ESTIMATION OF MINIMUM DEMAND THRESHOLDS:
AN APPLICATION OF COUNT DATA PROCEDURES
WITH THE EXISTENCE OF EXCESS ZERO OBSERVATIONS

by

Thomas R. Harris, Steven Yen,and Steven C. Deller

Thomas R. Harris is Director of the University Center for Economic Development and Foundation Professor in the Department of Applied Economics and Statistics at the University of Nevada, Reno

Steven Yen is an Associate Professor in the Department of Applied Economics and Statistics at the University of Nevada, Reno.

Steven C. Deller is an Associate Professor in the Department of Agricultural and Applied Economics at the University of Wisconsin, Madison

Abstract

Count data models that incorporate the existence of excess zero observations were employed to estimate minimum demand thresholds for rural Wisconsin retail sectors. It was found that single- and double-hurdle models improved the estimation of rural retail minimum demand thresholds.

Presented paper at the AAEA annual meeting in Tampa, Florida, July 30 - August 2000.

ESTIMATION OF MINIMUM DEMAND THRESHOLDS: AN APPLICATION OF COUNT DATA PROCEDURES WITH THE EXISTENCE OF EXCESS ZERO OBSERVATIONS

INTRODUCTION

Traditionally, rural economic development has been 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. Importantly 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 services for the new industry are incurred without an expanding tax base due to the tax moratoria (Tweeten and Brinkman 1976). Moreover, firms that are 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 recruit 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). Import substitution strategies strengthen linkages within the local economy because expenditures remain inside the local economy instead of being lost to imports. 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; 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; Foust and Pickett 1974; Murray and Harris 1978; Salyards and Leitner 1981). However all of these past studies employed ordinary least squares procedures and truncated data sets to estimate threshold levels for rural retail establishments. Studies by Harris et al. (1996), Harris and Shonkwiler (1996) and Wensley and Stabler (1998) have introduced use of count-data techniques when data is truncated.

This paper expands on previous count data research to incorporate hurdle model procedures when excess zero observations are realized. Specific objectives are to review demand threshold analysis, discuss count data procedures when excess zeroes exist and review results of threshold demand results when excess zeroes are not addressed by 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

$$P = \alpha B^{\beta} \tag{1}$$

where P is the place's population, B is the number of businesses of a particular type within the place and α and β are parameters to be estimated. The nonlinear specification follows from CPT. In practice, the estimated equation is a double-log model. Given estimates of α and β, 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 are non-negative and integer which would suggest count data procedures.

Harris et al. (1996) and Harris and Shonkwiler (1996) applied count data procedures to estimate minimum demand thresholds at the county level. Wensley and Stabler (1998) employed count data procedures to estimate demand thresholds for rural Saskatchewan at the local or community level. In so doing, they highlighted a common observation that rural areas are characterized by lower demand thresholds and, therefore, higher frequency of business establishments relative to areas that are more proximate to urban centers, other things being equal.

However, when using county data procedures, overdispersion can be associated with a prevalence of zero data. Employing regime-splitting zero inflated negative binomial or hurdle negative binomial the prevalence of zeroes in the data can be addressed. These procedures correct for any heteroscedasticity associated with the count data model without accounting for the qualitative difference between zero and non-zero outcomes in the data generating process.

STATISTICAL MODELS

The alternative count-data specifications considered in this study are based on the Poisson distribution of random variable Y_i, with parameter λ_i:

$$h(y_i;\lambda_i) = e^{-\lambda_i}\lambda_i^{y_i} / y_i!, \quad y_i = 0,1,2,... \tag{2}$$

In the context of regression, the parameter λ_i is allowed to vary according to

$$E(Y_i) = \mathrm{var}(Y_i) = \lambda_i = \exp(x_i'\beta) \tag{3}$$

where x_i is a vector of explanatory variables for observation i and β is corresponding parameter vector.

Poisson Model

The first statistical model we consider is the Poisson model. Using (2), the sample likelihood function for an independent sample of *n* observations is

$$L(\beta, y) = \prod_{i=1}^{n} h(y_i; \lambda_i) \tag{4}$$

where $y = [y_1, y_2, ..., y_n]'$. Prediction and effects of explanatory variables are based on the conditional mean expression (3).

Single-Hurdle Model

Although the Poisson distribution admits zero values in the dependent variable, the sample used in this study contains a large proportion of zeros, which exceed what would typically be predicted by the Poisson model. To accommodate these excessive zeros, we construct hurdle count models, which are motivated by the Gaussian hurdle specifications of Cragg (1971) and Blundell and Meghir (1987). Consider latent variables d_i^* for the binary outcome and y_i^* for the level outcome. For the single-hurdle model, the observed value of the dependent variable y_i relates to these latent variables such that[1]

$$\begin{aligned} y_i &= y_i^* && \text{if } d_i^* > 0 \\ &= 0 && \text{otherwise.} \end{aligned} \tag{5}$$

Assume the binary outcome $d_i^* > 0$ is governed by a Gaussian structure, with probit probability

[1] General procedure for constructing single-hurdle count model is discussed in Mullahy (1986) and Gurmu (1998). Pohlmeier and Ulrich use single-hurdle count model in modeling health-care demand.

$$\Pr(d_i^* > 0) = \Phi(z_i'\alpha) = \int_{-\infty}^{z_i'\alpha} \frac{1}{\sqrt{2\pi}} e^{-w^2/2} dw \tag{6}$$

where z_i and α are conformable vectors of explanatory variables and parameters. In (6), $\Phi(\cdot)$ is the standard normal cumulative distribution function.[2] Then, the probability mass for observation i takes the form

$$\begin{array}{ll} \Pr(d_i^* \le 0) & \text{if } y_i = 0 \\ h(y_i; \lambda_i \mid y_i > 0) \Pr(d_i^* > 0) & \text{if } y_i > 0. \end{array} \tag{7}$$

Using (2), (5), (6), and (7), the sample likelihood for the single-hurdle model is

$$L(\alpha, \beta; y) = \prod_{y_i = 0} \left[1 - \Phi_i(z_i'\alpha)\right] \prod_{y_i > 0} \Phi_i(z_i'\alpha)\, h(y_i; \lambda_i) \Big/ \left[1 - h(0; \lambda_i)\right]. \tag{8}$$

In (8), $h(y_i; \lambda_i) \big/ \left[1 - h(0; \lambda_i)\right]$ is the probability mass of the truncated Poisson where $h(0; \lambda_i) = e^{-\lambda_i}$, using (2). This model corresponds to the continuous (Gaussian) single-hurdle specification of Cragg (1971, equations (7) and (9)). Because the likelihood function (8) is separable in α and β, estimation can be carried out by a probit estimation (for α) using the full sample, and a truncated regression (for β) using the truncated sample ($y_i > 0$). Prediction and effects of explanatory variables can be based on the probability, conditional mean, and unconditional mean of y_i, respectively:

$$P(y_i > 0) = \Phi(z_i'\alpha) \tag{9}$$

$$E(y_i \mid y_i > 0) = \lambda_i \big/ \left[1 - h(0; \lambda_i)\right] \tag{10}$$

---

[2] The first-hurdle probability can be composed with alternative distributions such as the Poisson, negative binomial (2), and exponential distributions. It is easy to show that use of a count distribution for the binary outcome amounts to an upper censored (at one) count regression, but we find that use of discrete probability distributions presents no advantage over the Gaussian distribution in modeling the binary outcomes. Mullahy (1986) points out that use of the exponential distribution, along with an exponential parameterization similar to that of (3), yields the familiar logistic probability for the first hurdle.

$$E(y_i) = \Phi(z_i'\alpha)\ \lambda_i \big/ \left[1 - h(0;\lambda_i)\right] \tag{11}$$

Population threshold estimates are calculated from the conditional mean (10) and unconditional mean (11), respectively. Further, elasticities with respect to explanatory variables can be derived by differentiating these expressions. These expressions also imply that the elasticities of probability (9) and conditional mean (10) with respect to a variable (say x_{ij}, a common elements of z_i and x_i) add up to the elasticity of unconditional mean (11).

Double-Hurdle Model

Another specification, common in the continuous case literature but rarely (if any) considered in count modeling, is the double-hurdle model.[3] The double-hurdle model features an additional censoring mechanism to generate zeros. Drawing on the continuous case of Blundell and Meghir (1987) and Cragg (1971, equations (5) and (6)), the probability of a zero observation is

$$\Pr(d_i^* \le 0) + \Pr(d_i^* > 0)\Pr(y_i^* \le 0) = 1 - \Pr(d_i^* \le 0)\Pr(y_i^* \le 0) \tag{12}$$

The probability of a positive observation is

$$\Pr(y_i^* > 0)\, h(y_i^*, \lambda_i \mid y_i^* > 0)\Pr(d_i^* > 0) \tag{13}$$

Using (2), (5), (12), and (13), the sample likelihood for the double-hurdle model is

$$L(\alpha,\beta; y) = \prod_{y_i=0}\left\{1 - \Phi(z_i'\alpha)\left[1 - h(0;\lambda_i)\right]\right\} \prod_{y_i>0}\Phi(z_i'\alpha)\, h(y_i;\lambda_i) \tag{14}$$

It is interesting to note that by reparameterizing $\gamma = -\alpha$ so that $\Phi(z_i'\gamma) = 1 - \Phi(z_i'\alpha)$, the likelihood function (14) corresponds to the zero-inflated Poisson specification of Lambert

[3] One exception appears to be Shonkwiler and Shaw (1997), who propose double-hurdle model count models for recreational demand modeling. We are not aware of any empirical work based on double-hurdle count models.

(1992), which was proposed to accommodate excessive zeros in the sample. As in the single-hurdle model, prediction (population threshold estimates) and effects of explanatory variables are based on the probability, conditional mean, and unconditional mean:

$$P(y_i > 0) = \Phi(z_i'\alpha)\left[1-h(0)\right] \tag{15}$$

$$E(y_i \,|\, y_i > 0) = \lambda_i \big/ \left[1-h(0)\right] \tag{16}$$

$$E(y_i) = \Phi(z_i'\alpha)\lambda_i = \Phi(-z_i'\gamma)\lambda_i \tag{17}$$

As the Poisson, single-hurdle, and double-hurdle models are not nested, selection among these competing specifications can be carried out by Vuong's (1989) nonnested specification test. In particular, let f and g be n-vectors containing the log-likelihoods of competing models. ι be an n-vector of ones, and define $d = f - g$. Then, Vuong's standard normal statistic (Vuong 1989, equation (5.6)) can be calculated as $z = \iota' d / [d'd - (\iota' d)^2 / n]^{1/2}$.

DATA AND SAMPLE

Data used in this study are compiled from two primary sources: the 1990 Census of Population and Unemployment Compensation Insurance files (ES202) maintained by the Wisconsin Department of Workforce Development. Because of the nature of the firm count data (i.e., ES202) smaller firms that do not meet the minimum requirement for reporting are lost to the analysis. Generally, these small firms, commonly referred to as “mom n’ pop” operations, have limited impacts on local markets, hence any bias from undercounting firms is assumed to be minimal. Data are for all 1747 municipalities in Wisconsin. In this study we investigate the grocery sector, focusing on sampling units with population between 100 and 5,000. A smaller proportion of observations with missing data for important variables are excluded. This leaves a final sample of 1,588 observations for analysis.

The dependent variable is the number of establishments (grocery stores). The explanatory variables include population, percentages of population over 65 and under 18, medium household income, proportion of commuters, proportion of urban population, and the proportion who live under poverty as defined in the 1990 Census. Sample statistics for all variables are presented in Table 1.

Table 2 presents the frequency distribution of the dependent variable (number of grocery stores). Among the sample, 1197 towns (75%) record no grocery stores. The large proportion of zero observations suggest that failure to accommodate these excessive zeros are likely to lead to unreliable results.

RESULTS

Parameter Estimates

The Poisson, single-hurdle, and double-hurdle models are estimated with maximum-likelihood method, based on the likelihood functions (4), (8), and (14), respectively. Results of Vuong's nonnested tests, presented in Table 3, suggest that the single-hurdle and double-hurdle models both outperform the Poisson model. Further, contrary to findings in much of the Gaussian hurdle literature, the double-hurdle model does not perform better than the single-hurdle model at the 10% level of significance, suggesting that zeros are governed entirely by the Gaussian process and the additional hurdle in the double-hurdle model is not an effective censoring mechanism.

Maximum-likelihood estimates of all models are presented in Table 4. At the 10% level of significance, all but one variables are significant in the Poisson model, whereas significance is more sparse in the single-hurdle and double-hurdle models. Population, percent of population over 65, and proportion living under poverty are significant in both the binary and level processes in the double-hurdle model, while only population is significant in both processes in the single-hurdle model. Differences between the single-hurdle and double-hurdle results appear quite notable. For instance, the percent of population under 18 and medium household income are significant in the binary process but not in the level process, while the results are opposite for the double-hurdle model, with the two variables affecting level but not binary.

One of the major purposes in estimating these econometric models is calculation of minimum demand thresholds. These thresholds are calculated based on the mean expression for the Poisson model, and conditional and unconditional means for the two hurdle models suggested above. The results are presented in Table 5. The Poisson model gives higher threshold at one establishment but tends to give more conservative threshold estimates at higher counts than the two hurdle models. The single-hurdle model also produces higher "unconditional" demand thresholds than the double-hurdle model.

Threshold estimates are often calculated from more simplified empirical estimates such as the ordinary least-squares estimates. This highlights the importance of using the hurdle models to estimate population thresholds. The hurdle models allow calculation of "conditional" demand thresholds, which can be more useful than the unconditional estimates. For instance, the conditional threshold estimate suggests the population required to support two establishments, conditional on the "fact" that one establishment is already existent. The conditional results, also presented in Table 5, suggest uniformly more conservative threshold estimates than their unconditional counterparts between the two hurdle models considered.

As the three models are parameterized differently, effects of explanatory variables can be evaluated further by calculating elasticities. The results, presented in Table 6, suggest that the effects of explanatory variables on the probability and level (number) of grocery stores vary, notably so in some cases, across models. For instance, the elasticities of conditional level with respect to all variables are insignificant, whereas the elasticities of probability are significant with respect to all but two variables (proportion who commute, proportion living in urban area). Thus, the explanatory variables affect the number of grocery stores through probability and not level. Elasticities for the double-hurdle model, on the other hand, suggest that these variables can affect both the probability and number of grocery stores. For instance, the elasticities of probability, conditional level and unconditional level (number) of grocery stores with respect to percent of population under 18 years old are all significant and positive, whereas the corresponding elasticities with respect to proportion who commute are all negative and significant. The Poisson model suggests very different elasticities. For instance, the results suggest that when population increases by 1 percent, all else equal, the number of grocery stores increases by 0.63 percent. The corresponding (unconditional) elasticities are both under 0.30. The elasticity with respect to proportion under poverty (–0.47) is also notably different from the corresponding elasticities in the single-hurdle model and double-hurdle model.

CONCLUDING REMARKS

Previous analysis of demand thresholds are often based on over-simplified statistical procedure such as the ordinary least squares. More recent studies use count-data regression models. Although the probability mass functions used in the count-data model admits zero values in the dependent variable, the excessive zeros in some of the samples often cannot be predicted by traditional count-data models. The hurdle count specifications considered in this study accommodate the excessive zeros by allowing for separate stochastic processes that generate the zero and positive counts and provide the flexibility in modeling count outcomes. Our findings suggest that failure to accommodate the excessive zeros may cause notably different threshold estimates. While the hurdle specifications considered in this study are based on the basic Poisson distribution, further study might consider generalization of these framework to more generalized specifications such as the single-hurdle and double-hurdle negative binomial models.

REFERENCES

Berry, B. and W. Garrison. "A Note on Central Place Theory and the Range of Good." *Economic Geography*, 34 (1958a): 304-311.

Berry, B. and W. Garrison. "Recent Developments in Central Place Theory." *Proceedings of the Regional Science Association*. 4(1958b): 107-121.

Blundell, R.W., and C. Meghir. "Bivariate Alternatives to the Univariate Tobit Model." *Journal of Econometrics* 34(1987): 179–200.

Chrisman, J.J. "Population Change and Its Effect on the Retail Sector: An Exploratory Study," International Small Business Journal, 3 (1985):26-46.

Cragg, J.G. "Some Statistical Models for Limited Dependent Variables With Applications to the Demand for Durable Goods." *Econometrica* 39(1971):829–844.

Deller, S.C. and D.L. Chicoine. "Economic Diversification and the Rural Economy: Evidence from Consumer Behavior," *Regional Science Perspectives* 19 (1989):41-55.

Foust, B. and A. DeSouza. *The Wisconsin Urban System: Functional Size, Trade Area and Nesting*. Unpublished manuscript, Department of Geography, University of Wisconsin, Eau Claire, 1977.

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.

Gurmu, S. "Generalized Hurdle Count Data Regression Models." *Economics Letters* 58 (1998): 263–268.

Hansen, N. *Rural Poverty and the Urban Crisis*. Bloomington, Indiana: Indiana University, 1970.

Harris, Thomas R. and J. Scott Shonkwiler. "Application of Count Data Procedures to Estimate Thresholds for Rural Commercial Sectors." *Proceedings of the Western Agricultural Economics Association* 1993, pp. 410-415.

Harris, Thomas R., Kalyan Chakraborty, Lijuan Xiao and Rangesan Narayanan. "Application of Count Data Procedures to Estimate Thresholds for Rural Commercial Sectors." *Review of Regional Studies* 26(1996): 75-88.

Haworth, J. and P. Vincent. "The Stochastic Disturbance Specification and Its Implications for Log-Linear Regression." *Environment and Planning A*, II(1979):781-790.

Henderson, D. "Retail Sales and Consumer Expenditure Functions." *Journal of Agricultural Research*, 42(1990):27-34.

Kieschnick, M. *Taxes and Growth: Business Incentives and Economic Development*. Washington, D.C.: Corporation for Enterprise Development, 1981.

King, L.J. *Central Place Theory*. London: SAGE Publications, 1984.

Lambert, D. "Zero-Inflated Poisson Regression, With An Application to Defects in Manufacturing." *Technometrics* 34 (1) (1992): 1–14.

McNamara, K. and W. Krissel. "Assessing Local Industrial Development Potential." *Journal of the Community Development Society*, 24(1993):61-71.

Mullahy, J. "Specification and Testing of Some Modified Count Data Models." *Journal of Econometrics* 33 (1986): 341–365.

Murray, J. and J. Harris. A Regional Economic Analysis of the Turtle Mountain Indian Reservation: Determining Potential for Commercial Development. Prepared for the Federal Reserve Bank of Minneapolis, Minneapolis, Minnesota, 1978.

Parr, J. and K. Denike. "Theoretical Problems in Central Place Analysis." *Economic Geography*, 47(1970): 568-586.

Pohlmeier, W., and V. Ulrich. "An Econometric Model of the Two-Part Decisionmaking Process in the Demand for Health Care." *Journal of Human Resources* 30 (1995): 339–361.

Salyards, D.M. and K.R. Leitner. "Market Threshold Estimates: A Tool for Business Consulting in Minnesota." *American Journal of Small Business*, 6(1981): 26-32.

Shaffer, R. *Community Economics: Economic Structure and Change in Small Communities*. Ames, Iowa: The Iowa State Press, 1989.

Shonkwiler, J.S., and W.D. Shaw. "Hurdle Count-Data Models in Recreation Demand Analysis." *Journal of Agricultural and Resource Economics* 21 (1996): 210–219.

Smith, D. "Native American Economic Development: A Modern Approach." *The Review of Regional Studies*, 24(1994):87-102.

Smith, T. and W. Fox. "Economic Development Programs for States in the 1990's." *Economic Review*, July/August (1990):25-35.

Tweeten, L. and G. Brinkman. *Micropolitan Development: Theory and Practice of Greater Rural Economic Development*. Ames, Iowa: The Iowa State Press, 1976.

Vuong, Q.H. "Likelihood Ratio Tests for Model Selection and Nonnested Hypotheses." *Econometrica* 57(1989): 307–333.

Wensley, Mitch R.D. and Jack C. Stabler. "Demand Threshold Estimation for Business Activities in Rural Saskatchewan." *Journal of Regional Science* 38(1998):155-177.

Table 1. Sample Statistics (Sample Size = 1588)

Variable	Mean	Std. dev.
Number of grocery stores	0.33	0.67
Population (thousands)	1.11	0.90
Percent of population over 65	14.70	6.14
Percent of population under 18	28.10	4.57
Medium household income (thousands)	27.33	8.41
Proportion of commuters	18.56	9.05
Proportion living in urban area	1.97	12.95
Proportion living under poverty	6.20	4.55

SOURCE: 1990 Census of Population

Table 2. Frequency Distribution of Number of Establishments: Groceries

Number of Establishments	Frequency	Number of Establishments	Frequency
0	1197	3	21
1	287	4	5
2	77	5	1

SOURCE: Wisconsin Department of Workforce Development, Unemployment Compensation Insurance Files (ES202)

Table 3. Voung's Standard Normal Statistics for Nonnested Tests of Hypotheses

Model (log-likelihood)	Single-Hurdle	Double-Hurdle
Double-Hurdle (–949.06)		–
Single-Hurdle (–940.33)	–	1.22
Poisson (–973.57)	3.96	3.60

Note: Test statistic is distributed as standard normal.

Table 4. Maximum-Likelihood Estimates

	Poisson	Single-Hurdle Model		Double-Hurdle Model	
		Binary	Level	Binary	Level
Constant	–3.817‡	–3.997‡	–0.872	–0.480	–4.244‡
	(0.752)	(0.644)	(1.753)	(2.050)	(1.097)
Population	0.570‡	0.533‡	0.455‡	0.493‡	0.448‡
	(0.036)	(0.048)	(0.062)	(0.165)	(0.050)
Percent of population over 65	0.113‡	0.125‡	0.039	0.173‡	0.092‡
	(0.012)	(0.012)	(0.026)	(0.043)	(0.018)
Percent of population under 18	0.074‡	0.086‡	–0.030	–0.031	0.104‡
	(0.015)	(0.013)	(0.038)	(0.047)	(0.026)
Medium household income	–0.049†	–0.040‡	0.005	–0.034	–0.031‡
	(0.011)	(0.009)	(0.025)	(0.030)	(0.014)
Proportion who commute	–0.012†	–0.007	–0.036‡	0.017	–0.020‡
	(0.007)	(0.005)	(0.017)	(0.019)	(0.008)

Table 4 continued

Proportion living in urban area	–0.002	0.003	–0.138	0.014	–0.002
	(0.004)	(0.003)	(1.082)	(0.029)	(0.006)
Proportion under poverty	–0.075‡	–0.074‡	0.015	–0.093‡	–0.044‡
	(0.015)	(0.012)	(0.033)	(0.039)	(0.022)
Log-likelihood	–973.57	–940.33		–949.06	

Note: Asymptotic standard errors in parentheses. Daggers ‡ and † denote significance at the 5% and 10% levels, respectively.

Table 5. Population Thresholds

Number of Establishments	Poisson	Single-Hurdle		Double-Hurdle	
		Unconditional	Conditional	Unconditional	Conditional
1	3,729	3,582	–	3,383	–
2	4,945	5,350	4,790	4,848	3,603
3	5,657	6,386	6,047	5,741	4,878
4	6,161	7,068	6,770	6,381	5,613
5	6,553	7,575	7,290	6,878	6,140
6	6,873	7,981	7,701	7,285	6,557

Table 6. Elasticities of Conditional Mean with respect to Continuous Variables

	Poisson	Single-Hurdle Model			Double-Hurdle Model		
		Prob.	Cond. Level	Uncond. Level	Prob.	Cond. Level	Uncond. Level
Population	0.63‡	0.21‡	0.06	0.27‡	0.16	0.09‡	0.24†
	(0.04)	(0.02)	(0.13)	(0.13)	(0.14)	(0.01)	(0.14)
Percent of population 65+	1.66‡	0.65‡	0.07	0.72‡	–0.07	0.24‡	0.17
	(0.17)	(0.06)	(0.16)	(0.17)	(0.49)	(0.04)	(0.51)
Percent of population 18–	2.08‡	0.84‡	–0.11	0.74‡	2.81‡	0.52‡	3.33‡
	(0.43)	(0.13)	(0.25)	(0.29)	(1.06)	(0.15)	(1.18)
Medium household income	–1.35‡	–0.39‡	0.02	–0.37‡	–0.27	–0.15‡	–0.42
	(0.30)	(0.09)	(0.09)	(0.13)	(0.62)	(0.07)	(0.67)
Proportion who commute	–0.23†	–0.05	–0.08	–0.13	–0.46†	–0.07‡	–0.52‡
	(0.13)	(0.03)	(0.18)	(0.18)	(0.25)	(0.03)	(0.27)
Proportion living in urban area	–0.00	0.00	–0.03	–0.03	–0.02	–0.00	–0.02
	(0.01)	(0.00)	(0.20)	(0.20)	(0.03)	(0.00)	(0.03)
Proportion under poverty	–0.47‡	–0.16‡	0.01	–0.15‡	0.04	–0.05	–0.01‡
	(0.09)	(0.03)	(0.04)	(0.04)	(0.21)	(0.02)	(0.23)

Note: Asymptotic standard errors in parentheses. Daggers ‡ and † denote significance at the 5% and 10% levels, respectively.