Harris is a Professor in the Department of Agricultural Economics and Director of the University Center for Economic Development at the University of Nevada, Reno. Kalyan Chakraborty is a former Graduate Research Associate in the Department of Agricultural Economics. Lijuan Xiao is a former Graduate Student in the Department of Agricultural Economics. Rangesan Narayanan is Professor and Chair in the Department of Agricultural Economics at the University of Nevada, Reno. The authors would like to express gratitude to J.S. Shonkwiler for review of this article. This research was conducted under funding by the Nevada Agricultural Experiment Station Project No. 219 and the University Center for Economic Development.
services for the new industry are incurred without an expanding tax base due to the tax moratoria (Tweetan and Brinkman 1976). Moreover, firms willing to relocate because of incentives and tax abatements are also likely to leave the community if other communities offer better inducements. Results of recent surveys (Smith and Fox 1990; McNamara and Kriesel 1993) continue to show that planning commissions still emphasize the recruitment of export-oriented or goods-producing industries, while the pursuit of alternative economic development strategies, such as local services and retail sector development, are largely overlooked and often neglected.

Questions regarding the development and expansion of rural commercial sectors may be addressed by the economic development strategy of import substitution. Import substitution seeks to replace goods and services imported from outside the area with local sources of supply (Shaffer 1989). Because expenditures remain inside the local economy instead of being lost to imports, import substitution strategies strengthen linkages within the local economy. Also, keeping earned surplus within the local economy enhances local employment and incomes (Smith, 1994). For current and future time periods, local economic development strategies must give balanced emphasis to the formulation of import substitution strategies as well as relocation of goods-producing industries.

A commercial sector market analysis tool commonly used to estimate rural commercial sector activity is demand threshold analysis. The demand threshold is defined as the minimum market size required to support a particular good or service and still yield an acceptable rate of return for the business owner (Berry and Garrison 1958a, 1958b; Parr and Denike 1970; Salyards and Leitner 1981 and King 1984). The concept is based on the internal economy of the firm and the characteristics of consumer demand. As dictated by central place theory, (the foundation for threshold analysis), thresholds are not absolute, but vary by good and service. Demand thresholds are usually measured in terms of population required to support one or more firms of a certain type.

Empirical estimates of market thresholds are numerous (Berry and Garrison 1958a, 1958b; Fouest and Pickett 1974; Murray and Harris 1978; Salyards and Leitner 1981). However, all of the past studies employed ordinary least squares procedures and truncated data sets, which may provide biased estimates. This paper examines the use of count data procedures to estimate threshold levels for rural retail establishments. Specific objectives are to review demand threshold analysis, to discuss count data procedures, and review threshold demand results employing count data procedures.
A Review of Market Threshold Analysis

Threshold analysis is rooted in central place theory (CPT) in two ways. First, CPT predicts that there is a direct and positive relationship between the population of the central place and the number of firms. Here, number of functions can be proxied by the number of firms within the central place. In other words, as the population of the central place increases, so do the number of firms within the place.

Second, and perhaps more fundamental, CPT predicts that goods will have a specific limitation to the size of their market in a spatial sense. The radius of this market determines the range of the good. The larger the range of the good, the larger the spatial size of the market supporting that good. The key determinants of a good's range are the demand for the good and the cost of supplying the good. Specifically, the interaction of the Losch demand cone and the firm's average cost curve determines the range or market size of the good. Given that the cost structure facing the firm is determined exogenously from CPT (i.e., factor prices and good's production technology) the primary determinant of a good's range, or spatial market, will be the characteristics of the good's aggregate demand structure (i.e., demand cone). A spatial equilibrium is achieved when the dollar volume under the demand structure is just sufficient to cover operating costs and allow an acceptable rate of return.

Threshold analysis attempts to proxy the demand structure for a good by relating population to the number of functions (i.e., number of businesses) within a particular central place. Berry and Garrison (1958a, 1958b) suggested that this relationship can be expressed as

\[ P = \alpha B^\beta \]  \hspace{1cm} (1)

where \( P \) is the place's population, \( B \) is the number of businesses of a particular type within the place, and \( \alpha \) and \( \beta \) are parameters to be estimated. The nonlinear specification follows from CPT. In practice, the estimated equation is a double-log model. Given estimates of \( \alpha \) and \( \beta \), one may substitute \( B = 1 \) and solve for the population required to support one firm. Hence, a proxy measure for the size of the supporting demand structure for the good is provided.

The use of this specification for estimating market thresholds raises several problems. First, the use of a logarithmic transformation affects the nature of the estimates produced. The regression procedures estimate the logarithm of the number of businesses, not the number of businesses themselves. The antilog of these estimates are biased estimates of the number of businesses (Haworth and Vincent 1979).
A second difficulty arises by the use of the logarithmic transformation when a place's number of businesses for a particular type is zero. Since the logarithm of zero is negative infinity, a small positive number is usually added to all observations, or zero observations are removed from the sample. In rural areas where there are numerous places with no retail activity in some sectors, this difficulty can lead to serious problems. Adding a small positive number will result in upward, nonparallel shift of the relationship and biased estimates of threshold populations.

A third problem many past researchers seemed to share was a reversal of the logical cause-effect relationship between population and number of businesses (Chrisman 1985). Berry and Garrison (1958a) for example, regress number of businesses onto population. Because the number of businesses is the random variable within the problem, placing it on the right-hand side of the equation results in both biased and inconsistent estimates. Not all threshold studies, however, are subject to this shortcoming (Foust and de Souza 1977; Foust and Pickett 1974).

A fourth shortcoming of the bulk of the empirical threshold literature is the sparseness of the specification of the estimated equation. Numerous studies use population as the sole determinant of market demand. As argued by Murray and Harris (1978), the number of businesses supported by a given population is influenced by many factors. Other studies or retail activity have determined that socioeconomic factors, such as income levels and distribution, population density and spatial competition, can dramatically affect the size and shape of the market demand cone (Deller and Chicoine 1989; Henderson 1990). By omitting relevant variables, the parameter estimates will be biased.

A final problem concerns the use of OLS procedures to estimate numbers of businesses. Ordinary least squares assume that the number of businesses are normally distributed, which implies that the possible values which can be taken by the random variable are normally distributed around the estimate. There is little reason to suppose the values are normal. In fact, the number of firms is a non-negative integer which would suggest count data procedures.

**Count Data Procedures**

Count data procedures can be used when the dependent variable takes on values that are only non-negative integer values corresponding to the number of events (or number of stores) occurring in a given interval (or community). In such cases, the Poisson regression model has been used to analyze count data. The model stipulates that each \( Y_i \) is drawn from a Poisson distribution with parameter \( \lambda_i \), which can be parameterized to depend upon the regressors, \( X_i \)'s. The Poisson distribution is stated as:
\[
\text{Prob}(Y_i = y_i) = \exp(-\lambda_i) \frac{\lambda_i^{y_i}}{y_i!} \quad y_i = 0, 1, 2, \ldots
\] (2)

The most common formulation for \( \lambda_i \) is to assume \( \lambda_i \) is loglinearly dependent on explanatory variables or:

\[
\ln \lambda_i = \beta_0 + \sum_{j=1}^{P} \beta_j X_{ij}
\]

\( i = 1, 2, \ldots, n \)

\( j = 1, 2, \ldots, P \)

The Poisson distribution is simply a nonlinear regression. Estimation of parameters is made through maximum likelihood techniques or the log-likelihood function which is:

\[
\ln L = \sum_{i=1}^{n} -\lambda_i + \beta_0 \sum_{i=1}^{n} Y_i + \sum_{j=1}^{P} \beta_j \sum_{i=1}^{n} X_{ij} - \sum_{i=1}^{n} \ln(Y_i!)
\] (4)

The Poisson distribution was estimated using the LIMDEP package of Greene (1989).

Several studies have been completed using the Poisson regression model in econometric and statistical literature. Advances in the theory of estimation can be found in works by Jorgenson (1961); Frome, et al. (1973); Gourieroux et al. (1984); Terza (1985); and Madala (1983). The technique has been successfully applied by Flowerdew and Aitkin (1982); Hausman et al. (1984); Gilbert (1979) and Cameron et al. (1984). An excellent literature review and general overview of count data modeling is presented in a referenced study by Cameron and Trivedi (1986).

However, when using a Poisson distribution for count data analysis, it is assumed that the conditional mean of the dependent variable, \( \lambda_i \), is equal to the conditional variance; i.e., the variance-mean ratio is one. Overdispersion of the population is defined as the conditional variance of the dependent variable exceeding its conditional mean, giving a variance-mean ratio greater than one. Overdispersion is a form of heteroscedasticity. If the population is overdispersed, the Poisson model will yield consistent estimates of the parameters but downwardly biased estimates of the standard errors (Gourieroux et al. 1984). The generalization of the Poisson distribution, which is often used to model such counts, is the negative binomial distribution (Hausman et al. 1984; Cameron and Trivedi (1986).
The negative binomial distribution is an extension of the Poisson distribution which allows the variance of the process to differ from the mean. The negative binomial arises from the Poisson model by specifying:

\[ \ln \lambda_i = \beta X + \varepsilon, \]  

(5)

where \( \varepsilon \) has a gamma distribution with mean parameter \( \Phi \) and index (or precision) parameter \( \alpha_i \). The mean and variance of the negative binomial distribution are given by:

\[ E[Y_i] = \Phi_i \]  

(6)

\[ Var[Y_i] = \Phi_i + \frac{1}{\alpha_i} \Phi_i^2 \]  

(7)

Since \( \Phi_i > 0 \) and \( \alpha_i > 0 \), it is clear that the variance exceeds the mean, so the model allows for overdispersion. Using the LIMDEP software, maximum likelihood estimates for the negative binomial are obtained. Starting values are obtained from Poisson parameters and an estimate of the variance of \( Y_i \).

**Model Specification**

Not all states have sales taxes, but for those states that have a sales tax, retail sales may not be recorded at point of sale. The eight two-digit SIC retail sectors from the 1987 Census of Retail Trade (1990) were used to derive minimum demand thresholds.\(^1\) Data for the minimum demand analysis were gathered from the 1988 *City and County Data Book* (1988) and the 1987 *Census of Retail Trade* (1990). The final data set contained 2,126 communities. This study differs from most other market threshold studies because the unit of observation is the county as opposed to town or city. It would have been advantageous to complete the analysis at the community level; however, for some states the only reliable retail establishment numbers are produced at the county level by the Census of Retail Trade.\(^2\) Therefore it is anticipated that the market population estimates from this study may not be directly comparable to previous studies.

For this analysis, the double log, Poisson, and negative binomial models were employed where:

\[ B = f(POP, PCI, PPBEL, DEN, UNRATE, PESTiS, ADJ) \]  

(8)

where:
Application of Count Data Procedures to Estimate

B is the number of retail establishments by type
POP is population in 1986
PCI is per capita income in 1985
PPBEL is the percent of persons below poverty level (1979)
DEN is the population density (1986)
UNRATE is the unemployment rate (1986)
PESTiiS is the percent of total number of establishments in SICii that are small (defined as having no employees)
ADJ is the dummy variable for adjacency to metropolitan counties (1 = yes; 0 = no)

For this analysis, population, per capita income, and population density were transformed into log values. The inclusion of explanatory variables beyond population builds upon the models developed by Murray and Harris (1978). Per capita income, percent of persons below poverty level, and unemployment rate were intended to capture the buying power of consumers within a county. It is expected that per capita income will be positively related to number of firms because of enhanced buying power by local residents, while increases in poverty level and unemployment rate would be negatively related. Population density was a proxy for ruralness of a retail market and is expected to be positively related. The metropolitan adjacency dummy captures the spatial competition and is expected to be negatively related due to heightened spatial competition of metropolitan retailers.

Inclusion of the percentage of businesses that are deemed small—that is, having no employees—is designated to capture the mix of retail businesses within a particular sector. Previous market threshold studies have implicitly assumed that businesses within a retail sector have identical cost structures. However, in rural retail markets, this may be false. For example, one county of a certain size may have one large retail store, while another county of similar population size has four small retail stores supporting local demand. Specifically, one retail market will appear to be under-retailed while the other would be classified as over-retailed. Inclusion of this variable attempts to address this problem and is assumed to be positively related to the number of retail establishments.

Results

Table 1 shows the results of the negative binomial procedures where population, per capita income, and population density were expressed in logarithmic form for the eight two-digit retail SIC sectors. Of interest is the coefficient for the negative binomial that had significant t-values for all eight retail sectors. This in-
TABLE 1
Minimum Demand Threshold Estimates Using Negative Binomial Procedures

<table>
<thead>
<tr>
<th>Variable</th>
<th>SIC-52</th>
<th>SIC-53</th>
<th>SIC-54</th>
<th>SIC-55</th>
<th>SIC-56</th>
<th>SIC-57</th>
<th>SIC-58</th>
<th>SIC-59</th>
</tr>
</thead>
<tbody>
<tr>
<td>(9.83)</td>
<td>(7.50)</td>
<td>(18.66)</td>
<td>(12.94)</td>
<td>(30.98)</td>
<td>(21.92)</td>
<td>(18.94)</td>
<td>(28.65)</td>
<td></td>
</tr>
<tr>
<td>Population*</td>
<td>0.8908</td>
<td>0.8593</td>
<td>0.8799</td>
<td>0.9409</td>
<td>1.1294</td>
<td>1.0619</td>
<td>1.0384</td>
<td>0.9983</td>
</tr>
<tr>
<td>(71.78)</td>
<td>(53.86)</td>
<td>(82.67)</td>
<td>(88.45)</td>
<td>(63.03)</td>
<td>(78.03)</td>
<td>(75.74)</td>
<td>(96.92)</td>
<td></td>
</tr>
<tr>
<td>Per Capita Income*</td>
<td>0.0839</td>
<td>-0.0378</td>
<td>0.5306</td>
<td>0.0751</td>
<td>1.5203</td>
<td>0.7464</td>
<td>0.4397</td>
<td>0.8440</td>
</tr>
<tr>
<td>(1.20)</td>
<td>(0.41)</td>
<td>(8.52)</td>
<td>(1.42)</td>
<td>(20.29)</td>
<td>(10.64)</td>
<td>(8.33)</td>
<td>(18.00)</td>
<td></td>
</tr>
<tr>
<td>% Below Poverty</td>
<td>-0.0112</td>
<td>0.0119</td>
<td>0.0255</td>
<td>-0.0031</td>
<td>0.0197</td>
<td>0.0007</td>
<td>-0.0216</td>
<td>-0.0058</td>
</tr>
<tr>
<td>(6.41)</td>
<td>(5.45)</td>
<td>(16.51)</td>
<td>(2.31)</td>
<td>(8.92)</td>
<td>(0.39)</td>
<td>(14.20)</td>
<td>(4.13)</td>
<td></td>
</tr>
<tr>
<td>Population Density*</td>
<td>-0.0786</td>
<td>-0.0051</td>
<td>0.0169</td>
<td>-0.0294</td>
<td>-0.0469</td>
<td>0.0189</td>
<td>-0.1739</td>
<td>-0.0869</td>
</tr>
<tr>
<td>(8.61)</td>
<td>(0.43)</td>
<td>(2.10)</td>
<td>(3.52)</td>
<td>(3.80)</td>
<td>(2.05)</td>
<td>(16.57)</td>
<td>(11.52)</td>
<td></td>
</tr>
<tr>
<td>Unemployment Rate</td>
<td>-0.0178</td>
<td>-0.0038</td>
<td>0.0055</td>
<td>-0.0046</td>
<td>-0.0046</td>
<td>-0.0152</td>
<td>-0.0029</td>
<td>-0.0070</td>
</tr>
<tr>
<td>(9.36)</td>
<td>(1.56)</td>
<td>(3.33)</td>
<td>(2.84)</td>
<td>(1.97)</td>
<td>(7.09)</td>
<td>(1.56)</td>
<td>(5.30)</td>
<td></td>
</tr>
<tr>
<td>% Small Establishments</td>
<td>-0.0008</td>
<td>0.0045</td>
<td>0.0080</td>
<td>0.0091</td>
<td>0.0024</td>
<td>0.0023</td>
<td>0.0036</td>
<td>0.0027</td>
</tr>
<tr>
<td>(1.78)</td>
<td>(10.07)</td>
<td>(16.32)</td>
<td>(23.95)</td>
<td>(7.84)</td>
<td>(6.43)</td>
<td>(6.78)</td>
<td>(5.48)</td>
<td></td>
</tr>
<tr>
<td>Adjacent Dummy Variable</td>
<td>-0.0667</td>
<td>-0.0816</td>
<td>-0.0578</td>
<td>-0.0835</td>
<td>-0.3396</td>
<td>-0.2223</td>
<td>-0.1179</td>
<td>-0.1278</td>
</tr>
<tr>
<td>α</td>
<td>0.0320</td>
<td>0.0293</td>
<td>0.0456</td>
<td>0.0352</td>
<td>0.1082</td>
<td>0.0430</td>
<td>0.0968</td>
<td>0.0571</td>
</tr>
</tbody>
</table>

\*Values are in log
\bAbsolute values of the t-statistic in parenthesis

The performance of the individual variables is in general as expected. By far the most dominant variable, as suggested by theory and previous empirical work, was population. In each case population was positive as expected and significant. Also, where significant, per capita income is positive. For most retail sectors, the number of establishments tends to be more sensitive to changes in population than changes in per capita income. This result is consistent with a previous study by Henderson (1990). Performance of the remaining independent variables was mostly consistent with prior expectations. Also, the variable accounting for variation in establishment size mix, when significant, had the correct sign. For all eight retail sectors, counties adjacent to a metropolitan area tend to support fewer establishments, given all other variables held constant.

In Table 2, the population thresholds are reported for a given number of firms by retail SIC. Due to the inclusion of several socio-economic variables within the specification of the threshold equations, market population estimates can be uniquely derived for any given county. For illustrative purposes, assume
TABLE 2
Population Required to Support a Given Number of Firms by SIC Class Based on Negative Binomial Estimates.

<table>
<thead>
<tr>
<th>SIC</th>
<th>Number of Firms</th>
</tr>
</thead>
<tbody>
<tr>
<td>52 Building/Gardening Merchandise</td>
<td>1    2    3    4    5</td>
</tr>
<tr>
<td></td>
<td>1,060 2,309 3,640 5,027 6,458</td>
</tr>
<tr>
<td>53 General Merchandise</td>
<td>1,399 3,045 4,801 6,631 8,518</td>
</tr>
<tr>
<td>54 Food Store</td>
<td>151 332 527 731 942</td>
</tr>
<tr>
<td>55 Automobile Dealers and Gasoline Stations</td>
<td>378 791 1,216 1,651 2,093</td>
</tr>
<tr>
<td>56 Apparel and Accessory Stores</td>
<td>1,894 3,498 5,009 6,463 7,874</td>
</tr>
<tr>
<td>57 Furniture and Home Furnishings</td>
<td>1,507 2,894 4,240 5,559 6,859</td>
</tr>
<tr>
<td>58 Eating and Drinking Places</td>
<td>555 1,082 1,598 2,109 2,614</td>
</tr>
<tr>
<td>59 Miscellaneous Retail</td>
<td>285 570 856 1,142 1,427</td>
</tr>
</tbody>
</table>

that the analyst is interested in market population for an average non-adjacent county (Table 3). For Apparel and Accessory Stores, using negative binomial procedures, the minimum demand for one store is 1,894; for two it is 3,498; for three it is 5,009; for four it is 6,463, and for five it is 7,874 for average values of non-adjacent rural counties. If ordinary least squares procedures are employed, the minimum demand for one store is 2,109; for two it is 3,831; for three it is 5,433; for four it is 6,960 and for five it is 8,435. The calculated minimum demand threshold for the above example is lower when negative binomial procedures are employed. One reason is that the negative binomial does not truncate the data, but allows zero values to be incorporated in the estimation procedures. Also, because the number of retail establishments is not normally distributed, the analysis should be applied with count data procedures such as the negative binomial.

In addition, another application of negative binomial procedures is that the expected number of retail establishments for a given county can be estimated. For the average non-adjacent nonmetropolitan county, the expected number of apparel and accessory stores is 15.03. When the expected value is compared to actual establishment numbers, economic development professionals can use this information to pursue retail development strategies. In addition, count data procedures
TABLE 3
Independent Variable Values for an Average Non-Adjacent County

<table>
<thead>
<tr>
<th>Variables</th>
<th>Units</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Population</td>
<td>nos.</td>
<td>20,861</td>
</tr>
<tr>
<td>Per Capita Income</td>
<td>$</td>
<td>7,805</td>
</tr>
<tr>
<td>Percent Below Poverty</td>
<td>%</td>
<td>17.73</td>
</tr>
<tr>
<td>Population Density</td>
<td>nos.</td>
<td>32.09</td>
</tr>
<tr>
<td>Unemployment Rate</td>
<td>%</td>
<td>9.30</td>
</tr>
<tr>
<td>Percent Small Establishments for SIC 52</td>
<td>%</td>
<td>34.38</td>
</tr>
<tr>
<td>Percent Small Establishments for SIC 53</td>
<td>%</td>
<td>30.35</td>
</tr>
<tr>
<td>Percent Small Establishments for SIC 54</td>
<td>%</td>
<td>36.94</td>
</tr>
<tr>
<td>Percent Small Establishments for SIC 55</td>
<td>%</td>
<td>38.11</td>
</tr>
<tr>
<td>Percent Small Establishments for SIC 56</td>
<td>%</td>
<td>31.56</td>
</tr>
<tr>
<td>Percent Small Establishments for SIC 57</td>
<td>%</td>
<td>51.04</td>
</tr>
<tr>
<td>Percent Small Establishments for SIC 58</td>
<td>%</td>
<td>26.83</td>
</tr>
<tr>
<td>Percent Small Establishments for SIC 59</td>
<td>%</td>
<td>61.55</td>
</tr>
</tbody>
</table>

such as the negative binomial can derive the probability of a given number of apparel and accessory stores existing in an average nonadjacent nonmetropolitan county. For the apparel and accessory store example, the probability of no apparel and accessory stores is 0 percent, while for 13 stores it is 9.52 percent; for 14 stores it is 10.22 percent; for 15 stores it is 10.24 percent and for 16 stores it is 9.62 percent. Using the highest probability value, the example county has a threshold for 15 stores which correspond to the expected value.

Results from ordinary least squares procedures would produce minimum threshold values but could not estimate probabilities of a given number of establishments. The probability results from the negative binomial could be incorporated into a retail development strategy. This would give economic development professionals an estimate of potential risks in advocating a given retail development strategy. Past retail development strategies which employ minimum demand thresholds derived from ordinary least squares procedures have been unable to incorporate risk into a development plan.

Conclusions

The primary objective of this paper was to add upon previous minimum demand threshold studies in two regards. First, contemporary minimum demand threshold studies built around population have several disadvantages ranging from model misspecification to unnecessary structure due to a specific functional form. Second, an alternative approach to ordinary least squares procedures was applied
to estimate minimum thresholds. Because retail store establishment numbers are non-negative and integer values, the assumption that retail business counts are normally distributed is suspect. Business establishment numbers are characterized as count data where the Poisson and negative binomial distribution are better suited. Results from the negative binomial distribution indicated that the number of retail establishments were overdispersed, so the negative binomial was the appropriate procedure to employ in estimating minimum thresholds.

Following suggestions by Murray and Harris, (1978) additional socio-economic data besides population was used to derive minimum thresholds. Results indicated that county population and per capita income, when significant, were positive influences on the number of retail establishments. Also, results indicated that number of retail establishments were more sensitive to population change than per capita income. In addition, for all retail sectors analyzed, counties adjacent to a metropolitan county tend to support fewer establishments, given all other variables held constant.

Previous market threshold studies have not addressed the probability of the number of retail establishments, given socio-economic characteristics of a county. Results of the negative binomial can derive probabilities of a given number of retail firms existing in a given county. Through the use of count data procedures, economic development strategies can incorporate risk of promotion of retail development in their county economy.

ENDNOTES

1. The eight retail sectors employed in this analysis were Building Materials and Garden Supply Stores (SIC 52); General Merchandise Stores (SIC 53); Food Stores (SIC 54); Automotive Dealers and Gasoline Service Stations (SIC 55); Apparel and Accessory Stores (SIC 56); Furniture and Home Furnishing Stores (SIC 57); Eating and Drinking Establishments (SIC 58) and Miscellaneous Retail Establishments (SIC 59).

2. In some states, sales taxes are not collected, and where sales taxes are collected, retail sales may be recorded at regional or state headquarters rather than at point of sale. Therefore Census of Retail Trade establishment numbers were used because of reliability of establishment numbers. Of course, where available, community retail establishment numbers are preferred.

3. Minimum demand threshold models tend to be demand driven, assuming identical cost structures across commodity groups. One possible source of misspecification is the omission of supply-related factors, specifically operating costs, which will affect returns.
4. Minimum thresholds were derived for ordinary least squares and Poisson procedures. Because of space limitations results of these procedures are not presented.

REFERENCES


Application of Count Data Procedures to Estimate


