New Estimation Strategies for Demand Threshold Models in the Southern United States

Junpyo Hong
Graduate Research Assistant
Department of Agricultural Economics and Agribusiness
Louisiana State University
101 Ag. Administration Bldg
Baton Rouge, LA 70803
Jhong2@lsu.edu

J. Matthew Fannin
Assistant Professor
Department of Agricultural Economics and Agribusiness
Louisiana State University
101 Ag. Administration Bldg
Baton Rouge, LA 70803
mfannin@agcenter.lsu.edu

Selected Paper
Annual Meetings of the Southern Agricultural Economics Association
February 4th-6th, 2007
Mobile, AL

Abstract

This paper estimates demand threshold models using both first generation log-log models and second generation Tobit models to zip code areas in the Southern US. Results of own-place demographic and economic variables were consistent with previous studies but impacts of neighboring zip codes contrasted previous studies.

JEL Code(s): R120
Key Words: Demand Threshold Analysis, Central Place Theory

Introduction

The purpose of this paper is to present the application of new estimation strategies when modeling demand thresholds in the southern US. In particular, this paper estimates more recent Tobit models that include demographic as well as competing establishment effects on identifying the optimal number of retail sector establishments a specific geographic area can support. Further this paper uses the zip code as the geographic unit of analysis as compared to traditional county or city models.

This paper is organized as follows. First, a literature review is presented on historical modeling strategies for estimating demand thresholds. Next, the regression procedure (Tobit model) is presented including a detailed description of the data. Regression results are then obtained and compared to results from the more traditional log-log model. Finally implications of the research are discussed.

Literature Review

Demand threshold analysis (DTA) has been studied for several decades. A common definition for DTA is the minimum market size required to support a particular establishment. (Berry and Garrison 1958). The conceptual framework underlying DTA can be found in Central Place Theory (Christaller 1966). In Central Place Theory, individual consumers are spread across a homogeneous plane facing positive transportation costs and attempt to minimize the delivered price of the goods they purchase. Individual firms attempt to maximize demand. As a result, levels of central places are created that support specific types of businesses based on the minimum efficient scale for that business to operate. Smaller central places have businesses that need only a small number of customers to provide the output levels needed to meet minimum efficient scale. Larger order central places include these lower minimum efficient scale businesses as well as larger businesses that require selling larger levels of output in order to meet minimum efficient scale.

DTA has often been used for planning purposes by communities. Estimation methods of DTA for these purposes have involved such procedures as identifying the total number establishments of a particular retailer in the state and then dividing the level by that state's population (Deller and Ryan 1996). This approach represents an average threshold, or the average number of residents required to support a single establishment of a specific retail sector in a state.

The average threshold approach serves as a good rule of thumb for identifying threshold levels for the first establishment, but does not take into account increasing returns to scale for many retail establishments. That is, many retail establishments may expand the output level of their particular business to meet a growing population in a particular place. This expansion may continue until the establishment produces output at some point beyond minimum efficient scale where a second establishment can enter the market, compete, and maintain their existence alongside the first establishment.

To account for this characteristic in establishment growth, Berry and Garrison assumed a non-linear relationship between the number of business establishments and the populations required to support them

(1) $P = a\,(B)^{B}$

where P is population, B is the number of businesses, and a and b are parameters to be estimated.

This model can predict not just the population threshold needed to support one business establishment but also the population needed two support two, five or ten business establishments. These type of studies proliferated throughout the 1960s – 1980s as an effective single univariate model for planning (cf. Salyards and Leitner 1981).

More sophisticated modeling procedures developed in the 1990s to address the more technical details of the data. Models developed by Shonkwiler and Harris (1996) and Harris and Shonkwiler (1997) recognized the non-negative count nature of the dependent variable (establishment counts) and applied count data estimators such as the Tobit model. Wensley and Stabler (1998) and Henderson, Kelly and Taylor (2000) evaluated incorporation of proximity to urban areas and agglomeration economies in estimating demand thresholds. The most recent literature in this area has evaluated how neighboring (adjacent) places impact demand thresholds. Both Mushinski and Weiler (2002) and Thilmany et al (2005) incorporated the number of establishments and population of the adjacent place in estimating a model of own place demand thresholds.

Many of these modern second generation models have included exogenous variables that were not incorporated in the parsimonious first generation models. The purpose of this research is to compare the performance in prediction between second generation models based on count data estimators against the new first generation demand threshold models that now incorporate the new exogenous variables from second generation models. We estimate these models using a previously unanalyzed geographic unit of analysis in demand threshold modeling – zip codes.

Regression Model

Estimates of population threshold levels are first derived by ordinary least squares regression analysis similar to Salyards and Keitner (1981). In this analysis, the level of population to support a specific number of establishments is to be obtained.

Following the findings of several threshold studies indicating a monotonic curvilinear relationship between the number of establishments and the population of a place (e.g. Beckmann 1958), this study adopts log-log regression model as follows:

(2) $\ln EST_p = a_0 + b' \ln x + e$

where a_0 is a constant term, x is a vector of explanatory variables, and $e \sim iid(0,1)$.
Explanatory variables include the number of establishments of neighboring areas(EST_n), the total population of the place(POP_p), the total population of the neighboring areas (POP_n), and per capita income of the place($PCINC_p$). Note that the total population of neighboring areas is also included in the place equations, because neighboring areas might be a source of demand which is separate from the competitive effect of neighboring establishments captured by Thilmany et. al. (2005).

As a comparison to the revised first generation model above, a Tobit model is also estimated. Due to the count data characteristics of the dependent variable, the Tobit model regression employing maximum likelihood estimation can produce unbiased and consistent estimators (Amemiya 1973).

As posited by Mushinski and Weiler (2002), and Thilmany et al. (2005), a relationship between the observed number of establishments in a place (EST_p) and the observed number of establishments in neighboring areas (EST_n) can be captured through the own place equation. The equation for the place is described as follows:

(3) $EST_p = a_p EST_n + b' x + e$ if RHS > 0

$= 0$ if RHS = 0

where x is a vector of explanatory variables and $e \sim iid(0,1)$. Explanatory variables include the square of population of the place (POP_p^2) as well as those used in the log-log regression.

Previous studies analyzing demand threshold level in each establishment strongly suggest that population approximates the level of household demand in the place. Additionally, the square of the population is included in the regression equation for nonlinear specification and recognizes that declining rate of increase that demand has on total establishments.[1] Per capita income is included as a demand variable capturing the buying power of consumers in a place (Mushinski and Weiler 2002). Hence, we would expect that total population and per capita income to have a positive effect on the number of establishments in a place and the square of population, number of neighboring establishments and the population of the neighboring place to have a negative effect on own place establishments.

Data

The southern states on which this study was focused included Alabama, Arkansas, Florida, Georgia, Kansas, Kentucky, Louisiana, Mississippi, North Carolina, Oklahoma, South Carolina, Tennessee, and Texas. There are 8,709 total five-digit zip code areas within southern states.[2] The areas are classified according to whether they are MSA or non-MSA, resulting in 3,260 MSA and 5,449 non-MSA zip codes. To identify contiguous zip codes, a procedure was applied to the zip code polygon file from ESRI GIS dataset.

[1] While Mushinski and Weiler (2002) did not choose a quadratic specification due to a strong relationship between high number of population and "zero" number of establishments, this study did not identify this strong relationship. In fact, at the higher level of ninety percentile of total population(28060), there were only two hundred thirty eight zip-code areas reported to have "zero" number of establishments. This number is less than 0.05 percentage of total number of zip-code areas which have "zero" number of establishments and is regarded as a small portion enough to be neglected. Therefore, the square of total population is included for nonlinear specification.

[2] Numbers in parentheses indicate missing zip-code areas in each state: Alabama (71), Arkansas (65), Florida(136), Georgia(96), Kansas (50), Kentucky (109), Louisiana(82), Mississippi(74), North Carolina(112), Oklahoma (100), South Carolina(55), Tennessee(43), and Texas (259).

The total number of establishments for each of 13 North American Industry Classification System (NAICS) codes[3] was obtained from County Business Patterns (US Census Bureau 2005). It is notable that approximately half of the retail sectors do not obtain a single establishment in more than half of zip-code areas. Descriptive statistics are summarized in Table 1.[4]

Results and Discussion

Table 2 shows the results of log-log regression. Most estimates are significant at 1 % significance level and adjusted R-squares indicate that the specified log-log model was well fitted. It is notable that the signs of EST_n are positive, which were opposite expectations. A positive sign implies that more establishments in neighboring areas induce more establishments in the place -- for example, a Furniture & Home Furnishing store business in one zip code is complimentary to another Furniture & Home Furnishing store business in a contiguous zip code.

As expected, signs of POP_p are positive, indicating greater total demand generates additional establishments. However, unlike the expectation of Thilmany et al that the population of neighboring areas might be an additional source of demand for own place establishments, the log-log regression model does not confirm this relationship. All signs of POP_n in place equations are negative and indicate that increased population in neighboring zip codes reduces own zip code establishment demand.

[3] Thirteen three-digit *NAICS* retail codes include Motor vehicle & parts dealers (*NAICS: 441*), Furniture & home furnishing stores (*NAICS: 442*), Electronics & appliance stores (*NAICS: 443*), Building material & garden equipment & supplies dealers (*NAICS: 444*), Food & beverage stores (*NAICS: 445*), Health & personal care stores (*NAICS: 446*), Gasoline stations (*NAICS: 447*), Clothing & clothing accessories stores (*NAICS: 448*), Sporting goods, hobby, book & music stores (*NAICS: 451*), General merchandise stores (*NAICS: 452*), Miscellaneous store retailers (*NACIS: 453*), Nonstore retailers (*NAICS: 454*), Food services & drinking places (*NAICS: 722*)

[4] Both were downloaded from "Census 2000 Summary File 3 (SF 3) - Sample Data, Detailed Tables" at US Census Bureau and the web site is as below; http://factfinder.census.gov/servlet/DTSubjectShowTablesServlet?_ts=185192911453

In Table 3, the population required to support various numbers of establishments is presented. Specially, the demand thresholds focused on the minimum population required to support a single retail establishment are presented in bold. Every retail sector except Motor vehicle & parts dealers (*441*) exhibits higher demand threshold in MSA regions than in a non-MSA regions. Reversely, results show that greater population is required to support higher numbers of establishments in non-MSA zip codes than in MSA zip codes.

Following log-log regression, Tobit regression was performed for each of thirteen retail sectors and the results of the regression are presented in Table 4. Tests using the Breusch-Pagan LM statistic showed no problems with heteroskedesticity in the place equation. Additionally, the null hypotheses that all explanatory variables are simultaneously equal to zero were rejected by a Wald test at 1 % significance level in both non-MSA and MSA of each equation. Most explanatory variables are significant at 1 % significance level in both non-MSA and MSA. Even though there are a few insignificant variables[5], the values of coefficients of those are trivial enough to be neglected.

It is remarkable that EST_n in most of the retail sectors shows a positive sign, which was opposite to the results of Mushinski and Weiler (2002) and Thilmany et. al. (2005). While negative signs imply that the businesses are competitors, positive signs of EST_n in place equations imply that those businesses are complimentary to each other. Only one retail sector, Clothing & clothing accessories stores (*448*) in MSA, had a negative sign for EST_n, but the parameter estimate was insignificant.

---

[5] Insignificant variables include EST_n for Clothing & clothing accessories stores (*448*) in MSA, POP_n for Clothing & clothing accessories stores (*448*) in MSA, POP_n for General merchandise stores (*452*) in non-MSA, and $PCINC_p$ for Building material & garden equipment & supplies dealers (*444*) in both non-MSA and MSA.

The signs of own zip code exogenous variables were as expected. The positive signs of total population and negative signs of the square of total population in the own zip code equations assure our supposition that increases in population create increased demands for retail businesses and that this relationship occurs at a decreasing rate.

However, total population of neighboring areas has a negative effect on the number of establishments of the place in most retail sectors for both non-MSA and MSA. More people in a neighboring zip code reduce the number of retail establishment in one's own zip code. This result is inconsistent with our assumption and the result of Thilmany et al that neighboring areas might be a source of demand in the place.

Table 5 represents how many places have exactly, under, or over estimated establishments in each retail sector.[6] Under-estimation implies that there exist more establishments in a zip code than what the model predicts and over-estimation vice versa. In general, the log-log model overestimated establishment counts for all retail sectors in both MSA and non-MSA zip codes and the Tobit model over-estimated non-MSA establishment counts and under-estimated MSA establishment counts.

In particular, except for Clothing and accessory stores (448) in non-MSA zip codes, the log-log model had a higher percentage of over-estimated establishment counts than the Tobit

---

[6] For Tobit model, the values were computed by obtaining fitted values against the regression line. If the fitted values are between -0.5 and 0.5, those areas were recorded as exact estimation. If the fitted values are less than -0.5, those areas were recorded as over-estimation. For log-log model, the values were calculated by solving the Equation (2) with respect to EST_p, e.g., $\hat{EST}_p = \exp\left(\hat{a}_0 + \hat{b}_p' \ln \bar{x}_p\right)$, where ^ indicates estimates and – indicates mean values of each explanatory variable. After obtaining the nearest integer values of $\hat{EST}_p$ and the actual number of establishments of a place, we compared which one is greater than the other. For example, if $\hat{EST}_p$ is greater than the actual number of establishments in a place, those areas were recorded over-estimation, and vice versa.

model. In contrast, in MSA zip codes, the Tobit model under-estimated establishments a majority of the time in 10 of the 13 retail sectors evaluated. In terms of an exact match between actual and predicted establishment counts, neither model performed well. The Tobit model percentage of exact matches exceeded 10 percent in only three retail sectors and the log-log model only exceeded a 10 percent match in one retail category.

Implications and Conclusion

This paper attempts to compare the performance of second generation demand threshold models with their first generation counterparts adjusted by the inclusion of additional exogenous variables found significant in second generation models. These results were applied to a new geographic unit of analysis for demand threshold analysis, zip codes.

Regression results for own-place exogenous variables were consistent with results from models using city and county geographic units. However, neighboring establishment counts using zip code data were opposite expected signs from previous studies. Neither model generated a high probability of success in exactly predicting the number of establishments in each zip code. The log-log model over-estimated establishments for both non-MSA and MSA zip codes. The Tobit model over-estimated non-MSA establishments but typically under-estimated MSA zip codes establishments.

This research is its initial stages and is very much a work in progress. A number of planned analyses are expected to be performed to expand and refine the current analysis. First, the endogeneity of the neighboring establishment counts in the Tobit model that are ignored in this analysis will be accounted for through a simultaneous model following more closely the work of Thilmany et al. Second, a similar analysis will be conducted using counties as the geographic unit of analysis for southern states in order to compare the predictive power of the model using alternative geographic definitions. A spatial econometric model is also being considered to correct potential autocorrelation between neighboring geographic units in the regressions. Such improvements should help to increase predictive power as well as obtain a better understanding of how sensitive DTA is to the choice of geographic unit.

References

Amemiya, T. "Regression Analysis When The Dependent Variable is Truncated Normal", *Econometrica*, 41(1973): 997-1016.

Beckmann, M. "City Hierarchies and the Distribution of City Size", *Economic Development and Cultural Change*. 6(1958): 243-248.

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

Christaller, W. *Central Places of Southern Germany*. Englewood Cliffs, NJ: Prentice-Hall. 1966.

Deller, S.C. and W.F. Ryan. "Retail and Service Demand Thresholds for Wisconsin." Staff Paper 96.1. Center for Community Economic Development, University of Wisconsin. 1996.

Greene, W. H. *Econometric Analysis*, 5th Edition. Pearson Education, 2003.

Harris, T.R. and S.J. Shonkwiler. "Interdependence of Retail Businesses." *Growth and Change*, 27(1997): 520-533.

Henderson, J.W., T.M. Kelly, and B.A. Taylor. "The Impact of Agglomeration Economies on Estimated Demand Thresholds: An Extension of Wensley and Stabler." *Journal of Regional Science*, 40(2000): 719-733.

Mushinski, D. and S. Weiler. "A Note on the Geographic Interdependencies of Retail Market Areas", *Journal of Regional Science*, 42(2002): 75-86.

Salyards, D. M. and K. R. Leitner. “Market Threshold Estimates: A Tool For Business Consulting in Minnesota”, *American Journal of Small Business*, VI, 2(1981): 26-32.

Shonkwiler, J.S, and T.R. Harris. “Rural Retail Business Thresholds and Interdependencies.” *Journal of Regional Science*, 36(1996): 617-630..

Thilmany, D., N. McKenney, D. Mushinski, and S. Weiler. “Beggar-thy-Neighbor Economic Development: A Note on the Effect of Geographic Interdependencies in Rural Retail Markets”, *The Annals of Regional Science*, 39(2005): 593 – 605.

U.S. Census Bureau. *2002-2003 Zip Code Business Patterns*. On DVD. U.S. Department of Commerce. October 2005.

Veall, M. R., and K. F. Zimmermann. “Goodness of Fit Measures in the Tobit Model”, *Oxford Bulletin Economics and Statistics* 56, 4(1994): 485 – 499.

Wensley, M.D.R., and J.C.Stabler. “Demand-Threshold Estimation for Business Activities in Rural Saskatchewan”, *Journal of Regional Science* 38, 1(1998): 155-177.

Table 1. Descriptive Statistics Used in Demand Threshold Models.

		Mean	Median	Min	Max	Std. Dev.
Motor vehicle & parts dealers(*441*)	non-MSA	3.335	0	0	82	6.442
	MSA	7.311	3	0	64	9.481
Furniture & home furnishing stores(*442*)	non-MSA	1.428	0	0	53	3.518
	MSA	3.931	1	0	50	6.114
Electronics & appliance stores (*443*)	non-MSA	0.930	0	0	38	2.452
	MSA	2.494	1	0	40	3.974
Building material & garden equipment & supplies dealers (*444*)	non-MSA	2.287	1	0	39	3.948
	MSA	4.690	3	0	40	5.626
Food & beverage stores (*445)*	non-MSA	3.583	0	0	81	5.935
	MSA	6.975	4	0	58	7.311
Health & personal care stores (*446*)	non-MSA	1.862	0	0	67	4.030
	MSA	4.622	2	0	40	6.248
Gasoline stations (*447*)	non-MSA	3.807	2	0	46	5.734
	MSA	7.210	5	0	51	7.513
Clothing & clothing accessories stores (*448*)	non-MSA	3.108	0	0	185	10.133
	MSA	8.834	1	0	192	18.358
Sporting goods, hobby, book & music stores (*451*)	non-MSA	1.029	0	0	32	2.921
	MSA	3.425	1	0	50	5.701
General merchandise stores (*452*)	non-MSA	1.356	0	0	17	2.375
	MSA	2.397	1	0	20	3.222
Miscellaneous store retailers (*453*)	non-MSA	2.776	0	0	53	5.810
	MSA	7.266	3	0	135	9.867
Nonstore retailers (*454*)	non-MSA	0.884	0	0	24	1.814
	MSA	2.098	1	0	23	2.667
Food services & drinking places (*722*)	non-MSA	8.870	2	0	191	17.650
	MSA	23.882	12	0	283	28.701
Total Population	non-MSA	6611.164	2451	0	76146	9935.618
	MSA	15428.060	12040.5	0	113935	13226.900
Per Capita Income (dollars)	non-MSA	16010.830	15143.0	0	283189	6559.990
	MSA	19684.300	18022.5	0	85883	7693.715

Table 2. Log-log Regression Results.[7]

		ln(EST_n)	ln (POP_p)	ln(POP_n)	ln($PCINC_p$)	constant	Adjusted R^2
Motor vehicle & parts dealers(*441*)	non-MSA	-0.020	0.543***	-0.085***	-0.183***	-0.850***	0.581
	MSA	0.168***	0.700***	-0.167***	-0.400***	0.197***	0.529
Furniture & home furnishing stores(*442*)	non-MSA	0.090***	0.309***	-0.103***	0.016	-1.292***	0.425
	MSA	0.306***	0.435***	-0.256***	0.057*	-1.572***	0.423
Electronics & appliance stores (*443*)	non-MSA	0.078***	0.234***	-0.074***	0.031*	-1.206***	0.378
	MSA	0.179***	0.359***	-0.130***	0.035	-1.874***	0.372
Building material & garden equipment & supplies dealers (*444*)	non-MSA	0.056***	0.433***	-0.141***	-0.096***	-0.536***	0.540
	MSA	0.207***	0.533***	-0.160***	-0.154***	-1.072***	0.483
Food & beverage stores (*445*)	non-MSA	0.185***	0.502***	-0.172***	-0.110***	-0.771***	0.643
	MSA	0.349***	0.668***	-0.328***	-0.256***	0.357***	0.626
Health & personal care stores (*446*)	non-MSA	0.096***	0.381***	-0.117***	-0.013	-1.373***	0.526
	MSA	0.245***	0.547***	-0.234***	-0.029	-1.747***	0.494
Gasoline stations (*447*)	non-MSA	0.052**	0.541***	-0.111***	-0.158***	-0.822***	0.660
	MSA	0.125***	0.662***	-0.133***	-0.335***	-0.176***	0.602
Clothing & clothing accessories stores (*448*)	non-MSA	0.033***	0.433***	-0.075***	0.078***	-2.932***	0.442
	MSA	0.170***	0.605***	-0.173***	0.083*	-3.795***	0.378
Sporting goods, hobby, book & music stores (*451*)	non-MSA	0.121***	0.228***	-0.097***	0.103***	-1.659***	0.371
	MSA	0.268***	0.385***	-0.217***	0.082***	-1.738***	0.388
General merchandise stores (*452*)	non-MSA	-0.033**	0.347***	-0.081***	-0.157***	0.113***	0.512
	MSA	-0.039*	0.440***	-0.041*	-0.245***	-0.353***	0.412
Miscellaneous store retailers (*453*)	non-MSA	0.097***	0.462***	-0.146***	0.063***	-2.340***	0.532
	MSA	0.332***	0.627***	-0.325***	0.022	-2.060***	0.514
Nonstore retailers (454)	non-MSA	0.096***	0.204***	-0.064***	0.022	-0.997***	0.392
	MSA	0.273***	0.279***	-0.144***	0.031	-1.172***	0.410
Food services & drinking places (*722*)	non-MSA	0.277***	0.746***	-0.366***	-0.006	-1.823***	0.665
	MSA	0.527***	0.903***	-0.536***	-0.134***	-1.009***	0.647

*, **, and *** indicates that coefficient is significant at 10 %, 5%, and 1% significance level, respectively.

[7] EST_n = number of establishments in neighboring areas,
POP_p = total population in a place,
POP_n = total population in neighboring areas,
$PCINC_p$ = per capital income in a place(dollar).

Table 3. Population Required to Support Establishments Based on Log-Log Model.

		pop(1)	pop(2)	pop(3)	pop(4)	pop(5)	pop(10)	pop(20)
Motor vehicle & parts dealers(441)	non-MSA	**630**	2255	4754	8071	12169	43565	155960
	MSA	**350**	944	1684	2540	3495	9409	25334
Furniture & home furnishing stores(442)	non-MSA	**721**	6822	25392	64519	132988	1257723	1.19E+07
	MSA	**939**	4621	11738	22743	37988	186962	920155
Electronics & appliance stores (443)	non-MSA	**753**	14498	81765	278995	722845	1.39E+07	2.68E+08
	MSA	**1200**	8287	25658	57211	106563	735697	5079156
Building material & garden equipment & supplies dealers (444)	non-MSA	**671**	3329	8493	16505	27634	137001	679220
	MSA	**1077**	3952	8453	14497	22029	80812	296449
Food & beverage stores (445)	non-MSA	**467**	1856	4160	7374	11497	45679	181485
	MSA	**1006**	2839	5210	8014	11193	31592	89167
Health & personal care stores (446)	non-MSA	**710**	4375	12675	26962	48417	298387	1838894
	MSA	**1246**	4423	9279	15697	23600	83754	297233
Gasoline stations (*447*)	non-MSA	**501**	1803	3814	6489	9801	35279	126988
	MSA	**950**	2706	4993	7709	10798	30758	87612
Clothing & clothing accessories stores (448)	non-MSA	**768**	3799	9681	18798	31454	155642	770147
	MSA	**1174**	3692	7218	11614	16796	52832	166187
Sporting goods, hobby, book & music stores (451)	non-MSA	**625**	12997	76702	270274	717940	1.49E+07	3.10E+08
	MSA	**894**	5398	15460	32614	58193	351570	2123983
General merchandise stores (452)	non-MSA	**48**	352	1134	2601	4951	36580	270251
	MSA	**1987**	9610	24161	46473	77190	373302	1805341
Miscellaneous store retailers (453)	non-MSA	**675**	3023	7265	13533	21926	98156	439406
	MSA	**980**	2962	5656	8951	12778	38613	116681
Nonstore retailers (454)	non-MSA	**570**	17021	124128	508283	2E+06	4.53E+07	1.35E+09
	MSA	**682**	8182	34984	98082	218207	2615858	3.14E+07
Food services & drinking places (722)	non-MSA	**517**	1310	2255	3315	4471	11316	28642
	MSA	**654**	1408	2206	3033	3883	8365	18018

* Number in parentheses indicates the number of establishments.

Table 4. Tobit Regression Result of Place Equation.

		EST_n	POP_p	POP_p^2	POP_n	$PCINC_p$	constant	R^2
Motor vehicle & parts dealers(*441*)	non-MSA	0.08347***	0.00134***	-1.42E-08***	-6.6E-05***	-0.00008***	-4.565***	0.473
	MSA	0.04956***	0.00099***	-8.37E-09***	-2E-05***	-0.00014***	-3.746***	0.412
Furniture & home furnishing stores(*442*)	non-MSA	0.11935***	0.00085***	-9.73E-09***	-3.9E-05***	0.00006***	-7.360***	0.400
	MSA	0.09021***	0.00067***	-6.21E-09***	-1.9E-05***	0.00018***	-10.170***	0.392
Electronics & appliance stores (*443*)	non-MSA	0.10982***	0.00065***	-7.09E-09***	-2.9E-05***	0.00006***	-6.631***	0.387
	MSA	0.06275***	0.00049***	-4.32E-09***	-7.54E-06***	0.00011***	-7.974***	0.390
Building material & garden equipment & supplies dealers (*444*)	non-MSA	0.09235***	0.00082***	-9.17E-09***	-4.7E-05***	-0.00001	-2.752***	0.458
	MSA	0.07426***	0.00054***	-4.61E-09***	-1.5E-05***	0.00000	-3.439***	0.396
Food & beverage stores (*445*)	non-MSA	0.12019***	0.00077***	-4.87E-09***	-6.1E-05***	-0.00003***	-1.629***	0.524
	MSA	0.11216***	0.00063***	-3.76E-09***	-4.6E-05***	-0.00003**	-1.606***	0.498
Health & personal care stores (*446*)	non-MSA	0.11078***	0.00077***	-7.43E-09***	-4.4E-05***	0.00005***	-5.156***	0.457
	MSA	0.05336***	0.00068***	-5.87E-09***	-1.8E-05***	0.00012***	-7.402***	0.421
Gasoline stations (*447*)	non-MSA	0.07332***	0.00100***	-9.45E-09***	-5.9E-05***	-0.00004***	-1.102***	0.526
	MSA	0.05724***	0.00066***	-3.67E-09***	-3.2E-05***	-0.00008***	0.050***	0.476
Clothing & clothing accessories stores (*448*)	non-MSA	0.03551***	0.00222***	-2.38E-08***	-5E-05***	0.00034***	-25.595***	0.365
	MSA	-0.00161	0.00203***	-1.93E-08***	3.24E-06	0.00064***	-36.867***	0.319
Sporting goods, hobby, book & music stores (*451*)	non-MSA	0.10135***	0.00074***	-7.58E-09***	-3.3E-05***	0.00011***	-8.776***	0.379
	MSA	0.06825***	0.00067***	-6.17E-09***	-1.4E-05***	0.00018***	-11.177***	0.374
General merchandise stores (*452*)	non-MSA	0.04680***	0.00060***	-6.89E-09***	-2.6E-05***	-0.00004***	-1.998***	0.425
	MSA	-0.03256***	0.00043***	-3.84E-09***	-4.26E-06**	-0.00005***	-2.207***	0.390
Miscellaneous store retailers (*453*)	non-MSA	0.10528***	0.00115***	-1.17E-08***	-6.9E-05***	0.00013***	-7.680***	0.471
	MSA	0.08479***	0.00097***	-8.75E-09***	-4.1E-05***	0.00022***	-10.376***	0.402
Nonstore retailers (*454*)	non-MSA	0.11077***	0.00039***	-4.55E-09***	-1.6E-05***	0.00003***	-3.328***	0.369
	MSA	0.11225***	0.00025***	-2.22E-09***	-7.63E-06***	0.00005***	-3.485***	0.398
Food services & drinking places (*722*)	non-MSA	0.11916***	0.00233***	-1.64E-08***	-0.00019***	0.00025***	-10.320***	0.543
	MSA	0.09769***	0.00210***	-1.17E-08***	-0.00014***	0.00035***	-13.529***	0.456

R^2 represents Aldrich and Nelson measure of goodness-of-fit(Veall et.al. 1994).

*, **, and *** indicates that coefficient is significant at 10 %, 5%, and 1% significance level, respectively.

Table 5. Distribution of Exact, Under, or Over Estimation of the Number of Establishments in Place Equations.

		Exact-estimation		Under-estimation		Over-estimation	
		Tobit (%)	log-log (%)	Tobit (%)	log-log (%)	Tobit (%)	log-log (%)
Motor vehicle & parts dealers(*441*)	non-MSA	192 (3.52)	169 (3.10)	1643 (30.15)	1187 (21.78)	3614 (66.32)	4093(75.11)
	MSA	141 (4.33)	54 (1.66)	**2063 (63.28)**	575 (17.64)	1056 (32.39)	2631(80.71)
Furniture & home furnishing stores(*442*)	non-MSA	171 (3.14)	306 (5.63)	1056 (19.42)	869 (15.98)	4212 (77.44)	4274(78.58)
	MSA	143 (4.39)	196 (6.01)	**1654 (50.74)**	1077 (33.04)	1463 (44.88)	1987(60.95)
Electronics & appliance stores (*443*)	non-MSA	158 (2.90)	246 (4.51)	895 (16.43)	634 (11.64)	4396 (80.68)	4569(83.85)
	MSA	171 (5.25)	281 (8.62)	1495 (45.86)	1011 (31.01)	1594 (48.90)	1968(60.37)
Building material & garden equipment & supplies dealers (*444*)	non-MSA	471 (8.64)	310 (5.69)	1793 (32.91)	1113 (20.43)	3185 (58.45)	4026(73.89)
	MSA	263 (8.07)	219 (6.72)	**2127 (65.25)**	1191 (36.53)	870 (26.69)	1850(56.75)
Food & beverage stores (*445*)	non-MSA	837 (15.36)	295 (5.41)	**2441 (44.80)**	1280 (23.49)	2171 (39.84)	3874(71.10)
	MSA	262 (8.04)	159 (4.88)	**2520 (77.30)**	1296 (39.75)	478 (14.66)	1850(56.75)
Health & personal care stores (*446*)	non-MSA	220 (4.04)	386 (7.08)	1310 (24.04)	1118 (20.52)	3919 (71.92)	3945(72.40)
	MSA	145 (4.45)	169 (5.18)	**1871 (57.39)**	1123 (34.45)	1244 (38.16)	1968(60.37)
Gasoline stations (*447*)	non-MSA	922 (16.92)	284 (5.21)	**2650 (48.63)**	1417 (26.00)	1877 (34.45)	3748(68.78)
	MSA	324 (9.94)	162 (4.97)	**2687 (82.42)**	1308 (40.12)	249 (7.64)	1790(54.91)
Clothing & clothing accessories stores (*448*)	non-MSA	53 (0.97)	138 (2.53)	968 (17.76)	923 (16.94)	4428 (81.26)	4388(80.53)
	MSA	49 (1.50)	89 (2.73)	1507 (46.23)	1063 (32.61)	1704 (52.27)	2108(64.66)
Sporting goods, hobby, book & music stores (*451*)	non-MSA	127 (2.33)	208 (3.82)	849 (15.58)	629 (11.54)	4473 (82.09)	4612(84.64)
	MSA	130 (3.99)	191 (5.86)	1535 (47.09)	928 (28.47)	1595 (48.93)	2141(65.67)
General merchandise stores (*452*)	non-MSA	391 (7.18)	121 (2.22)	1375 (25.23)	266 (4.88)	3683 (67.59)	5062(92.90)
	MSA	250 (7.67)	379 (11.63)	**1663 (51.01)**	1107 (33.96)	1347 (41.32)	1774(54.42)
Miscellaneous store retailers (*453*)	non-MSA	160 (2.94)	219 (4.02)	1452 (26.65)	1160 (21.29)	3837 (70.42)	4070(74.69)
	MSA	131 (4.02)	117 (3.59)	**2010 (61.66)**	1185 (36.35)	1119 (34.33)	1958(60.06)
Nonstore retailers (*454*)	non-MSA	378 (6.94)	394 (7.23)	1086 (19.93)	627 (11.51)	3985 (73.13)	4428(81.26)
	MSA	381 (11.69)	395 (12.12)	**1748 (53.62)**	1024 (31.41)	1131 (34.69)	1841(56.47)
Food services & drinking places (*722*)	non-MSA	219 (4.02)	98 (1.80)	2128 (39.05)	1419 (26.04)	3102 (56.93)	3932(72.16)
	MSA	61 (1.87)	39 (1.20)	**2504 (76.81)**	1401 (42.98)	695 (21.32)	1820(55.83)

Bold indicates higher percentage of under-estimation than over-estimation.