

X = A matrix of other demand-related variables;

v = An error term that captures randomness in the data (hence, the frontier is stochastic); and

u = An error term whose elements capture firms that are willing to accept a lower rate of return.

This latter component of the error structure (u) is important because it explicitly captures our hypothesis of overretailing in rural markets. Adding u, as opposed to subtracting, to the equation envelops the bottom of the data; thus, true market minimums as opposed to market averages are derived.

To estimate this minimum market stochastic frontier equation, we employ the method suggested by Aigner, Lovell, and Schmidt (1977). One of the system parameters to be estimated in the Aigner, Lovell, and Schmidt approach is of special interest to this study:  $\lambda (= \sigma_u/\sigma_v)$ . If the data do not support the hypothesis of overretailing in rural markets, then  $\sigma_u$  will be statistically equivalent to zero, or  $\lambda = 0$ . This hypothesis can be tested by computing a t-statistic for the system parameter  $\lambda$ . If the computed t-statistic suggests that  $\lambda = 0$ , then the frontier estimate and the ordinary least squares estimate are essentially the same, and there is no evidence of overretailing. On the other hand, if  $\lambda$  is statistically different from zero, there exists necessary, but not sufficient, evidence to suggest that some rural markets are overretailed due to lower acceptable rates of return.

### Model Specification

Data used to estimate both average and minimum market thresholds were gathered from the *1988 City and County Data Book*, the *1987 Census of Retail*, and the USDA, Economic Research Service. The final data set contained observations on 2,126 rural U.S. counties. This study differs from most other market threshold studies because the unit of observation is the county as opposed to the town or city.<sup>3</sup>

To avoid the complication of zero elements and anti-log biases within the log-log specification, a nonlinear quadratic specification in population of the model assumed:

---

<sup>3</sup> While the county may not appear to be consistent with the notion of central places, it better reflects the shopping area of rural residents. In other words, in rural areas, the interpretation of a central place as a specific town may be too narrow.

$$(3) B = \alpha_0 + \beta_1 \text{POP} + \beta_2 \text{POP}^2 + \beta_3 \text{PCI} + \beta_4 \text{PPBEL} + \beta_5 \text{DEN} + \beta_6 \text{UNRATE} \\ + \beta_7 \text{PESTS}_{ii} + \beta_8 \text{ADJ} + u + v.$$

where:

**B** = The total number of retail establishments by type, including those businesses both with and without employees.

The vector **X** includes the following:

- POP** = Population (1986)
- POP<sup>2</sup>** = Population squared;
- PCI** = Per capita income (1985);
- PPBEL** = Percent of persons below poverty level (1979);
- DEN** = Population density;
- UNRATE** = Unemployment rate (1986);
- PESTS<sub>ii</sub>** = Percent of total number of establishments in SIC ii that are small (defined as no employees);
- ADJ** = Dummy for adjacent to metro area (1 = yes, 0 = no).

The inclusion of variables beyond population builds on the models of Murray and Harris (1978).<sup>4</sup> Income, the poverty measure, and unemployment rate are intended to capture the buying power of consumers within the market. We expect income to be related positively due to the enhanced buying power of county residents, but we expect negative coefficients on the poverty measure and the unemployment rate for similar reasons. Population density is a proxy for how rural the market is, and it is expected to be positive. The metro-adjacent dummy captures spatial competition and is expected to have a negative coefficient due to the heightened spatial competition afforded by nearby retailers. Further, we also expect the coefficients on population and population squared to be positive due to agglomeration economies in retailing.

We include the percentages of businesses that are deemed to be small (i.e., no employees) to capture the mix of businesses within a particular sector. All studies of market threshold to date implicitly have assumed that because businesses within a sector are characterized by identical cost structures, they will be similar in size. In rural areas this need not be the case. For example, in one place of a certain population there may be one large supermarket servicing the market area. A neighboring market area of equal size (population) may have three smaller grocers supporting local demand. To treat these two markets the same

---

<sup>4</sup> While we rationalize the empirical specification of the model, a detailed theoretical derivation of the specification is beyond the scope of this analysis.

will result in misspecification of the model. One market will appear to be underretailed, while the other will appear to be overretailed. Inclusion of this variable partially addresses the problem, and we expect the coefficient to be positive.<sup>5</sup>

### **Empirical Results**

The frontier estimates for the eight two digit SIC retail sectors are presented in Table 1. For the OLS analysis (not reported here), the intercept term is positive and usually significant across all eight sectors.<sup>6</sup> The OLS results imply that a barren county would support a certain number of establishments (the value of the intercept term). While this result is an economic absurdity, it is consistent with the hypothesis that rural areas may be overretailed. When OLS procedures are applied to log-log specifications, the problem of a positive intercept disappears. One wonders if the popularity of the double-log specification for market threshold analysis is a result of the elimination of the positive intercept problem. If this is the case, the existing literature also may suffer from functional form bias.

The estimated frontier system parameter  $\lambda$  is significantly different from zero at the 99 percent level of confidence for each of the eight threshold equations. This result confirms our hypothesis that some rural retail markets are overretailed due to lower rates of acceptable return. Thus, the county data support the notion that enveloping the bottom of the data yields a more conservative and arguably more realistic estimate of market population.

The performance of the individual variables is, in general, as one would expect. The dominant variable, as suggested by theory and earlier empirical studies, is population. In each case, the coefficient is positive, as expected, and highly significant. Population squared is generally positive, although small in magnitude, and statistically significant. This results hint at agglomeration economies in retailing, as predicted by theory. Income is positive and also significant. In most cases the number of establishments tends to be more sensitive to changes in population than to changes in per capita income. This result is consistent with the work of Henderson (1990). Performance of the remaining variables tends to be consistent with prior expectations, with a few exceptions. The variable accounting for variation in establishment size mix is of the correct sign when significant. Also, in each case,

---

<sup>5</sup> One possible source of misspecification is the omission of supply-related factors, specifically operating costs, that will affect rates of return. These models, however, tend to be demand driven, assuming identical cost structures across commodity groups.

<sup>6</sup> See Table 2 for a complete list of the SIC sectors used.

counties that are adjacent to metropolitan areas tend to support fewer establishments, everything else held constant. Space limitations prevent a more detailed discussion of each individual coefficient.

Table 2 reports the population threshold for a given number of firms by retail SIC. Due to the inclusion of several socioeconomic variables in the specification of the threshold equations, market population estimates can be uniquely constructed for any particular county. For illustrative purposes assume that the analyst is interested in the market population for a typical (i.e., sample mean) nonadjacent county. One can estimate the population required to support one merchandising store using the estimated equation for general merchandising. The market population threshold is 7,250 persons using the frontier estimates, while the OLS estimates suggests a population threshold of only 450 persons. Given the positive intercept using OLS, this latter result (450 persons) is suspect, but it again supports the notion of overretailing in rural markets. The frontier estimator provides consistently more conservative (many times profoundly more conservative) estimates than those reported in the literature.

To further illustrate the potential level of overretailing in many rural markets, the observed number of establishments is compared to the estimated minimum derived from the frontier estimates. Four counties are examined: Elko and Churchill in Nevada and Aroostook and Hancock in Maine. The characteristics of each of these four counties and the estimated minimum and observed number of establishments are provided in Table 3. We can construct a simple measure of overretailing by dividing the estimated minimum by the observed number of establishments. The closer the measure is to one, the closer is the rate of return earned in the market to a normal rate of return. The smaller the measure, the lower is the rate of return due to overretailing. If the measure is greater than one, the data suggest that there are market opportunities that are not being realized.

In all cases except the number of eating and drinking establishments in Aroostook County, Maine (the minimum-to-observed ratio is greater than one: 1.15), the number of observed establishments is greater than the true minimum that can be supported by the market. We posit that the greater the divergence between the minimum and the observed (or the smaller the minimum to observed ratio), the greater is the degree of overretailing and the lower the rates of return an entrepreneur should expect to earn in a given market.<sup>7</sup>

---

<sup>7</sup> The estimated minimum number of establishments reported in the lower portion of Table 3 should not be confused with the population thresholds reported in Table 2. For example, Table 2 would lead one to expect that Churchill, Nevada should be able to support three general merchandise establishments, but Table 3 reports that the estimated minimum is only one establishment. Table 3 is more refined—income

Two important points must be highlighted based on the evidence provided in Table 3. First, a higher degree of overretailing does not necessarily imply that the market could not support another establishment. Rather, the higher degrees of overretailing suggest that firms in the market are accepting lower rates of return than those in other markets. Prospective entrepreneurs need to be aware of this market position. Second, unique characteristics of individual markets must be taken into account. Hancock County, Maine appears to be heavily overretailed in eating and drinking establishments and miscellaneous retail. Hancock County is the home of Bar Harbour and Acadia National Park, the most visited national park on the east coast. In short, Hancock County supports an enormous tourist economy. Our simple frontier equation does not capture this important market condition.

### **Conclusions and Implications**

The primary objective of this research is to build upon the existing demand threshold literature in two regards. First, current population estimates are plagued with several shortcomings, ranging from model misspecification with respect to right side variables to unnecessary structure due to the selection of a specific functional form. Each of these shortcomings has been addressed in this analysis. Second, and perhaps more fundamentally, an alternative to the traditional application of the OLS estimator to the threshold problem has been suggested and applied. There is sufficient *a priori* reasoning to suspect that the distribution of the number of retail establishments is not normal, but is positively skewed. Lower opportunities in rural areas for alternative business diversification may force some retailers to accept a lower rate of return. The result is overretailing in some rural areas.

Market threshold estimates should reflect this lower rate of acceptable return by discounting areas that could be described as overretailed. This can be accomplished by putting the traditional least squares estimator to the side in favor of a method, such as the stochastic frontier estimator, that envelops the bottom of the data. This approach provides not only a more conservative estimate of market threshold, but also gives an indirect measure of market overretailing. Prospective entrepreneurs can use this information to develop a proxy measure of the expected rate of return within each market. Just as lower opportunity costs in rural areas imply lower rates of return and overretailing, empirical evidence of overretailing implies lower rates of return. We postulate that a direct relation exists that can be exploited in policies designed to revitalize downtown area across rural America.

---

and spatial competition explicitly is accounted for in the estimated minimum.

## References

1. Aigner, D.T., C.A.K. Lovell, and P.J. Schmidt, "Formulation and Estimation of Stochastic Frontier Production Function Models," *Journal of Econometrics*, 6 (1977), pp. 21-37.
2. Berry, B., and W. Garrison, "A Note on Central Place Theory and the Range of a Good," *Economic Geography*, 34 (1958a), pp. 304-311.
3. Berry, B., and W. Garrison, "Recent Developments in Central Place Theory," *Proceedings of the Regional Science Association*, 4 (1958b), pp. 107-121.
4. Chrisman, J.J., "Population Change and Its Effect on the Retail Sector: An Exploratory Study," *International Small Business Journal*, 3, no. 4 (1985), pp. 26-46.
5. Daniels, T.L., "Small Town Economic Development: Growth or Survival?" *Journal of Planning Literature*, 4, no. 4 (1989), pp. 413-429.
6. Daniels, T.L., J.W. Keller, and M.D. Lapping, *The Small Town Planning Handbook* (Chicago: APA Planners Press, 1988).
7. Deller, S.C., and D.L. Chicoine, "Economic Diversification and the Rural Economy: Evidence from Consumer Behavior," *Regional Science Perspectives*, 19, no. 2 (1989), pp. 41-55.
8. Deller, S.C., J.C. McConnon, Jr. J. Holden, and K.E. Stone, "The Measurement of a Community's Retail Market," *Journal of the Community Development Society*, 22, no. 2 (1991), pp. 68-83.
9. Foust, B., and E. Pickett, "Threshold Estimates: A Tool for Small Business Planning in Wisconsin," unpublished manuscript, Department of Geography, University of Wisconsin, Eau Claire, 1974.
10. Haworth, J., and P. Vincent, "The Stochastic Disturbance Specification and Its Implication for Log-Linear Regression," *Environment and Planning A*, 2 (1979), pp. 781-790.
11. Henderson, D., "Rural Retail Sales and Consumer Expenditure Functions," *Journal of Agricultural Economics Research*, 42, no. 3 (1990).
12. Hustedde, R., R. Shaffer, and G. Pulver, *Community Economic Analysis* (Ames, Iowa: North Central Regional Center for Rural Development, 1984).
13. King, L.J. *Central Place Theory* (London: SAGE Publications, 1984).
14. McNamara, K. and G. Green, "Local and Regional Economic Development Planning and the Role of Community Development Practitioners," *Journal of the Community Development Society*, 19 (1988), pp. 4-55.
15. Murray, J. and J. Harris, *A Regional Economic Analysis of the Turtle Mountain Indian Reservation: Determining the Potential for*

*Commercial Development* (Minneapolis, MN: Federal Reserve Bank of Minneapolis, 1978).

16. Parr, J. and K. Denike, "Theoretical Problems in Central Place Analysis," *Economic Geography*, 46 (1970), pp. 568-586.

17. Salyards, D.M. and K.R. Leitner, "Market Threshold Estimates: A Tool for Business Consulting in Minnesota," *American Journal of Small Business*, 6, no. 2 (1981), pp. 26-32.

18. Schuker, A.V. and F.L. Leistritz, "Threshold Population Levels for Rural Retail Businesses in North Dakota," unpublished paper, Department for Agricultural Economics, North Dakota State University, Fargo. 1991.

19. Shaffer, R. *Community Economics: Economic Structure and Change in Smaller Communities*, Ames (Iowa: Iowa State University Press, 1989).

20. Stone, K.E. and J.C. McConnon, Jr., "Retail Sales Migration in the Midwestern United States," paper presented at the 1980 AAEA meetings, University of Illinois, Urbana, 1980.

21. Stone, K.E., and J.C. McConnon, Jr., "Trade Area Analysis Extension Program: A Catalyst for Community Development," in *Proceedings of Realizing Your Potential as an Agricultural Economist in Extension* (Ithaca, NY, 1984).

**Table 1—Stochastic Frontier Market Population Threshold Estimates**

Variable	SIC-52	SIC-53	SIC-54	SIC-55	SIC-56	SIC-57	SIC-58	SIC-59
Intercept	0.3799 (0.24) <sup>a</sup>	-2.1450 (2.26)	-26.7921 (6.50)	1.9265 (0.47)	-22.0911 (9.04)	-11.4213 (5.89)	-20.333 (3.48)	-62.5012 (6.72)
Population	0.0005 (42.88)	0.0002 (25.38)	0.0011 (37.3900)	0.0017 (37.03)	0.0005 (25.46)	0.0006 (41.00)	0.0014 (33.55)	0.0026 (47.75)
Population Squared	4E-10 (5.59)	5E-11 (0.64)	-1E-10 (0.78)	-2E-09 (5.81)	1E-09 (12.40)	7E-10 (9.03)	5E-09 (19.00)	6E-09 (18.04)
Per Capita Income	0.0003 (2.02)	-0.0001 (1.07)	0.0007 (1.85)	-0.0007 (2.05)	0.0018 (9.62)	0.0011 (6.96)	0.0027 (6.08)	0.0067 (11.05)
% of Persons Below Poverty	-0.1270 (5.05)	0.0582 (3.63)	0.7305 (11.05)	0.1419 (2.00)	0.1927 (4.63)	-0.0032 (0.10)	-0.5767 (5.71)	-0.2214 (1.61)
Population Density	-0.0250 (10.35)	0.0056 (2.79)	0.0418 (6.04)	0.0373 (4.80)	0.0140 (3.58)	0.0282 (9.19)	-0.1134 (11.80)	-0.1063 (8.99)
Unemployment Rate	-0.1800 (5.89)	0.0105 (0.67)	0.3505 (5.12)	-0.0849 (1.00)	-0.0813 (1.48)	-0.2401 (5.95)	-0.0033 (0.03)	-0.6946 (4.64)
% Small Establishments	0.0221 (3.79)	0.0159 (5.00)	0.0086 (4.23)	-0.0005 (0.31)	-0.0014 (0.56)	-0.0013 (0.19)	0.1256 (3.53)	0.1771 (2.97)
Adjacent Dummy Variable	-0.6791 (2.75)	-0.5209 (3.55)	-1.0565 (1.69)	-2.6055 (3.91)	-4.4325 (9.75)	-3.2616 (9.60)	-4.3921 (4.72)	-7.7928 (6.03)
$\lambda$	2.1244 (32.12)	2.7020 (26.92)	2.4370 (31.21)	2.2460 (34.25)	3.7970 (23.29)	2.8877 (32.43)	3.2956 (27.42)	2.9376 (35.79)

<sup>a</sup>Absolute value of the t-statistics in parentheses

**Table 2—Population Estimate Needed to Support a Given Number of Firms by SIC Class Based on Stochastic Frontier Estimates.<sup>a</sup>**

SIC	Number of Firms				
	1	2	3	4	5
52-Building/Gardening	4,250	6,250	8,000	9,750	11,750
53-General Merchandise	7,250	10,250	13,500	16,750	20,000
54-Food	4,000	4,750	5,500	6,250	7,000
55-Auto	5,000	5,500	6,000	6,500	7,000
56-Apparel	9,500	10,750	12,500	14,250	16,000
57-Furniture	7,750	9,000	10,500	12,000	13,500
58-Eat/Drink	4,250	5,250	5,750	6,250	7,000
59-Misc.	3,750	4,250	4,500	4,750	5,000

<sup>a</sup> Estimates rounded to nearest 250 people

**Table 3—An Example of the Minimum Market Threshold Estimates Applied to Four Rural Counties.**

	Elko, Nevada	Churchill, Nevada	Aroostook, Maine	Hancock, Maine
Population	21,000	15,300	87,900	44,000
Per Capita Income	9,271	8,746	7,556	8,456
Poverty Rate	11.0	11.2	16.2	14.6
Pop. Density	1.3	3.1	13.1	28.6
Unemployment Rate	6.0	7.0	8.6	6.3
Adjacency	nonadjacent	adjacent	nonadjacent	adjacent
	minimum/ observations	minimum/ observations	minimum/ observations	minimum/ observations
52-Building/Gardening	13/14 = .93	8/10 = .80	46/57 = .81	22/46 = .48
53-General Merchandise	2/4 = .50	0/1 = .00	17/39 = .44	7/29 = .24
54-Food	13/29 = .45	6/19 = .32	90/172 = .52	41/122 = .34
55-Auto	28/38 = .74	16/26 = .62	133/167 = .80	67/111 = .60
56-Apparel	7/25 = .28	0/13 = .00	46/75 = .61	15/35 = .43
57-Furniture	10/18 = .56	3/10 = .30	53/59 = .90	22/35 = .63
58-Eat/Drink	32/63 = .51	17/44 = .39	154/134 = 1.15	60/179 = .34
59-Misc.	62/104 = .60	33/61 = .54	263/308 = .85	113/324 = .35