Ann Reg Sci (2005) 39: 593–605
DOI: 10.1007/s00168-005-0229-x

The Annals of
Regional Science

Beggar-thy-neighbor economic development: A note on the effect of geographic interdependencies in rural retail markets

Dawn Thilmany[1], Nicole McKenney[2], David Mushinski[2], Stephan Weiler[3]

[1]Department of Agricultural and Resource Economics, Colorado State University, Fort Collins, CO, 80523-1172, USA (e-mail: thilmany@lamar.colostate.edu)
[2]Department of Economics, Colorado State University, Fort Collins, CO, 80523-1172
[3]Center for the study of Rural America, Federal Reserve Bank Kansas City, MO 64198, USA (e-mail: Stephan.Weiler@kc.frb.org)

Received: March 2003 / Accepted: September 2004

Abstract. Mushinski and Weiler (2002) updated a vein of rural economic geography literature by estimating empirically the importance for retail development of geographic interdependencies between places and their neighboring areas. This note extends interpretation of their empirical results by considering the influence of neighboring areas and establishments on retail thresholds in a place, and the policy and economic development implications of their results. This note discusses the nature of spatial competition and considers how retail establishments might act as regional base industries, absorbing shopping flows from outlying residents in a fashion similar to traditional export industries. Moreover, it signals that understanding geographic interdependencies is important for economic development planning, and suggests there may be merit in more regional coordination of retail firm recruitment in relatively small and/or isolated rural areas.

1. Introduction

As towns and cities grow, local retail establishments appear, then increase in number. Central place theory posits a distinct relationship between characteristics of local areas and retail development in those areas. Economists, planners, and regional scientists have explored this process of regional growth, in part, to inform economic development-oriented public policies. While the size and relative status of an area have been identified as key determinants of the number and type of industries present in the area, regional economists have also explicitly identified the relationship between the number of establishments in an area and a variety of socioeconomic and demographic variables, including the population in an area. Mushinski and Weiler (2002) extended the analysis by seeking to measure geographic interrelationships between a central place and its neighboring market areas.

Mushinski and Weiler (2002) updated a literature concerning geographic market interdependencies by using recent census data and econometric techniques to ascertain the importance of such interdependencies for rural retail development. The article did not explore various economic and practical interpretations of their empirical results. For example, they did not consider the economic development implications of those results, or the relationship between neighboring areas and the population thresholds at which retail establishments appear in a place. This note delves into these unexplored aspects of their empirical research.

2. Literature review

Central place theory explains the hierarchical organization of different levels of agglomeration, or "central places," from hamlets to towns to metropolitan areas (Fik 1988a; Mulligan 1984). Among other implications, the theory predicts a direct and positive relationship between population and the number and type of establishments present in a community. The estimation of population thresholds for particular industries extends central place theory by identifying the specific critical population mass for various types of activities in a settlement (see, e.g. Mulligan, Wallace and Plane 1985; Deller and Harris 1993). From a policy perspective, population thresholds provide economic development planners with an idea of the types of retail industries they may expect to find in their communities or may they have the best potential of recruiting to an area. Deller and Ryan (1996) note that trade area analysis and demand threshold analysis are two approaches that state extension programs are taking to support retail-based economic development efforts. They argue that these methods are helpful in assessing development strategies that may assist potential firm owners with business planning.

Currently, several issues are prominent in retail firm economic geography. Gibson, Albrecht and Evans (2003) posit that the transition to a knowledge-based economy requires the field to rethink regional development efforts. Retail creates a direct economic impact, and may also enhance the "quality of life" that knowledge-based firms seek out to attract employees. They also suggest regional economic welfare can be improved through retail sector growth by meeting local demand (that would otherwise leak to other areas) and securing export income (for those that may come to shop in the region). Fik's (1988a and b) retail work shows that space and location matter in price competition among retail food markets, supporting the theory that isolation may be a significant factor in determining the demand schedule firms face.

Wensley and Stabler (1998) found a higher frequency of retail business establishments in more isolated rural areas relative to areas adjacent to urban centers. They argue that geographic isolation is a significant factor in threshold determination because transportation/travel costs make the relative costs of rural firms more competitive with urban areas. Henderson, Kelly and Taylor (2000) found the reverse, showing that agglomeration economies of urban-proximate markets may allow them to be more competitive. These studies suggest that geographic interactions between population centers and establishments may themselves vary by sector.

Mushinski and Weiler (2002) fit into a literature which includes the works of Thill (1996), Deller and Harris (1993) and Shonkwiler and Harris (1996), which extended central place theory by exploring relationships between markets in proximate geographic locations. Harris et al. (1996) explored commercial retail activity levels by extending previous research with new empirical methods. They found both population and per capita income positively affected the number of retail firms, with population change generally being more influential and counties adjacent to metro areas supporting fewer firms. Mushinski and Weiler (2002) built upon these efforts by examining interdependencies between a place and neighboring areas.

3. Discussion of Regression Results

Mushinski and Weiler's (2002) article involved econometrically modelling and estimating geographic interdependencies. We summarize the model, which is developed fully in their article. They posited a simultaneous relationship between the observed number of establishments in a place (EST_p) and the observed number of establishments in neighboring areas (EST_n). The expected number of establishments in a place and the expected number of establishments in neighboring areas were latent variables which could be positive or negative. Establishments appeared in an area only when the latent variable became positive. The relationships observed by the econometrician were captured by two equations. The equation for the place (the *Place equation*) may be represented as follows:

$$\begin{aligned} y_p &= \lambda_p y_n + \beta_p' x_p + u_p \quad && \text{if RHS } > 0 \\ &= 0 && \text{if RHS } \leq 0 \end{aligned} \tag{1}$$

where y_p was EST_p, y_n was EST_n, x_p was a vector of exogenous regressors which were demand and supply factors expected to affect y_p and u_p was a disturbance. The second equation, for the neighboring area (the *Neighbor equation*), may be represented as

$$\begin{aligned} y_n &= \lambda_n y_p + \beta_n' x_n + u_n \quad && \text{if RHS } > 0 \\ &= 0 && \text{if RHS } \leq 0 \end{aligned} \tag{2}$$

where x_{ni} was a vector of exogenous regressors expected to affect y_n and u_n was a disturbance. The equations were estimated in a simultaneous equation Tobit framework using maximum likelihood techniques.

Mushinski and Weiler (2002) focused on non-metropolitan communities in the Intermountain West region: including Arizona, Colorado, Idaho, Montana, Nevada, New Mexico, Utah and Wyoming. Those areas contain a large number of geographically isolated communities distinct from MSAs and separated by significant amounts of sparsely-populated hinterland. The communities were non-metropolitan in that they resided in counties that did not contain MSAs.

The data used in the analysis was drawn from the 1992 Census of Retail Trade and the 1992 City and County Data Book. The reader is directed to Mushinski and Weiler (2002) for further details about the data. The data set contained 183 rural communities within 129 counties, with each county containing at least one incorporated town of at least 2500 inhabitants. Variables for place were drawn from town data. Neighboring areas were defined as the value of that variable in the county in which the town resided minus the value of that variable in the town. Tables 1 and 2, which contain the regression results, identify the retail sectors included in the analysis.

Table 1. Place equation estimates for nine industries[a]

Industry Variable	Building Stores	Merchandise Stores	Food Stores	Auto Dealers	Gas Stations	Apparel Stores	Furniture Stores	Eating & Drinking	Drug Stores
Constant	9.833*	5.898*	17.24*	19.51*	22.14*	2.98	9.84*	35.85 *	3.11*
	(1.361)	(1.082)	(2.79)	(2.34)	(2.73)	(6.41)	(2.56)	(7.66)	(0.71)
$Est_{neighbor}$	−0.019	−0.109*	−0.112	−0.118*	−0.284*	−0.132*	−0.082*	−0.163**	−0.052
	(0.042)	(0.047)	(0.071)	(0.051)	(0.074)	(0.060)	(0.037)	(0.094)	(0.060)
$(1/POP_{place})$	−28.68*	−14.7*	−40.05*	−44.99*	−39.37*	−59.79*	−52.45*	−126.25*	−11.46*
	(3.210)	(2.024)	(5.62)	(4.97)	(5.12)	(8.58)	(6.40)	(20.13)	(1.52)
Density	0.001	0.017	0.049	0.021	−0.022	0.064	0.038	0.078	0.017
(÷100)	(0.037)	(0.018)	(0.033)	(0.038)	(0.029)	(0.060)	(0.05)	(0.128)	(0.017)
Per Capita	0.103	−0.046	−0.005	−0.127	−0.081	1.168*	0.44*	1.67*	0.060**
Income	(0.69)	(0.052)	(0.095)	(0.108)	(0.12)	(0.376)	(0.12)	(0.50)	(0.035)
($1000s)									
Close to MSA	−1.626*	−0.047	−0.143	−0.780	−0.223	−1.48	−1.34	0.32	0.098
(Yes = 1)	(0.587)	(0.336)	(0.97)	(0.875)	(0.921)	(1.60)	(1.02)	(3.95)	(0.308)
$1/POP_{neighbor}$	−1.462	−1.071	−10.91	−5.73	−16.66*	−7.30	−4.16	−56.08**	1.27
	(2.670)	(1.533)	(8.08)	(4.77)	(7.37)	(6.18)	(4.37)	(32.31)	(1.60)
Rho	0.039	0.292	0.48*	0.199	0.68*	0.233	0.020	0.626*	0.17
	(0.198)	(0.194)	(0.22)	(0.181)	(0.14)	(0.184)	(0.145)	(0.214)	(0.27)
Log-Likelihood	−1119	−797	−1312	−1168	−1216	−1154	−1106	−1758	−740
Restricted Log-Likelihood[b]	−1208	−894	−1405	−1276	−1321	−1261	−1228	−1852	−823
Pseudo R^2 [c]	0.49	0.51	0.5	0.54	0.53	0.54	0.57	0.51	0.48

[a] Standard errors are reported in parentheses below coefficient estimates. Quasi-maximum likelihood standard errors are reported (see: Mittelhammer et al. 2003:251).

[b] Assumes that all slope coefficients equal zero.

[c] This is the Aldrich and Nelson measure of goodness-of-fit (see: Veall et al. (1994)). It equals $[2 \bullet (L(\beta) - L(0))]/[2 \bullet (L(\beta) - L(0)) + N]$, where $L(\beta)$ is the log-likelihood for the regression, $L(0)$ is the restricted log-likelihood and N is the number of observations.

* 95% Wald confidence region does not include zero. Thus, a zero null hypothesis with a two-sided test would be rejected at a 5% level of significance.

** 90% Wald confidence region does not include zero. Thus, a zero null hypothesis with a two-sided test would be rejected at a 10% level of significance.

Table 2. Neighbor equation estimates for nine industries[a]

Industry Variable	Building Stores	Merchandise Stores	Food Stores	Auto Dealers	Gas Stations	Apparel Stores	Furniture Stores	Eating & Drinking	Drug Stores
Constant	8.256*	9.24*	24.43*	15.46*	19.62*	−3.56	−1.35	16.65	1.22
	(3.756)	(2.24)	(5.37)	(6.28)	(4.7)	(5.17)	(5.20)	(16.11)	(1.80)
Est_{place}	−0.101	−0.76*	−0.46	−0.28	−0.38	−0.58**	−0.32	−0.39	−0.70**
	(0.396)	(0.33)	(0.38)	(0.26)	(0.29)	(0.31)	(0.28)	(0.38)	(0.41)
$(1/POP_{neighbor})$	−93.39*	−53.48*	−110.8*	−107.7*	−81.4*	−265.8*	−177.7*	−279.7*	−46.6*
	(20.55)	(10.09)	(17.2)	(22.2)	(13.6)	(47.2)	(33.48)	(55.19)	(14.14)
Density	20.17*	3.77	36.9*	44.9*	29.28*	39.54*	46.27*	144.1*	8.72**
(÷100)	(9.78)	(4.15)	(14.3)	(14.32)	(7.72)	(16.8)	(13.26)	(53.13)	(5.05)
Per Capita Income	0.425	−0.02	0.16	0.12	0.04	2.15*	1.33*	3.43*	0.33*
($1000s)	(0.247)	(0.12)	(0.24)	(0.35)	(0.21)	(0.41)	(0.33)	(1.08)	(0.13)
Close to MSA	2.98	2.25*	3.71	1.82	3.76	1.57	3.54	17.6	1.56
(Yes = 1)	(2.55)	(1.08)	(3.11)	(2.58)	(2.62)	(3.83)	(3.05)	(11.2)	(1.00)

[a] Standard errors are reported in parentheses below coefficient estimates. Quasi-maximum likelihood standard errors are reported (see: Mittelhammer et al. 2003:251).

* 95% Wald confidence region does not include zero. Thus, a zero null hypothesis with a two-sided test would be rejected at a 5% level of significance.

** 90% Wald confidence region does not include zero. Thus, a zero null hypothesis with a two-sided test would be rejected at a 10% level of significance.

We now turn to the regressors used in the analysis, which are contained in Tables 1 and 2. The presence of EST_N in the Place equation and EST_P in the Neighbor equation captures possible spatial competition interdependencies between markets and implies simultaneity between the equations. A dummy variable for adjacency to an MSA captured any remaining spatial competition from larger metropolitan markets (see: Deller and Harris 1993). The inverse of the population in the place ($POP_P{}^{-1}$) and the inverse of the population in neighboring areas ($POP_N{}^{-1}$) in the Place and Neighbor equations, respectively, permitted a non-linear relationship between the dependent variable and population in both equations. The $POP_N{}^{-1}$ variable was included in the Place equation because neighboring areas might be a source of demand which is separate from the competitive effect of neighboring areas captured by EST_N. Population densities in the place and in neighboring areas captured the density of demand, with the assumption that less dense demand implies more travel costs and, thus, lower effective demand. Separate per capita income measures for the place and neighboring areas were not available. Thus, the same county per capita income measure was included in each equation.

Place and Neighbor equations were estimated for each of nine retail sectors. Table 1 contains coefficient estimates for the Place equations, while Table 2 identifies coefficient estimates for the Neighbor equations. Mushinski and Weiler (2002) organized their Place equation results in terms of a taxonomy which depended on the statistical significance of the (Est_N) and ($POP_N{}^{-1}$) variables in the Place equation. While organizing our discussion of Place equation results around that taxonomy, we discuss those results in greater depth.

For the first group of industries, which included Drug Stores, Food Stores and Building Stores, both variables were statistically insignificant. It is notable that MSA adjacency is statistically significant in the Building Supply Stores equation. An example firm, Home Depot stores, average 130,000 square feet and 40–50,000 product lines (Value Line 2001), and big-box retailers like these are driven by the cost structures of the industry (Munroe, 2001). Finally, the statistically significant positive correlation coefficient in the Food Stores equation suggests that when the number of establishments in the place is above its expected value, the number of establishments in the neighboring area is also above its expected value. One potential explanation is that seasonal tourism or intercounty traffic and commuting patterns may draw outside population to the county, affecting both areas within a county similarly.

For a second group of outlets (Merchandise Stores, Auto Dealers, Apparel Stores and Furniture Stores) only EST_n was statistically significant in the Place equation. The statistical significance of this variable implies supply-side interdependence between a place and its neighboring areas. This group of industries comprises stores selling more durable products, with less frequent visits and/or purchases relative to the other sectors. It should also be noted that many establishments in these sectors are not locally owned, so that locational decisions and allocations are made at a regional or national level. For instance, car dealers are generally guaranteed a certain geographic or population market size they will service under a franchise agreement with the manufacturer. Yet, the choice of place for durable good establishments may lead to inter-jurisdiction competition or recruitment given the relatively large tax potential for that sector.

Both EST_n and $POP_n{}^{-1}$ were statistically significant in the Place equation for the third group of industries (Gas Stations and Eating & Drinking). The statistical significance of both variables implies both supply- and demand-side interdependencies. This group features industries with localized service/consumption, more frequent purchases, and a more mobile clientele. In short, convenience plays a large role in consumer demand for these products. More gas stations and restaurant/bars in outlying areas reduce the need for going "into town" for these services, but more people in the surrounding county also increases the likelihood of stopping through a place for precisely these same needs. You may even see a "trading" of clientele among the areas to support one another and increase the diversity of offerings in the case of eating/drinking establishments.

A new focus for this note is interpretation of Neighbor equation estimates, summarized in Table 2. Neighboring areas are predominantly rural hinterlands with lower population densities than a place. Once again, the estimation results support central place theory; the population variable has the correct sign and is statistically significant in all equations. The most striking new finding is the virtual absence of geographical interdependence between neighbor establishments and place. The EST_p variable in the Neighbor equation was negative and statistically significant in only three equations, two of which exhibited spatial competition effects in the Place equation (Merchandise and Apparel) while the other exhibited no effect among its Place establishments (Drug Stores).

This result would suggest that establishments in neighboring areas depend primarily on local demand conditions through population. The statistical significance of the density variable in the Neighbor equations supports this view. Adjacency to metropolitan areas has a positive effect on the number of establishments in the Merchandise and Apparel sectors. There is anecdotal evidence that some rural areas support the development of outlet mall centers that contain merchandise, apparel and furniture stores, with developers simultaneously taking advantage of lower tax rates and a sizable proximate demand agglomeration.

The higher incidence of statistical significance for the density variable in the Neighbor equations may also suggest that travel times (which are likely to be negatively related to density) are a more important determinant of firm establishment in outlying county areas than regional competition. In fact, density is statistically significant almost as often as population. This finding also implies that the lower population densities inherent in regions with large areas of public lands in many Western counties may inhibit retail sector growth.

The similarity of results in the Place and the Neighbor equations with respect to per capita income suggests that this variable has the same impact across geography. Per capita income was statistically significant for Drug Stores, Apparel Stores, Furniture Stores, and Eating and Drinking Places in both equations. Consumers of these goods/services generally have preferences for variety that increase with income. Wealthier areas tend to feature a greater number of restaurant varieties, as well as broader ranges of clothing and furniture stores. The greater sensitivity of the number of establishments in these industries to per capita income indicates that regional development planners should focus on attracting these industries only if they believe that the income of their populace will be high enough to sustain new firms. Alternatively, it may suggest that these sectors are an important prerequisite for recruiting knowledge-based workers who choose location based on quality of life (see: Gibson, Albrecht, and Evans' 2003).

4. Impact of population neighboring areas on population thresholds in the place

The regression results in Mushinski and Weiler (2002) permit evaluation of the relationship between exogenous and endogenous variables in their regressions. Because we are interested in the impact of neighboring areas on a place, we focus on the relationship between the population in neighboring areas and population thresholds in the place. The population threshold in a place is the population in the place at which the expected number of establishments in an industry in the place equals one (see: Wensley and Stabler 1998). Intuitively, we expect a tradeoff between a larger neighboring population and a smaller place population for a given initial establishment and, thus, a negative relationship between population thresholds and and neighbour population levels. Figures 1 and 2 identify the relationship between population thresholds and population levels in neighboring areas for eight of the industries in the analysis.[1] Because calculation of $E(y_p)$ is somewhat complicated, we identify the equation for $E(y_p)$ in the Appendix.

Analysis of the relationships between population thresholds in the place and the population in neighboring areas indicated that the industries can be placed, generally, into one of two groups.[2] One group of industries, represented in Figure 1, is characterized by a negative relationship between the population threshold in the place and population in neighboring areas until population approaches between 5,000 and 10,000 people with no impact from further population increases. By contrast, the industries in Fig. 2 show population thresholds and population in neighboring areas become positively related at a similar population interval (5,000–10,000), even with negative relationships at lower population levels (similar to Fig. 1).

The initial negative relationship in both Figs. 1 and 2 confirms the broader market area hypotheses underlying this research note. Thresholds in a place for a first establishment decline with a larger neighboring customer base (the first firm locates with a relatively smaller population), as these stores appear to consider the full customer base from the combined place and neighboring market areas. Thus, first-mover establishments seem to examine both core and peripheral market areas to assess potential business viability and can be assumed to serve as basic industries for the place, drawing in customers and spending flows from the neighboring areas.

Figure 2 establishments' ensuing shift to a slightly positive relationship is particularly intriguing, yet can be interpreted through their contrasts with Fig. 1 establishments. Figure 1 sectors, such as Gas Stations and Food Stores, tend to be relatively homogeneous in scale and scope in smaller rural communities, with outlying populations only being influential in the more

[1] A population threshold for an industry was calculated by fixing the population in neighboring areas at some level and then identifying the population in the place at which $E(y_p)$ goes from less than one to one or more. We calculated such thresholds for a variety of population levels in neighboring areas for that industry. In calculating all thresholds, we held all other quantitative exogenous variables constant at their mean levels and the exogenous dummy variable constant at its median level.

[2] Drug Stores did not fall into either of the two groups, which may be partially due to the positive coefficient on the $1/POP_{neighbor}$ variable in the Place Equation for Drug Stores.

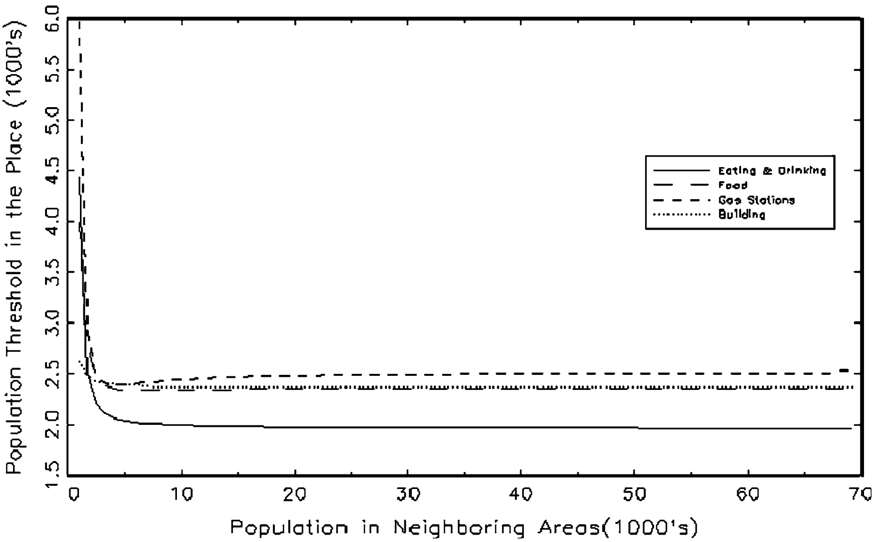

Fig. 1. Relationship between population thresholds and population in neighboring areas

likely initial choice to locate in a smaller place if it is by a larger neighboring population. Thereafter, larger outlying places have little effect on place establishment decisions, as these outlying places themselves are then likely to attract their own outlets.

In contrast, the first regional outlet for Fig. 2 sectors, such as furniture stores and auto dealers, could be expected to vary considerably in size. For example, small-scale furniture and clothing stores initially see the same

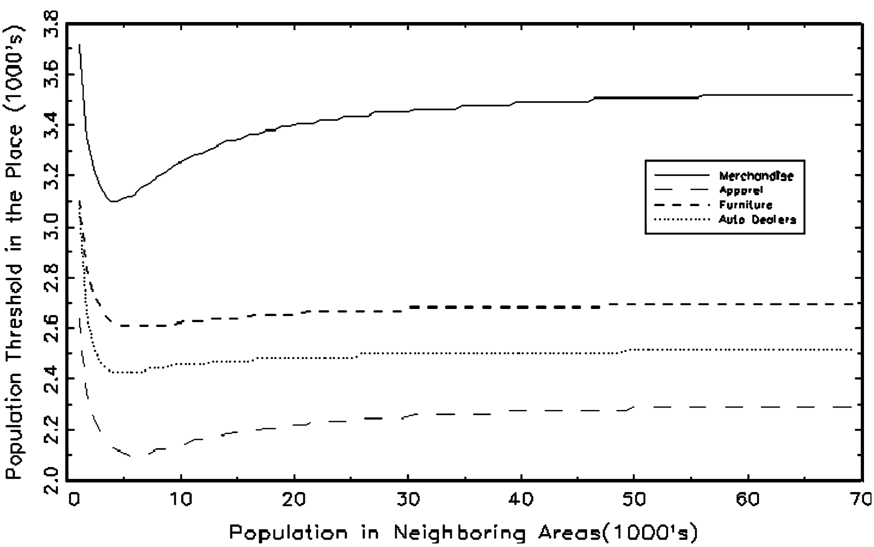

Fig. 2. Relationship between population thresholds and population in neighboring areas

tradeoff as above between a smaller place and a larger outlying population. However, after the noted minimum point for this relationship, Munroe's work suggests that threshold benchmarks for both place and neighbor populations might well increase as regional managers consider big-box retailing models (due to consumers desire for more diverse product lines or cost structure). This finding underscores the importance of viewing retail outlets in an interdependent market context, as both firm location and customer shopping decisions shape the regional flow of retail sales, income, tax revenue, and economic activity.

5. Implications for economic development strategies and policy

Retail establishments are generally considered to be non-basic local industries, which do not generate new regional income. But, retail sectors can play potentially significant export roles in creating net flows of new basic income from population located in a broader region, or, may indirectly improve perceived quality of life, thereby assisting in the attraction of higher income industries (Gibson et al. 2003). The interregional dynamics highlighted in this note also suggest a potential for politics to affect economic outcomes through the recruitment and retention of retail establishments.

Interregional battles for tax-producing shops, especially along urban fringes, have become increasingly fierce. Tax policy itself can become part of such skirmishes, which underlines the utility of Mushinski and Weiler's use of county lines as effective political *and* economic boundaries for retail placement. While the distributional effect among neighboring localities can easily result in potential zero-sum competition, game-theory suggests that interregional cooperation can lead to Pareto superior solutions. Our results may serve to better inform the nature and implementation of cooperative strategies among regions, or in the public sector, and economic development policies and programs that facilitate such cooperation at the supra-county level (Weiler et al. 2001).

The importance of the density findings in analysis of neighboring areas may be of concern in less dense Western areas that draw people by quality of life, even though many perceive that environmental attributes alone attract in-migrants (Power and Barrett 2001). Still, the presence of at least a base retail sector may be of some importance to those who are accustomed to convenient access to retail trade and service activity. Thus, retail sector development may complement other economic development strategies, including tourism, extended visitation by those with retirement status, and entrepreneurial firms that can choose location based on their quality of life preferences (Gibson et al. 2003).

6. Conclusions

Understanding the process of development in a locality requires an understanding of necessary threshold demand conditions for businesses alongside the related importance of spatial competition. This note explored the development and policy implications of Mushinski and Weiler's (2002) work. The implications of proximate population and firms are especially important for particular retail sectors and Western rural areas, given the relative isolation of the latter and the market size requirements of the former. There is diversity among sector results, signaling differing levels of importance with respect to population-based demand and interspatial competition variables. Therefore, different retail sectors must be considered unique.

In addition to highlighting subtle differences in retail subsectors, the general comments also highlight Blumenthal's (1955) seminal view of market area orientation. The present work focuses on the early stages of retail development at low establishment count levels in small towns. Blumenthal's prediction that such firms would be unusually likely to look at markets beyond their immediate areas are confirmed by the results, underscoring the importance of geographic interdependencies in rural retail analysis.

There are limitations to this research. First, it is well known that political boundaries are not a good proxy for the true geographical sense of place, but county level data are the most readily available for analysis and also reflect some political economy (e.g., tax) distinctions. Also, this analysis did not control for demographic variables that may affect retail location strategies (eg., stage of life and ethnicity), infrastructure and commuting patterns (that may facilitate or hinder cross-jursidiction buying patterns) and the effects of the significant increases in Internet-based consumer purchases.

References

Blumenthal H (1955) The economic base of the metropolis. Journal of the American Institute of Planners 21: 114–132

Deller SC, Ryan WF (1996) Retail and service demand thresholds for Wisconsin. Staff Paper 96–1. University of Wisconsin, Center for Community Economic Development

Deller SC, Harris TR (1993) Estimation of minimum market thresholds for rural commercial sectors using stochastic frontier estimators. Regional Science Perspectives 23: 3–17

Fik TJ (1988a) Spatial competition and price reporting in retail food markets. Economic Geography 64: 29–44

Fik TJ (1988b) Hierarchical interaction: the modeling of a competing central place system. Annals of Regional Science 22: 48–69

Gibson LJ, Albrecht B, Evans B (2003) Is retail trade a focus for real economic development in the knowledge-based economy? Applied Research in Economic Development 1: 44–55

Harris TR, Chakraborty K, Xiao L, Narayanan R (1996) Application of count data procedures to estimate thresholds for rural commercial sectors. The Review of Regional Studies 26: 75–88

Henderson JW, Kelly TM, Taylor BA (2000) The impact of agglomeration economies on estimated demand thresholds: An extension of Wensley and Stabler. Journal of Regional Science 40: 719–733

Maddala GS (1983) Limited-Dependent and qualitative variables in econometrics. Cambridge University Press, Cambridge, UK

Mittelhammer RC, Judge GC, Miller DJ (2000) Econometric foundations. Cambridge University Press, Cambridge, England

Mulligan GF (1984) Agglomeration and central place theory: A review of the literature. International Regional Science Review 9(1): 1–42

Mulligan GF, Wallace ML, Plane DA (1985) A general model for estimating the number of tertiary establishments in communities: An Arizona perspective. The Social Science Journal 22: 77–93

Munroe S (2001) Retail structural dynamics and the forces behind big-box retailing. Annals of Regional Science 35: 357–373

Mushinski D, Weiler S (2002) A note on the geographical interdependencies of retail market areas. Journal of Regional Science 42(1): 75–86

Perloff JM, Sickles RC (1987) Union wage, hours, and earnings differentials in the construction industry. Journal of Labor Economics 5(2): 174–210
Power TM, Barrett R (2001) Post-cowboy economics: Pay and prosperity in the New West. Island Press.
Shonkwiler JS, Harris TR (1996) Rural retail business thresholds and interdependencies. Journal of Regional Science 36(4): 617–630
Thill JC (1986) A note on multipurpose and multistop shopping, sales, and market areas of firms. Journal of Regional Science 26: 775–778
U.S. Department of Commerce (1994) 1992 census of retail trade: Utah, Nevada, Wyoming, Colorado, New Mexico, Idaho, Montana and Arizona geographic area series. Bureau of the Census, Washington, DC
U.S. Department of Commerce (1996) Land area, population, and density for states and counties and places: 1990. Bureau of the Census, Washington, DC
Value Line (2000) Value Line Investment Survey. A. Bernhard, New York
Veall MR, Zimmerman KF (1994) Goodness of fit measures in the tobit model. Oxford Bulletin of Economics and Statistics 56(4): 485–499
Weiler S, Thompson E, Ozawa T (2001) The evolution of a new industrial district: The automobile industry in the American Southeast. Planning and Markets 4: 24–29
Wensley MRD, Stabler JC (1998) Demand-threshold estimation for business activities in rural Saskatchewan. Journal of Regional Science 38: 155–177

Appendix

The unconditional expectation of number of establishments in the place

The econometric model underlying our analysis is

$$\begin{aligned} y_p &= \lambda_p y_n + \beta_p' x_p + u_p && \text{if RHS } > 0 \\ &= 0 && \text{if RHS} \leq 0 \end{aligned} \tag{1}$$

$$\begin{aligned} y_n &= \lambda_n y_p + \beta_n' x_n + u_n && \text{if RHS } > 0 \\ &= 0 && \text{if RHS} \leq 0 \end{aligned} \tag{2}$$

where x_p and x_n are column vectors of exogenous variables, β_p and β_n are column vectors of coefficients, RHS means "right-hand side" and $(u_p,u_n) \sim \phi_2(0,0,\sigma_p^2,\sigma_n^2,\sigma_{pn})$ where $\phi_2(\bullet,\bullet,\sigma_p^2,\sigma_n^2,\sigma_{pn})$ is the bivariate normal probability density function with variances σ_p^2 and σ_n^2 and covariance σ_{pn}. (Observation subscripts have been suppressed for notational convenience.)

The general formula for the unconditional expected value of y_p is

$$\begin{aligned} \mathrm{E}(y_p) = {} & \mathrm{E}(y_p/y_p > 0, y_n > 0) \bullet \Pr(y_p > 0, y_n > 0) \\ & + \mathrm{E}(y_p/y_p > 0, y_n \leq 0) \bullet \Pr(y_p > 0, y_n \leq 0). \end{aligned}$$

The unconditional expected value of y_p may be calculated using the representation of the four sets of reduced-form coefficients for this model contained in Maddala (1983:206–208) and the formula for the expected value of a random variable drawn from a double truncated bivariate normal distribution contained in Perloff and Sickles (1987:182).

Partition x_p and x_n as $x_p' = (x_{pp}', x')$ and $x_n' = (x_{nn}', x')$ where x is the set of exogenous variables common to both equations, x_{pp} is the set of exogenous variables which appear only in Eq. (1) and x_{nn} is the set of exogenous variables which appear only in Eq. (2).

Let $x_t' = (x_{pp}', x_{nn}', x')$.

Partition the coefficient vectors – β_p and β_n – in a similar fashion:

$$\beta_p' = (\beta_{pp}', \beta_{px}') \text{ and } \beta_n' = (\beta_{nn}', \beta_{nx}').$$

Let $\Delta = (1 - \lambda_p \bullet \lambda_n)$, $\theta_p' = (\beta_{pp}', \lambda_p \bullet \beta_{nn}', \beta_{px}' + \lambda_p \bullet \beta_{nx}')$, $\theta_n' = (\lambda_n \bullet \beta_{pp}', \beta_{nn}', \lambda_n \bullet \beta_{px}' + \beta_{nx}')$, $\alpha_p' = (1/\Delta)\theta_p'$ and $\alpha_n' = (1/\Delta)\theta_n'$.

Define disturbances

$$v_p = u_p + \lambda_p \bullet u_n \text{ and } v_n = \lambda_n \bullet u_p + u_n \quad \text{with } (v_p, v_n) \sim \phi_2(0, 0, \omega_p^2, \omega_n^2, \omega_{pn}).$$

From the definitions of v_p and v_n, we have

$$\omega_p^2 = \sigma_p^2 + \lambda_p^2 \bullet \sigma_n^2 + 2 \bullet \lambda_p \bullet \sigma_{pn} \qquad \omega_n^2 = \lambda_n^2 \bullet \sigma_p^2 + \sigma_n^2 + 2 \bullet \lambda_n \bullet \sigma_{pn}$$
$$\omega_{pn} = \lambda_n \bullet \sigma_p^2 + \lambda_p \bullet \sigma_n^2 + (1 + \lambda_n \bullet \lambda_p)\sigma_{pn}$$

Let

$$\xi_p^2 = \omega_p^2/\Delta^2 \qquad \xi_n^2 = \omega_n^2/\Delta^2 \qquad \xi_{pn}^2 = \omega_{pn}^2/\Delta^2$$

Finally, note that the covariance of u_p and v_n is $\quad v_{np} = \lambda_n \bullet \sigma_p^2 + \sigma_{pn}$.

Define the following

$$\mu_{p1} = \alpha_p' x_t \qquad \mu_{p2} = \beta_p' x_p. \qquad \mu_{n1} = \alpha_n' x_t \qquad \mu_{n2} = \theta_n' x_t$$

Let $\phi(\bullet)$ be the standard normal probability density function, $\Phi(\bullet)$ be the standard normal cumulative distribution function and $\Phi_2(\bullet, \bullet; \rho)$ be the standard bivariate normal cumulative distribution function with correlation coefficient ρ.

Let

$$P_1 = \Phi_2(\mu_{p1}/\xi_p, \mu_{n1}/\xi_n; \xi_{pn}/(\xi_p \bullet \xi_n))$$
$$P_2 = \Phi_2(\mu_{p2}/\sigma_p, -\mu_{n2}/\omega_n; -v_{np}/(\sigma_p \bullet \omega_n)).$$

The unconditional expected value of y_p for our simultaneous equation Tobit model is

$$\begin{aligned} \mathrm{E}(y_p) = {} & \mu_{p1} \bullet P_1 + \xi_p \bullet \phi(\mu_{p1}/\xi_p) \bullet \Phi((\mu_{n1} - (\xi_{pn}/\xi_p^2) \bullet \mu_{p1})/(\xi_n^2 - \xi_{pn}^2/\xi_p^2)^{1/2}) \\ & + (\xi_{pn}/\xi_n) \bullet \phi(\mu_{n1}/\xi_n) \bullet \Phi((\mu_{p1} - (\xi_{pn}/\xi_n^2) \bullet \mu_{n1})/(\xi_p^2 - \xi_{pn}^2/\xi_n^2)^{1/2}) \\ & + \mu_{p2} \bullet P_2 + \sigma_p \bullet \phi(\mu_{p2}/\sigma_p) \bullet \Phi((-\mu_{n2} + (v_{np}/\sigma_p^2) \bullet \mu_{p2})/(\omega_n^2 - v_{np}^2/\sigma_p^2)^{1/2}) \\ & - (v_{np}/\omega_n) \bullet \phi(\mu_{n2}/\omega_n) \bullet \Phi((\mu_{p2} - (v_{np}/\omega_n^2) \bullet \mu_{n2})/(\sigma_p^2 - v_{np}^2/\omega_n^2)^{1/2}). \end{aligned}$$