View metadata, citation and similar papers at core.ac.uk

brought to you by CORE

provided by Research Papers in Economics

A Tractable Approach to the Firm Location Decision Problem[1]

Paulo Guimarães
Universidade do Minho and NIMA
Braga, Portugal
Email: paulog@eeg.uminho.pt

Octávio Figueiredo
Universidade do Porto and CEMPRE
Porto, Portugal

Douglas Woodward
University of South Carolina
Columbia, South Carolina USA
Email: Woodward@darla.badm.sc.edu

July 14, 2000

[1]Paper presented at the 40th European Congress, European Regional Science Association, 1 September, 2000.

1 Introduction

The location of economic activity spawns a large empirical literature. A central concern of urban and regional economists, location studies inform important public policy debates; for example, by addressing the contentious issue of spatially targeted tax incentives used throughout the world to promote economic development.[1] Firm and industry location research also attracts economists interested in the spatial distribution of foreign direct investment, where countries, not cities or regions, form the location choice set.[2] Increasingly, the literature has turned to model location decision probabilities against many spatial choices, just as firms face when making actual site selection decisions. Discrete choice econometric advances have complemented the increasing availability of more accurate micro data sets.

From a theoretical standpoint, interregional and international location issues remain alive as well. Location has been a principal feature of urban and regional theory since European scholars of the early 20th century, notably August Lösch and Alfred Weber, began to develop models grounded in least-cost, optimal choice behavior. Over the years theoretical models have incorporated agglomeration or spatial externalities along with traditional costs like wages and transportation. The "new economic geography" that emerged in United States during the early 1990s revived old questions about location dynamics and the influence of firm site selection decisions on growth and development. Agglomeration economies were recast in formal models advanced by some of contemporary economics' most prominent theorists and prolific writers [Krugman (1991*a*, 1991*b*), Porter (1994), Arthur

(1994), Venables (1996), Hanson (1997), Krugman (1998); for a critique see Martin (1999)].

Given extensive analysis and enduring relevance, the underlying determinants of the firm and plant location decision should be well established. We should know a lot about the relative importance of economic factors (such as transport costs, wages, and agglomeration economies) vis-à-vis policy influences (infrastructure provision and tax incentives). But the results of this vast empirical literature vary widely across studies.[3] Moreover, the basic questions keep getting recast in different models. Is agglomeration really the dominant force in location that theory would predict? Which industries are more influenced by labor, energy, or transport costs? What is the real efficacy of tax incentives on location? Almost invariably, the motivation for the research is that major questions remain unanswered.[4]

Clearly, a systematic approach to location modeling has not been found. One problem is that the spatial scale tested in the empirical literature extends from neighborhoods to nation states. In any case, proper location modeling requires highly disaggregated, microlevel industrial and spatial data. The underlying research program should have a clear microfoundation and conform with actual firm decision making—ideally based on micro data sets that reflect detailed information about local areas, which then can be nested in decision models that look at the determinants of larger areas. Statistical advances and computer applications make testing large micro data sets feasible.

The most appealing approach to empirical location research was pioneered by Carlton (1979, 1983), who tested the probability that a branch plant (in one of three narrowly defined industries) would chose a metropolitan location in the United States. This seminal work established the methodological basis for the subsequent applied location research reviewed throughout this paper. Carlton's significant and lasting contributions were two-fold. First, his work was based on a rich micro data base that focused the location decision problem on narrowly defined industries and geographic areas. Second, Carlton applied a discrete choice (conditional logit) model, opening up a new paradigm in applied location research. The paper suggested that location decision probabilities could be modeled in a partial equilibrium framework, following a logical and verifiable economic process that results from profit maximizing behavior across spatial choices.

This paper will argue that problems arose in the aftermath of Carlton's path-breaking work. These problems hindered further progress and refinement in an otherwise promising line of research. Major econometric drawbacks arose when the model had the decision maker (the firm) facing complex choice scenarios with many spatial alternatives. Moreover, following Carlton's logit approach, researchers had to confront the Independence of Irrelevant Alternatives (IIA) assumption, which, in a spatial context, states that decision makers look at all locations as similar, after controlling for the observable characteristics tested in the model.

In the following we contend that the Poisson regression is a tractable solution to these problems. In fact, we demonstrate that the coefficients of the conditional logit can be equivalently estimated using the Poisson regression. As a consequence, the Poisson regression is compliant with the Random Profit Maximization framework. At the same time, it easily handles the above mentioned problems of Carlton's logit approach to the location decision problem. Thus, the Poisson regression could be established as the basis for the new work that seeks to exploit micro data sets, more likely to be obtainable as more information about location is provided by government statistical agencies.

The rest of the paper is comprised of three sections. The following reviews the discrete choice model and attendant problems when applied to location. Section 3 then explores the relation between the conditional logit model and the Poisson regression. Section 4 offers some concluding comments and points to directions for future research.

2 The Discrete Choice Model

The discrete choice model is now well established as the prevailing empirical method underlying industrial location studies. This modelling approach was first implemented when Carlton (1979) realized that McFadden's multinomial logit model could be easily adapted to the firm location decision problem. Most subsequent research on this topic has relied on the discrete choice methodology [e.g. Carlton (1983), Bartik (1985), Luger & Shetty (1985), Hansen (1987), Schmenner, Huber & Cook (1987), Coughlin, Terza & Arromdee (1991), Woodward (1992), Friedman, Gerlowski & Silberman (1992), Head, Ries & Swenson (1995), Guimarães, Rolfe & Woodward (1998), Guimarães, Figueiredo & Woodward (2000)].

The popularity of this approach resides in the fact that the resulting econometric specification is obtained directly from the Random Utility (Profit) Maximization framework developed by McFadden (1974). If we consider the existence of J spatial choices with $j = 1, ..., J$ and N investors with $i = 1, ..., N$, then the profit derived by investor i if he locates at area j is given by

$$\pi_{ij} = \boldsymbol{\beta}'\mathbf{z}_{ij} + \varepsilon_{ij} \tag{1}$$

where $\boldsymbol{\beta}$ is a vector of unknown parameters, $\mathbf{z}_{ij}$ is a vector of explanatory variables and ε_{ij} is a random term. Thus, the profit for investor i of locating at j is composed of a deterministic and a stochastic component. The investor will choose the area that will yield him the highest expected profit.

If the ε_{ij} are independent and Weibull distributed, then it can be shown that

$$p_{ij} = \frac{\exp(\boldsymbol{\beta}'\mathbf{z}_{ij})}{\sum_{j=1}^{J} \exp(\boldsymbol{\beta}'\mathbf{z}_{ij})} \tag{2}$$

where p_{ij} is the probability that investor i locates at j. If we let $d_{ij} = 1$ in case individual i picks choice j and $d_{ij} = 0$ otherwise, then we can write the log-likelihood of the conditional logit model as

$$\log L_{cl} = \sum_{i=1}^{N} \sum_{j=1}^{J} d_{ij} \log p_{ij}. \tag{3}$$

In practice, the application of this approach to industrial location studies poses several questions. The first one is related to the spatial choice set. Several authors [Bartik (1985), Coughlin, Terza & Arromdee (1991), Friedman, Gerlowski & Silberman (1992), Friedman, Fung, Gerlowski & Silberman (1996), Head, Ries & Swenson (1995)] have modelled location choices among highly aggregated regions such as U.S. states, large geographic units that encompass substantial heterogeneity within themselves. Ideally, narrow areas should be used because factors usually identified as relevant for location decisions (such as agglomeration economies, labor market conditions, or the cost of land) apply to a local level and consequently they can not be adequately accounted for when the model considers large areas in the spatial choice set.[5] This problem was recognized by the pioneers of empirical location studies such as Carlton (1983), who used very narrowly defined geographic regions in the United States[6], and Hansen (1987), who used cities in the São Paulo state in Brazil. Woodward (1992) used separated conditional logit models to test location decisions in both states and counties across the United States. More recent studies [Guimarães, Rolfe & Woodward (1998), Guimarães, Figueiredo & Woodward (2000)] resumed the narrowly defined spatial choice set approach.[7]

An econometric difficulty raised by the use of narrowly defined regions has to do with the handling of large choice sets. In this case it becomes cumbersome to estimate a conditional logit model. In the past some researchers [Hansen (1987), Woodward (1992), Friedman, Gerlowski & Silberman (1992), Guimarães, Rolfe & Woodward (1998), Guimarães, Figueiredo & Woodward (2000)] have followed a suggestion given by McFadden (1978), in which the logit model could still be estimated by using smaller choice sets which were randomly selected from the full choice set. The estimators will still be consistent but not much is known about its small sample properties which may be very different from the asymptotic ones. Clearly, they should be less efficient because they disregard useful information.[8] An additional drawback of the estimates obtained by sampling alternatives is that they cannot be independently replicated.

A second problem of the conditional logit model has to do with the Independence of Irrelevant Alternatives (IIA) assumption. Conditional logit models rely on the idea that the ε_{ij} are independent across individuals and choices and consequently that all locations are symmetric substitutes after controlling for observable characteristics. While some researchers have simply ignored this potential problem others have controlled for the existence of unobservable correlation across choices.[9] Hansen (1987), Ondrich & Wasylenko (1993) and Guimarães, Rolfe & Woodward (1998) have estimated a two-step limited information nested logit while other authors such as Bartik (1985), Woodward (1992) and Luker (1998) have controlled for the IIA by introducing dummy variables for larger regions. However, both approaches are only valid if we are willing to assume that IIA holds within subsets of the choice set (lower level nests for the nested logit and larger regions for the dummy procedure).

Another problem often cited in the literature is related to the common problem that no investment decisions are observed in some particular areas. While this is not an issue from an econometric standpoint, it is a practical problem in the sense that existing software may require that all choices be selected. This has led some authors to drop spatial choices from their data set [Woodward (1992), Head, Ries & Swenson (1995), Luker (1998)], throwing away useful information. Others, in response to this problem, approached the location problem differently, applying "tobit" regression to the number of investments in a given region [Smith & Florida (1994), Ó'hUallacháin & Reid (1997)]. Although "tobit" is a common alternative to deal with the "zero problem," it is hardly justified in this location context. Importantly, it lacks a theoretical underpinning such as the Random Utility Maximization framework. Also, the dependent variable is discrete and the zero observation is a natural outcome of the variable being modelled.

In the following we show that an equivalent approach to the conditional logit model is the Poisson regression. As it turns out, the Poisson regression will produce exactly the same results as the conditional logit model. Moreover, under certain circumstances, present in industrial location decisions studies, the Poisson regression is substantially simpler to implement given that it naturally handles the "zero" and the "large choice set" problems. We also show that the fixed-effects version of this model can be used to more effectively control for the potential IIA problem resulting from unequal substitutability of elemental spatial choices.

3 The Relation between the Conditional Logit Model and the Poisson Regression

A well-known relation between the Poisson and the multinomial distribution (Johnson, Kotz & Balakrishnan 1997) states that if $X_1, X_2, ..., X_J$ are independent Poisson random variables with expected values $\lambda_1, \lambda_2, ..., \lambda_J$, respectively, then, conditional on $\sum_{j=1}^{J} X_j = N$, the joint distribution of $X_1, X_2, ..., X_J$ is multinomial with parameters $(N; p_1, p_2, ..., p_J)$ where p_j, the probability of occurrence of event j, is given by $p_j = \frac{\lambda_j}{\sum_{j=1}^{J} \lambda_j}$. It is clear that if the usual parametrization for the Poisson model is used, $\lambda_j = \exp(\alpha + \boldsymbol{\beta}'\mathbf{z}_j)$, then p_j reflects the conditional logit specification.[10] We will explore this relation and verify how it applies to the cases commonly dealt with in the industrial location literature.

3.1 Case 1: $\mathbf{z}_{ij} = \mathbf{z}_j$

Let us start by assuming that individual decisions are based exclusively in a vector of choice-specific attribute variables common to all decision-makers, as in Bartik (1985), Coughlin, Terza & Arromdee (1991), Woodward (1992) and Guimarães, Rolfe & Woodward (1998). In this case, $\mathbf{z}_{ij} = \mathbf{z}_j$ and so the log-likelihood for the conditional logit model equals,

$$\log L_{cl} = \sum_{i=1}^{N} \sum_{j=1}^{J} d_{ij} \log p_{ij} = \sum_{j=1}^{J} n_j \log p_j, \qquad (4)$$

where n_j is the number of investments placed in location j.

Alternatively, we can let the n_j be independently Poisson distributed with,

$$E(n_j) = \lambda_j = \exp(\alpha + \boldsymbol{\beta}'\mathbf{z}_j) \tag{5}$$

then we can write the log-likelihood function as,

$$\begin{aligned} \log L_P &= \sum_{j=1}^{J} \left[-\lambda_j + n_j \log(\lambda_j) - \log(n_j!) \right] \\ &= \sum_{j=1}^{J} \left[-\exp(\alpha + \boldsymbol{\beta}'\mathbf{z}_j) + n_j (\alpha + \boldsymbol{\beta}'\mathbf{z}_j) - \log(n_j!) \right]. \end{aligned}$$

From the first order condition with respect to α we obtain,

$$\frac{\partial \log L_P}{\partial \alpha} = \sum_{j=1}^{J} \left[n_j - \exp(\alpha + \boldsymbol{\beta}'\mathbf{z}_j) \right] = 0$$

and so,

$$\exp(\alpha) = \frac{N}{\sum_{j=1}^{J} \exp(\boldsymbol{\beta}'\mathbf{z}_j)}.$$

If we replace α back into the log-likelihood we obtain the concentrated log-likelihood,

$$
\begin{aligned}
\log L_{Pc} &= -N + N\log(N) - \sum_{j=1}^{J} n_j \log\left(\sum_{j=1}^{J}\exp(\boldsymbol{\beta}'\mathbf{z}_j)\right) + \\
&\quad + \sum_{j=1}^{J} n_j \boldsymbol{\beta}'\mathbf{z}_j - \sum_{j=1}^{J}\log(n_j!) \qquad (6)\\
&= \sum_{j=1}^{J} n_j \log p_j - N + N\log(N) - \sum_{j=1}^{J}\log(n_j!)\,.
\end{aligned}
$$

The first term in the expression is the log-likelihood of the conditional logit model and the additional terms are constants. Consequently, the estimates obtained for $\boldsymbol{\beta}$ are the same in both models. The estimated covariance matrix will also be identical in both models provided the estimator is the negative inverted of the empirical Hessian (Davidson & MacKinnon 1993).

Thus, we can conclude that results such as those obtained in Bartik (1985), Coughlin, Terza & Arromdee (1991) and Woodward (1992) could be identically obtained by running a simple Poisson model with the number of investments in each location as a dependent variable and $\mathbf{z}_j$ as explanatory variables. It is also now evident that the zero observations constitute no problem and could have been considered in the estimation procedure. It should also be clear that the estimation of the lower level nests in Woodward (1992) and Guimarães, Rolfe & Woodward (1998) would have benefited if the authors had considered the Poisson regression approach as an alternative to the random based technique to overcome the large number of choices. Note also that our result shows that the number of choices in the conditional logit equals the number of observations in the Poisson regression. Since from a purely statistical point of view a larger number of observations (choices) is desirable, the statistical evidence offered by studies that have modelled location choices among highly aggregated regions [such as Bartik (1985) and Coughlin, Terza & Arromdee (1991)] is limited.

3.2 Case 2: $\mathbf{z}_{ij} = \mathbf{z}_{jg}$, with $g = 1, 2, ..., G$

Next consider a more complex approach in which each location decision is based in a vector of choice-specific attribute variables common to groups of individuals. In that more general case $\mathbf{z}_{ij} = \mathbf{z}_{jg}$, with $g = 1, 2, ..., G$, where G is the number of different groups of investors. This modelling approach was followed in Hansen (1987), Friedman, Gerlowski & Silberman (1992), Head, Ries & Swenson (1995) and Guimarães, Figueiredo & Woodward (2000). Hansen's (1987) specification includes choice specific explanatory variables that also change with the two-digit manufacturing sector of the investors while in Friedman, Gerlowski & Silberman (1992) the explanatory variables are grouped in three distinct time periods according to the date of the investments. In Head, Ries & Swenson (1995) and Guimarães, Figueiredo & Woodward (2000) studies investors are grouped in both dimensions, time periods and industrial sectors.

In this second case, the log-likelihood for the conditional logit model is given by

$$\log L_{cl} = \sum_{i=1}^{N} \sum_{j=1}^{J} d_{ij} \log p_{ij} = \sum_{g=1}^{G} \sum_{j=1}^{J} n_{jg} \log p_{jg} \tag{7}$$

where n_{jg} is the number of firms from group g that select location j.

Alternatively, we can let the n_{jg} be independently Poisson distributed with

$$E(n_{jg}) = \lambda_{jg} = \exp\left(\boldsymbol{\alpha}'\mathbf{d}_{jg} + \boldsymbol{\beta}'\mathbf{z}_{jg}\right) \tag{8}$$

where $[\boldsymbol{\alpha}, \boldsymbol{\beta}]$ is the vector of parameters to be estimated and $\mathbf{d}_{jg}$ is a vector of G dummy variables each one assuming one if the observation belongs to group g. Consequently, the log-likelihood for the Poisson model is

$$\begin{aligned}\log L_P &= \sum_{g=1}^{G}\sum_{j=1}^{J}\left[-\lambda_{jg} + n_{jg}\log\left(\lambda_{jg}\right) - \log\left(n_{jg}!\right)\right] \\ &= \sum_{g=1}^{G}\sum_{j=1}^{J}\left[-\exp(\boldsymbol{\alpha}'\mathbf{d}_{jg} + \boldsymbol{\beta}'\mathbf{z}_{jg}) + n_{jg}\left(\boldsymbol{\alpha}'\mathbf{d}_{jg} + \boldsymbol{\beta}'\mathbf{z}_{jg}\right) - \log\left(n_{jg}!\right)\right]\end{aligned}$$

From the first order conditions with respect to the α_gs we obtain

$$\frac{\partial \log L_P}{\partial \boldsymbol{\alpha}_g} = \sum_{j=1}^{J}\left[n_{jg} - \exp(\alpha_g + \boldsymbol{\beta}'\mathbf{z}_{jg})\right] = 0$$

and so, $\exp(\alpha_g) = \frac{n_g}{\sum_{j=1}^{J}\exp(\boldsymbol{\beta}'\mathbf{z}_{jg})}$, where we let $n_g = \sum_{j=1}^{J} n_{jg}$.

Now, we can "concentrate-out" the α_gs to obtain

$$\log L_{Pc} = \sum_{g=1}^{G}\sum_{j=1}^{J} n_{jg}\log p_{jg} - N + \sum_{g=1}^{G} n_g\log(n_g) - \sum_{g=1}^{G}\sum_{j=1}^{J}\log(n_{jg}!) \tag{9}$$

Again, the Poisson concentrated log-likelihood is identical to the log-likelihood function of the conditional logit model plus a set of constants. The estimates obtained from any of the two models are equivalent. Hence, the above comments regarding the use of the random procedure, the discarding of zeros, and the modelling location choices among highly aggregated regions apply equally well to this second case.[11] Thus, many previous studies, including Hansen (1987), Friedman, Gerlowski & Silberman (1992), Head, Ries & Swenson (1995) and Guimarães, Figueiredo & Woodward (2000) would also have benefited if they had considered the Poisson regression as an alternative to the conditional logit model.

3.3 Case 3: Controlling for the IIA Violation

Typically, industrial location researchers have focused on the potential problem caused by the existence of unobserved correlation across elemental choices that can generate a form of the IIA violation.[12] Train (1986) shows that adding alternative specific constants for each individual choice enables the use of the logit specification in the presence of the above mentioned type of IIA violation. This can be justified if we assume that the IIA problem is motivated by the existence of unobservable choice characteristics. As indicated in the introduction, some studies, in line with this suggestion, have introduced dummy variables for groups of elemental choices[13]. However, this approach is not satisfactory because it assumes that the IIA assumption still holds within the defined groups. As shown in Train (1986) to more effectively control for the potential violation of the IIA assumption one should include a dummy variable for each individual choice. This amounts to a specification of the following type,

$$\pi_{jg} = V_{jg} + \varepsilon_{jg} = \delta_j + \boldsymbol{\beta}'\mathbf{z}_{jg} + \varepsilon_{jg} \tag{10}$$

where the δ_js are alternative specific constants introduced to absorb factors which are specific to each particular choice. In this case all explanatory variables (observable or non-observable) that only change across choices (i.e. of type $\mathbf{z}_j$) are absorbed by the alternative specific constants. However, in the presence of a large choice set the implementation of this suggestion is impractical because of the large number of parameters to be estimated. On the other hand, in light of the relation shown above (Case 2) it becomes clear that the alternative specific constant is a fixed-effect in a Poisson regression model. Thus, as shown in Hausman, Hall & Griliches (1984), the $[\boldsymbol{\alpha}, \boldsymbol{\beta}]$ vector can be estimated regardless of the number of δ_j parameters by running a Poisson regression with expected value equal to

$$E(n_{jg}) = \lambda_{jg} = \exp\left(\boldsymbol{\alpha}'\mathbf{d}_{jg} + \boldsymbol{\beta}'\mathbf{z}_{jg} + \delta_j\right) \tag{11}$$

That is, the above problem can be modelled as a Poisson regression with fixed-effects (the δ_js) where we include as explanatory variables the $\mathbf{z}_{jg}$ as well as a dummy variable for each group (the $\mathbf{d}_{jg}$).[14]

4 Conclusion

This paper demonstrates that the Poisson regression is an equivalent approach that offers some advantages to the commonly used conditional logit location model. Actually the coefficients of the conditional logit model can be equivalently estimated using a Poisson regression. Most recent empirical studies of industrial location decisions could have benefited if the authors were aware of this relation. Moreover, this discovery may prove particularly useful for further research in partial equilibrium location modelling. The increasing availability of detailed micro data sets will probably stimulate studies using more narrowly defined choice sets, since from a theoretical standpoint the use of small areas is desirable. In fact, factors usually deemed relevant for location decisions (such as agglomeration economies, labor market conditions, or the cost of land) apply to a local level and cannot be adequately accounted for when the specified model only considers large areas in the spatial choice set. Yet, in the presence of large choice sets, it becomes cumbersome to estimate the conditional logit model. In practice estimators based on existing methods to accommodate large choice sets within the conditional logit model are not efficient. The Poisson regression offers a tractable alternative that will produce exactly the same results as the conditional logit and easily handles the "large choice set" problem.

As demonstrated in this paper the use of narrowly defined areas is also desirable from an econometric point of view because increasing the number of choices in the conditional logit model is equivalent to increasing the number of observations in the Poisson regression. Thus, our result also shows that the statistical evidence supplied by most of the major past studies is limited since they have modelled location choices among highly aggregated regions. Consequently, a new generation of studies using the more detailed spatial data set currently available is necessary.

An econometric difficulty raised by the use of narrowly defined regions results from the fact that if the IIA violation occurs it is more likely to be relevant when dealing with small geographical units. However, in the modelling approach of the more complex and general of the two cases commonly dealt with in the industrial location literature (Case 2), the fixed effect version of the Poisson regression can be used to more effectively control for the potential IIA violation resulting from unequal substitutability of elemental spatial choices.

This paper has focused on the firm location decision problem. However, our results may prove equally useful in applications to other problems whenever the logit model is required. At the same time, we indirectly show that the coefficients of the Poisson model can be given an economic interpretation compatible with the Random Utility (Profit) Maximization framework.

Acknowledgments:

This research was supported with a grant from FCT: (PRAXIS/C/ECO/13134/1998), Portugal.

We would like to thank Santos Silva for helpful suggestions.

References

Anderson, John E. & Robert W. Wassmer. 2000. *Bidding for Business: The Efficacy of Local Economic Development Incentives in a Metropolitan Area.* Kalamazoo, MI: W.E. Upjohn Institute For Employment Research.

Arthur, W.B. 1994. Increasing Returns and Path Dependence in the Economy. In *Increasing Returns and Path Dependence in the Economy*, ed. W.B. Arthur. Michigan: Michigan University Press pp. 99–110.

Bartik, Timothy. 1985. "Business Location Decisions in the United States: Estimates of the Effects of Unionization, Taxes, and Other Characteristics of States." *Journal of Business and Economic Statistics* 3(1):14–22.

Bartik, Timothy. 1991. *Who Benefits from State and Local Economic Development Policies?* Kalamazoo, Michigan: W. E. Upjohn Institute for Employment Research.

Ben-Akiva, Moshe & Steven Lerman. 1985. *Discrete Choice Analysis: Theory and Application to Travel Demand.* Cambridge, MA: The MIT Press.

Carlton, Dennis. 1979. Why New Firms Locate Where They Do: An Econometric Model. In *Interregional Movements and Regional Growth*, ed. W. Wheaton. Washington, D.C.: The Urban Institute pp. 13–50.

Carlton, Dennis. 1983. “The Location and Employment Choices of New Firms: An Econometric Model with Discrete and Continuous Endogenous Variables.” *Review of Economics and Statistics* 65(3):440–449.

Caves, Richard E. 1996. *Multinational Enterprise and Economic Analysis*. Cambridge, U.K.: Cambridge University Press.

Coughlin, Cletus, Joseph Terza & Vachira Arromdee. 1991. “State Characteristics and the Location of Foreign Direct Investment Within the United States.” *Review of Economics and Statistics* 73(4):675–683.

Davidson, Russell & James MacKinnon. 1993. *Estimation and Inference in Econometrics*. New York: Oxford University Press.

Fisher, Peter S. & Alan H. Peters. 1999. *Industrial Incentives: Competition Among American States and Cities*. Kalamazoo, MI: W.E. Upjohn Institute for Employment Research.

Friedman, Joseph, Daniel Gerlowski & Jonathan Silberman. 1992. “What Attracts Foreign Multinational Corporations? Evidence from Branch Plant Location in the United States.” *Journal of Regional Science* 32(4):403–418.

Friedman, Joseph, Hung-Gay Fung, Daniel Gerlowski & Jonathan Silberman. 1996. “A Note on ”State Characteristics and the Location of Foreign Direct Investment Within the United States.”.” *Review of Economics and Statistics* 78(2):367–368.

Guimarães, Paulo, Octávio Figueiredo & Douglas Woodward. 2000. "Agglomeration and the Location of Foreign Direct Investment in Portugal." *Journal of Urban Economics* 47(1):115–135.

Guimarães, Paulo, Robert Rolfe & Douglas Woodward. 1998. "Regional Incentives and Industrial Location in Puerto Rico." *International Regional Science Review* 21(2):119–138.

Hansen, Eric. 1987. "Industrial Location Choice in Sao Paulo, Brazil: A Nested Logit Model." *Regional Science and Urban Economics* 17(1):89–108.

Hanson, G.H. 1997. "Localization Economies, Vertical Organization, and Trade." *American Economic Review* 87:1266–78.

Hausman, Jerry, Brownyn Hall & Zvi Griliches. 1984. "Economic Models for Count Data with an Application to the Patents-RD Relationship." *Econometrica* 52:909–938.

Head, Keith, John Ries & Deborah Swenson. 1995. "Agglomeration Benefits and Location Choice: Evidence from Japanese Manufacturing Investments in the United States." *Journal of International Economics* 38(3-4):223–247.

Johnson, Norman, Samuel Kotz & N. Balakrishnan. 1997. *Discrete Multivariate Distributions.* Wiley Series in Probability and Statistics New York: John Wiley and Sons.

Krugman, Paul. 1991*a*. *Geography and Trade*. Cambridge, MA.: M.I.T. Press.

Krugman, Paul. 1991*b*. "Increasing Returns and Economic Geography." *Journal of Political Economy* 99(3):483–499.

Krugman, Paul. 1998. "What's New About the New Economic Geography." *Oxford Review of Economic Policy* 14:7–17.

Luger, Michael & Sudhir Shetty. 1985. "Determinants of Foreign Plant Start-Ups in the United States: Lessons for Policymakers in the Southeast." *Vanderbuilt Journal of Transnational Law* 18:223–245.

Luker, Bill. 1998. "Foreign Investment in the Nonmetropolitan U.S. South and Midwest: A Case of Mimetic Location Behavior?" *International Regional Science Review* 21(2):163–184.

Martin, Ron. 1999. "The New 'Geographical Turn' in Economics: Some Critical Reflections." *Cambridge Journal of Economics* 23(1):65–91.

McFadden, Daniel. 1974. Conditional Logit Analysis of Qualitative Choice Behavior. In *Frontiers in Econometrics*, ed. Paul Zarembka. New York: Academic Press chapter 4, pp. 105–142.

McFadden, Daniel. 1978. Modelling the Choice of Residential Location. In *Spatial Interaction Theory and Planning Models*, ed. A. Karquist, L. Lundqvist, F. Snickars & J. Weibull. Amsterdam: North-Holland chapter 3, pp. 75–96.

Ó'hUallacháin, Breandán & Neil Reid. 1997. "Acquisition versus Greenfield Investment: The Location and Growth of Japanese Manufacturers in the United States." *Regional Studies* 31(4):403–416.

Ondrich, Jan & Michael Wasylenko. 1993. *Foreign Direct Investment in the United States*. Kalamazoo, Michigan: W. E. Upjohn Institute for Employment Research.

Porter, M. 1994. "The Role of Location in Competition." *Journal of the Economics of Business* 1(1):35–39.

Schmenner, Roger, Joel Huber & Randall Cook. 1987. "Geographic Differences and the Location of New Manufacturing Facilities." *Journal of Urban Economics* 21(1):83–104.

Smith, Donald & Richard Florida. 1994. "Agglomeration and Industrial Location: An Econometric Analysis of Japanese-Affiliated Manufacturing Establishments in Automotive-Related Industries." *Journal of Urban Economics* 36(1):23–41.

Train, Kenneth. 1986. *Qualitative Choice Analysis: Theory, Econometrics, and an Application to Automobile Demand.* Cambridge, Massachussetts: The MIT Press.

Venables, A. 1996. "Loacalisation of Industry and Trade Performace." *Oxford Economic Policy Review* 12(3):52–60.

Woodward, Douglas. 1992. "Locational Determinants of Japanese Manufacturing Start-Ups in the United States." *Southern Economic Journal* 58(3):690–708.

Notes

[1]Recent urban/regional incentive studies include Anderson & Wassmer (2000) and Fisher & Peters (1999). For a useful topology of location incentives and promotion policies see Bartik (1991). The author was also an early contributor to the conditional logit location modeling exhaustively reviewed in this paper (see Bartik (1985)).

[2]A succinct review of the foreign direct investment literature can be found in Caves (1996).

[3]Studies that have highlighted these conflicting results are, for example, Schmenner, Huber & Cook (1987), Coughlin, Terza & Arromdee (1991), Ondrich & Wasylenko (1993) and Guimarães, Figueiredo & Woodward (2000).

[4]Ample examples of this can be found in this Review. Carlton (1983, p. 440) writes "economists know very little about the factors influencing new business location." Refering to the determinants of foreign direct investment in manufacturing, Coughlin, Terza & Arromdee (1991, p. 675) voiced a similar concern. For a more recent location study in this Review, see Friedman et al. (1996).

[5]Consider for example the state of California. If the large number of firms

choosing to locate in this state are drawn by the agglomeration economies of Silicon Valley, a model that considers the state as the unit of decision could be unable to pick up the influence of state agglomeration economies. This would happen because the effect of the local agglomeration economies was diluted in the state variable.

[6]Carlton (1983) modelled the location probabilities for 3 industries defined at 4-digit SICs using SMSAs as the spatial choice set. However, he restricted the alternatives to "those SMSAs in which about 70% of all branch plant births occurred in the industries under study"(p. 443). This restriction constrained the number of spatial choices to 39 for SIC 3079, 24 for SIC 3662 and 26 for SIC 3679.

[7]The increasing availability of micro data sets will probably stimulate a surge in studies using narrowly specified choices.

[8]Train (1986) also notes that the estimator based on a subset of alternatives is not efficient. The same logic applies to the estimates based on the aggregation of alternatives [see McFadden (1978) or Ben-Akiva & Lerman (1985)]. This solution was initially proposed in the context of industrial location decision studies by Bartik (1985) who justified the choice of states as resulting from the aggregation of the true alternatives considered by firms.

[9]Note that if this problem occurs it is more likely to be relevant when dealing with small geographical units. For example, one would expect two adjacent counties to be closer substitutes than two adjacent states.

[10]However, α cancels out of the expression and can not be identified.

[11]Note that in case 2 the number of observations in the Poisson regression equals the number of choices (J) times the number of groups (G).

[12]It is also conceivable that unobserved characteristics of the choosers might make some choices closer substitutes for certain investors.

[13]For example, groups of states in a state choice set analysis.

[14]In practice, for identification reasons, one of the α coefficients has to be normalized to zero.