ECONSTOR
Make Your Publications Visible.

A Service of

ZBW Leibniz-Informationszentrum Wirtschaft
Leibniz Information Centre for Economics

Guimaraes, Paulo; Figueiredo, Octávio; Woodward, Douglas

Conference Paper

Modeling industrial location decisions in U.S. counties

42nd Congress of the European Regional Science Association: "From Industry to Advanced Services - Perspectives of European Metropolitan Regions", August 27th - 31st, 2002, Dortmund, Germany

Provided in Cooperation with:

European Regional Science Association (ERSA)

Suggested Citation: Guimaraes, Paulo; Figueiredo, Octávio; Woodward, Douglas (2002) : Modeling industrial location decisions in U.S. counties, 42nd Congress of the European Regional Science Association: "From Industry to Advanced Services - Perspectives of European Metropolitan Regions", August 27th - 31st, 2002, Dortmund, Germany, European Regional Science Association (ERSA), Louvain-la-Neuve

This Version is available at:
https://hdl.handle.net/10419/115585

Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

Terms of use:

Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.

WWW.ECONSTOR.EU

Mitglied der Leibniz-Gemeinschaft

Modeling Industrial Location Decisions in U.S. Counties[1]

Octávio Figueiredo

Universidade do Porto and CEMPRE

Paulo Guimarães

Universidade do Minho and NIMA

Douglas Woodward[2]

University of South Carolina

April 18, 2002

[1]This research was made during a sabbatical leave by the two first authors at the Division of Research of The Darla Moore School of Business, University of South Carolina, and was partially supported by FCT, the Portuguese Foundation for Science and Technology.

[2]Corresponding author. Address: The Darla Moore School of Business, University of South Carolina, Columbia, South Carolina. E-mail: woodward@darla.badm.sc.edu .

Abstract

Given its sound theoretical underpinnings, the Random Utility Maximization-based conditional logit model (CLM) serves as the principal method for applied research on industrial location decisions. Studies that implemented this methodology, however, had to confront the underlying Independence of Irrelevant Alternatives (IIA) assumption and were unable to fully accommodate this problem. This paper shows that by taking advantage of an equivalent relation between the CLM and Poisson regression likelihood functions one can more effectively control for the potential IIA violation in complex choice scenarios where the decision-maker confronts a large number of spatial alternatives. The paper also provides an illustration, demonstrating the advantages of this relation in investigation of location determinants of new manufacturing plant births in the U.S. counties.

JEL classification: C25, R12, R39.

1 Introduction

The location of economic activity represents a logical and testable case of firm behavior. Not surprisingly, the subject continues to spawn an enormous literature, covering both theoretical and empirical research. While many studies examine intraurban and international location decisions, most research focuses on firm location decisions among regions.[1]

Indeed, from the standpoint of optimal choice theory, location is the oldest branch of regional science. Alfred Weber and August Lösch developed well-known interregional models of profit-maximizing location emphasizing transport costs in the early 1900s, while Edgar Hoover, Walter Isard and Melvin Greenhut, among others, refined the theory at mid-century. Over the years location choice theory has incorporated agglomeration (spatial externalities) along with demand conditions and factor costs. More recently, the "new economic geography" that emerged during the early 1990s revived old questions about location dynamics and the influence of firm site selection decisions on economic growth and development. Agglomeration economies and other spatial forces 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 (1996), Krugman (1998); for a critique see Martin (1999)].

At the same time, empirical studies seeking to identifying the factors that underlie location decisions (markets, agglomeration economies, factor costs) continue to proliferate. Spurring more sophisticated empirical work on location, econometric advances have complemented the increasing availability

of more detailed micro data sets. Increasingly, the empirical literature has turned to model location probabilities against many spatial choices, just as firms face when making site selection decisions. Potentially, these studies contain important findings that can be confirmed or rejected through studies in both similar and different spatial contexts. Reliable estimates across studies can help inform important public policy debates; for example, by assessing the influence of local taxes compared with other regional factors.[2]

Given extensive analysis, the determinants of firm and plant location decisions should be well established. We should know a lot about the relative importance of economic factors (such as factor costs and agglomeration economies) vis-à-vis policy influences (eg. taxes and promotional policies). But the results of the vast location empirical literature vary widely.[3] Moreover, the basic questions keep getting recast in different models. Is agglomeration really the dominant force in location that theory would predict? Do labor and land costs matter? What is the real efficacy of tax abatements on location? Almost invariably, the motivation for more empirical research is that these and other major questions remain unanswered.

Unfortunately, then, a systematic approach to empirical location modeling has not been found. One reason is that the spatial scale tested in the empirical literature extends from neighborhoods to nation states. Location factors (wages and taxes, for example) exert distinct influences on intraurban and international decisions. Even within interregional location studies, however, there seems to be little commonality among the estimates. In part, this is because various econometric approaches have been employed (linear regression models, limited dependent models, and categorical mod-

els). Moreover, the research often fails to take advantage of all available information, including disaggregated data sets that capture microlevel industrial and spatial characteristics. In many cases, the econometric analysis lacks a clear theoretical foundation–in particular, profit maximizing behavior.

In this connection, the most appealing approach to recent interregional 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. 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 the conditional logit model (CLM) for the first time, opening up new possibilities for applied location research. Based on McFadden's (1974) Random Utility Maximization framework, the paper suggested that location decision probabilities could be modeled in a partial equilibrium setting, following a verifiable economic process that results from profit maximizing behavior across spatial choices.

This paper argues that despite the advantages of the CLM, problems arose in the aftermath of Carlton's work. These problems hindered further progress and refinement in an otherwise promising line of research. Specifically, studies that followed the conditional logit approach 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. The

assumption of independent errors is an important one, because, if violated, it can lead to biased coefficient estimates. In practice, as shown in this paper, the empirical studies of location have been unable to fully accommodate the IIA problem within the CLM. Also, the proposed solutions to accommodate complex choice scenarios with the decision maker confronting many (narrowly defined) spatial alternatives have been unsatisfactory. More recent studies on industrial location have tackled this later problem by applying Poisson (count) models. Yet this direction in empirical modeling has not been cast as part of the Random Utility Maximization framework, a main advantage of the McFadden-Carlton approach since it links empirical work to theory.

Here we show how one can more effectively control for the potential IIA violation in complex choice scenarios, regardless of the spatial choice set dimension. This is done by taking advantage of an equivalence relation between the likelihood functions of the conditional logit model and the Poisson regression (Guimarães, Figueiredo & Woodward 2002). We also provide an empirical illustration, wherein we demonstrate how that relation can be helpful to provide more reliable estimates for the location determinants of start-up manufacturing plants in the United States counties. We find strong evidence that agglomeration economies (both urbanization and localization), as well as taxes, influence location decisions. These relations hold across all tested specifications, even when we add stringent controls to account for omitted relevant variables. The evidence concerning other factors (labor costs, land costs, and local markets) is not as conclusive.

The rest of the paper is comprised of four sections. The next section

reviews previous research on industrial location decisions, pointing to perceived problems with CLM and Poisson models. Section 3 proposes solutions to these problems in econometric location modeling. Section 4 offers the empirical illustration, providing evidence for location factors affecting locational choices among U.S. counties. Section 5 summarizes the paper and points to directions for further research.

2 Previous Research on Industrial Location Decisions

Most recent empirically based interregional location papers have relied on the CLM. The virtue in this line of research is a profit maximizing model linking the site selection decision to specific area characteristics. The probability of a new plant being opened at a particular site depends on the relative level of profits that can be derived in this site and hence on the site's attributes compared with those of all other alternatives. As stated in the introduction, this approach was pioneered by Carlton (1983), who modeled the location of new branch plants across standard metropolitan statistical areas (SMSAs) in the United States. With the exception of Hansen (1987), Woodward (1992), and more recently Guimarães, Figueiredo & Woodward (2000), who also relied on narrowly defined spatial choice sets, subsequent research has modeled location choices among highly aggregated regions, such as U.S. states [Bartik (1985), Coughlin, Terza & Arromdee (1991), Friedman, Gerlowski & Silberman (1992), Friedman, Fung, Gerlowski & Silberman (1996), Head, Ries & Swenson (1995), Levinson (1996), Head, Ries & Swen-

son (1999)]. The small number of studies carried out on a narrowly defined spatial scale may be justified by the lack of available data sets (although the information available is growing). Also, the challenge posed by modeling large spatial choice sets within the CLM may have constituted a significant hurdle. When confronted with the large data set problem, researchers have followed McFadden's (1978)suggestion to work with a smaller sample of alternatives randomly drawn from the full choice set [Hansen (1987), Woodward (1992), Friedman, Gerlowski & Silberman (1992)and Guimarães, Figueiredo & Woodward (2000)].[4] A different approach (aggregation alternatives) was proposed by Bartik (1985)who justified the choice of U.S. states as resulting from the aggregation of the true alternatives considered by firms. However, these solutions to overcome the large data set problem are unsatisfactory because they disregard useful information. The resulting estimators are clearly less efficient.

An econometric problem posed by the CLM in the use of narrowly defined spatial sets is that the Independence of Irrelevant Alternatives (IIA) assumption is more likely to be violated. Conditional logit models rely on the assumption that the error terms are independent across individuals and choices. Typically, industrial location researchers have acknowledged the potential problem caused by the existence of unobserved site characteristics that may induce correlation across choices and therefore a violation of the IIA assumption.[5] When dealing with small geographical units, this problem may be more important because site characteristics that are unaccounted for can more easily extend their influence beyond the boundaries of the considered spatial units.[6] Some researchers have attempted to control for the

existence of unobservable correlation across choices. Two different methodologies have been used. Hansen (1987), Ondrich & Wasylenko (1993) and Guimarães, Rolfe & Woodward (1998) estimated a two-step limited information nested logit. The difficulty here resides in the identification of the upper levels as they may constitute unrealistic scenarios for the decision-maker. Moreover, it is sometimes difficult to conceive of regional characteristics that affect upper level location choices in ways different from the elemental choices. Consequently, most authors [e.g. Bartik (1985), Woodward (1992), Luker (1998), Levinson (1996) and Head, Ries & Swenson (1999)] have attempted to control for the IIA violation by introducing dummy variables for larger regions.[7] Both approaches, however, and importantly, are unsatisfactory because they are only valid if one is willing to assume that the IIA assumption holds within subsets of the choice set (lower level nests for the nested logit solution and larger regions for the dummy procedure).

A recent strand of empirical research has modeled the firm location decision problem using Poisson (count) models and microlevel spatial data sets [Papke (1991), Wu (1999), Coughlin & Segev (2000)and List (2001)]. These Poisson studies approached the location problem differently than the CLM. They relate the number of new plants being opened at a particular site to a vector of area attributes. The Poisson regression is particularly advantageous in dealing with large spatial choice sets. Thus, what was perceived as a drawback in the CLM model becomes an advantage in the context of count models.[8] At the same time, the authors claim that extensions of the Poisson regression model can be used to address known problems that surface when applied to location studies. In particular, this is the case of the overdisper-

sion problem caused by the prevalence of zeros [List (2001)] or originated by an excessive spatial concentration of firms [Wu (1999)and Coughlin & Segev (2000)]. Papke (1991)use a fixed-effects Poisson regression to control for unobserved state heterogeneity. Meanwhile, and despite these attractive features of the Poisson regression model, it lacks a theoretical underpinning such as the Random Utility Maximization framework for the CLM.

The link between the CLM and the Poisson regression has been addressed in a recent paper by Guimarães, Figueiredo & Woodward (2002). Tthe paper shows that, under some circumstances, the coefficients of the Poisson model can be given an economic interpretation compatible with the Random Utility Maximization framework. The next section of this paper explores this relation's deeper implications for regional location research, positing instruments to more effectively control for the potential IIA violation in complex choice scenarios with a large number of spatial alternatives.

3 Econometric Aspects of Location Modeling

To show the connection between the received empirical location model, we posit a general profit function for firms in a particular industry and location. Let us start by considering an economy with K different industrial sectors $(k = 1, ..., K)$. There are N investors $(i = 1, ..., N)$ who independently select a location j from a set of J potential locations $(j = 1, ..., J)$. The profit the investor will derive if he selects sector k and locates at area j is assumed to be,

$$\pi_{ijk} = \boldsymbol{\gamma}'\mathbf{x}_k + \boldsymbol{\theta}'\mathbf{y}_j + \boldsymbol{\beta}'\mathbf{z}_{jk} + \varepsilon_{ijk}, \tag{1}$$

where $\boldsymbol{\gamma}$, θ and $\boldsymbol{\beta}$ are vectors of unknown parameters, $\mathbf{x}_k$ is a vector of sector specific variables (e.g. entry barriers or concentration ratios), $\mathbf{y}_j$ is a vector of location specific variables (such as agglomeration economies, land costs or local taxes), and $\mathbf{z}_{jk}$ is a vector of explanatory variables that change simultaneously with the region and the sector (e.g. wages or localization economies). ε_{ijk} is an identically and independently distributed random term assumed to have an Extreme Value Type I distribution. This random term reflects the idiosyncrasies specific to each investor, as well as unobserved attributes of the choices. Based on McFadden (1974)we can show that if investor i is profit oriented then his probability of selecting location j, conditional on his choice of sector k, equals,[9]

$$p_{j/k} = \frac{\exp(\boldsymbol{\theta}'\mathbf{y}_j + \boldsymbol{\beta}'\mathbf{z}_{jk})}{\sum_{j=1}^{J}\exp(\boldsymbol{\theta}'\mathbf{y}_j + \boldsymbol{\beta}'\mathbf{z}_{jk})}. \tag{2}$$

This expresses the familiar CLM formulation. Let us denote by n_{jk} the number of investments in region j and sector k. Then, we can estimate the parameters of the above equation by maximizing the following log-likelihood:

$$\log L_{cl} = \sum_{k=1}^{K}\sum_{j=1}^{J} n_{jk}\log p_{j/k}. \tag{3}$$

As shown in Guimarães, Figueiredo & Woodward (2002)the above log-likelihood function is equivalent to that of a Poisson model which takes as a dependent variable n_{jk} and includes as explanatory variables the $\mathbf{y}_j$ and $\mathbf{z}_{jk}$ vectors plus a set of dummy variables for each sector. That is, we will obtain the same results if we admit that n_{jk} follows a Poisson distribution

with,

$$E(n_{jk}) = \lambda_{jk} = \exp(\alpha_k + \boldsymbol{\theta}' z\mathbf{y}_j + \boldsymbol{\beta}'\mathbf{z}_{jk}), \tag{4}$$

where α_k is a dummy taking the value 1 for sector k.

Our main interest centers on the potential problem caused by the omission of unobserved explanatory variables, which can cause a violation of the IIA assumption. To address this problem, as indicated before, authors such as Bartik (1985), Woodward (1992), Levinson (1996)and Head, Ries & Swenson (1999)have included dummy variables for groups of elemental alternatives. Within the context of the Poisson regression this amounts to adding an additional dummy variable for each group and is equivalent to admitting that each investor restricts his choice set to the group of alternatives where the investment was observed.[10] However, as stated earlier, by doing this one is still assuming that the IIA assumption holds within the groups of alternatives.

To more effectively control for the potential violation of the IIA assumption one should include an additional effect specific to each alternative. This way, we should be able to absorb all the unaccounted for factors affecting the firm location decision. In terms of our model this amounts to adding an additional term to the profit function, γ_j, such that,

$$\pi_{ijk} = \boldsymbol{\gamma}'\mathbf{x}_k + \boldsymbol{\theta}'\mathbf{y}_j + \boldsymbol{\beta}'\mathbf{z}_{jk} + \gamma_j + \varepsilon_{ijk} \tag{5}$$

If we assume that γ_j is a random variable then, conditional on γ_j, the

probability of an investor selecting location j can be expressed as,

$$p_{j/k\gamma} = \frac{\exp(\boldsymbol{\theta}'\mathbf{y}_j + \boldsymbol{\beta}'\mathbf{z}_{jk} + \gamma_j)}{\sum_{j=1}^{J} \exp(\boldsymbol{\theta}'\mathbf{y}_j + \boldsymbol{\beta}'\mathbf{z}_{jk} + \gamma_j)}. \tag{6}$$

The above formulation may be interpreted as a variant of the mixed logit model, where the attributes of the characteristics which are not explicitly modeled are assumed to reside in the error terms.[11] On the other hand, in light of their relation between the CLM and the Poisson regression, one can estimate the model above by means of a Poisson model with random effects.[12] If we assume that $\exp(\gamma_j)$ follows an i.i.d. gamma distribution with $(\delta^{-1}, \delta^{-1})$ parameters and consequently that $E(\exp(\gamma_j)) = 1$ and $V(\exp(\gamma_j)) = \delta$, then, as shown by Hausman, Hall & Griliches (1984), the resulting Poisson model with gamma distributed random effects has an analytically tractable log-likelihood. In the pure cross-section case, this later model collapses to a standard negative binomial regression [Cameron & Trivedi (1998)]. Thus, if our specification does not include sectorial effects (i.e. $\mathbf{z}_{jk}$ variables) one can estimate (6) by applying the negative binomial model. More recently, there have been studies using the negative binomial regression to model location decisions [Wu (1999)and Coughlin & Segev (2000)] but the authors failed to note the compatibility of their approach with the Random Utility Maximization framework.

The CLM with random effects relies on the assumption that the alternative specific effects are uncorrelated with the explanatory variables. This is a questionable assumption for dealing with the IIA problem in location studies. Omitted factors which are supposedly accounted for by the random

effects, such as natural advantages, may be correlated with, for example, density of economic activity.

An alternative approach is to assume that γ_j is a fixed effect. This amounts to including a dummy variable for each elemental alternative (an alternative specific constant). In this case the dummies absorb the effects of the $\mathbf{y}_j$ variables and we may write,

$$p_{j/k} = \frac{\exp(\boldsymbol{\beta}'\mathbf{z}_{jk} + \gamma_j)}{\sum_{j=1}^{J} \exp(\boldsymbol{\beta}'\mathbf{z}_{jk} + \gamma_j)}. \tag{7}$$

However, in the presence of a large choice set the implementation of this specification is impractical because of the large number of parameters to be estimated. On the other hand, in light of the equivalence relation between the log-likelihoods of the CLM and the Poisson regression, the alternative specific constant can be viewed as a fixed-effect in a Poisson regression. Consequently, these effects can be "conditioned-out"and one can still obtain estimates for the $\boldsymbol{\beta}$ vector regardless of the number of parameters (see Appendix B).

The problem with the above approach is that we rely on sectoral variation to estimate the model and consequently are unable to identify the impact of variables that only exhibit intraregional variation (i.e. the $\mathbf{y}_j$ vector). The marginal impact of these variables is of particular interest in location studies. However, as long as we have available data for different time periods exhibiting sufficient time-series variation, one can still obtain estimates for

all parameters of interest. To see this, let

$$p_{tj/k} = \frac{\exp(\boldsymbol{\theta}'\mathbf{y}_{tj} + \boldsymbol{\beta}'\mathbf{z}_{tjk} + \gamma_j)}{\sum_{j=1}^{J} \exp(\boldsymbol{\theta}'\mathbf{y}_{tj} + \boldsymbol{\beta}'\mathbf{z}_{tjk} + \gamma_j)} \tag{8}$$

be the probability that the investor at time t selects location j, conditional on his choice of sector k. Proceeding in a similar fashion as above we can "condition-out" the local fixed effects and obtain estimates for the $\boldsymbol{\beta}$ and $\boldsymbol{\theta}$ vectors.[13]

4 An Empirical Application: Locational Determinants of Manufacturing Plant Births Across the U.S. Counties

4.1 Data and variables

To demonstrate ways to exploit the Poisson-CLM relation as described in the last section, we give an illustration of firm location decisions where there are many spatial choices. Specifically, we model the location determinants of manufacturing plant births for the 3,066 counties belonging to the 48 contiguous U.S. states[14]. To take advantage of the relation between the CLM and Poisson regression, the dependent variable formed for the tests is the number of establishment births for each county by industry (2-digit SIC code for all establishments in the manufacturing sector). We use special U.S. Census Bureau tabulations of the *Standard Statistical Establishments List* encompassing the universe of all new known openings for the years of 1989 and 1997. In Tables 1 and 2 we show the industry sector and spatial

distributions of these new plants. As can be seen, the distributions are relatively stable over time and exhibit a substantial degree of concentration. For both years, the same five most important sectors account for approximately 57 percent of all investments. A similar pattern can be found for the spatial distribution, as the same ten states concentrate 56 percent of new plants births for any of the considered years.[15].

[insert Tables 1 and 2]

The independent variables include the county characteristics that can affect the firm profit function. These characteristics can affect profits both from the cost and revenue side. On the cost side of the profit function we test the cost of labor, land, and capital. The county labor cost is measured by the wage and salary earnings per job in 1988 and 1996 (*LABOR COSTS*).[16] Since industrial and residential users compete for land, when modeling with small areas, as in our case, land costs can be proxied by population density. Consequently, we use population density for the years of 1988 and 1996 to approximate land costs (*LAND COSTS*). Per capita property taxes for 1987 and 1997 are included in the model to account for the tax business climate in each county (*TAXES*). Property taxes affect all private investments made in United States, and vary significantly across counties. Incentives can change effective payments to local governments in some cases, but for the majority of the new plants in our dataset the average county property tax captures a relevant cost of doing business. To account for the revenue (demand) side of the profit function, the model needs to include a measure of market size. As such, we use total county personal income for the years of 1988 and 1996

as an explanatory variable (*MARKET SIZE*).

Over the years theoretical models have incorporated agglomeration or spatial externalities along with factor costs and market dimension. Agglomeration includes both localization economies and urbanization economies. Urbanization economies, i.e. externalities that are common to all firms, are proxied by the county density of manufacturing and service establishments per squared kilometer in 1988 and 1996 (*URBANIZATION ECONOMIES*).[17] Localization economies, external economies that benefit firms in the same industry, are measured by the number of establishments in the same 2-digit SIC industry as the investor per squared kilometer for the same years (*LOCALIZATION ECONOMIES*).[18]

Additional regressors include dummy variables for the states (to account for observable and unobservable state level characteristics) as well as a set of dummies for each combination of year and 2-digit SIC sector to ensure compatibility between the CLM and Poisson approaches.

4.2 Empirical Results

In Table 3 we present the results of our regression analysis. We ran several models. The first one, corresponding to columns 1 and 2, is a standard CLM estimated by means of the equivalence relation with the Poisson regression. In the first specification (column 1) all variables are highly significant and with the expected signs. We find evidence that the costs of production factors (labor costs, land costs and taxes) impact negatively on the probability of location in a given county. Of all these costs, the cost of land has the highest impact. Everything else constant, a 1 percent increase in land costs

leads to an 0.81 percent decrease in the number of new plant births while the same elasticities for labor costs and taxes are -0.46 and -0.26, respectively.[19] We also find evidence that the county market size matters and that agglomeration economies (both localization and urbanization) are associated with higher numbers of plant births. Apparently, of the two agglomeration measures, urbanization economies have the strongest impact.

It may be argued that investment decisions are also affected by state level variables. Consequently, our results in column 1 may be substantially biased. While it could be possible to add some observable state level variables, such as right-to-work (open shop) legislation or state taxes, we opted instead to control for these effects by including "state fixed-effects." By doing this, we are also controlling for unobservable state characteristics and, as argued by some authors, mitigating the IIA problem. The results for this specification are presented in column 2. As expected, the increase in the log-likelihood is statistically significant providing evidence on the relevance of state level characteristics. Notwithstanding, all coefficient estimates remain practically unchanged.

[insert Table 3]

As argued in section 3, to more effectively control for the potential violation of the IIA assumption one should include "county specific-effects." In a first step, we estimate a mixed logit model (with and without "state fixed-effects") by means of a Poisson regression with county random effects. The results are shown in columns 3 and 4. The difference between the log-likelihoods of the model with random effects and the comparable Poisson

regression is statistically significant, providing evidence that the inclusion of random county effects makes sense. At the same time, as can be seen, the results remained remarkably stable. There is no change in the sign and significance of the coefficients, despite the reduction in the magnitude of the values. However, as stated earlier, this model relies on the lack of correlation between the county random effects and the explanatory variables. Based on Hausman, Hall & Griliches (1984), one can test this hypothesis by means of an Hausman test that evaluates the random and the fixed effects estimators. When applied to this setting, the test provides indirect evidence on the correlation between the random effects and the explanatory variables. The statistic equals 390.7 and thus we can not reject the null hypothesis at the 1 percent level of significance.

Therefore, in a final specification, we use an alternative approach to deal with the potential violation of the IIA assumption caused by the omission of relevant variables. We estimate a CLM where we include a dummy variable for each U.S. county. The estimation of this CLM is made by means of a Poisson regression with fixed effects (see column 5, Table 3). The estimates exhibit some noticeable changes. Agglomeration economies (both urbanization and localization) are still significant and with the right sign. The same is true for property taxes. However, the evidence on the significance of local markets and the costs of land and labor disappears. A possible explanation for the observed changes is that the estimates for this model are based exclusively on time series variation. The time variability of our data may be insufficient to identify the importance of these variables.

In sum, when controlling for "county specific-effects" we find strong ev-

idence that agglomeration economies (both urbanization and localization) are relevant factors for explaining location decisions across U.S. counties. Apparently, urbanization economies have a higher impact. Similar evidence about the positive impact of agglomeration economies on interregional and interurban location was found by Carlton (1983), Bartik (1985), Hansen (1987), Levinson (1996)and Figueiredo, Guimarães & Woodward (2002).

When controlling for "county specific-effects" we also find strong evidence that higher property taxes deter investments across U.S. counties. Property taxes in the United States remain a controversial policy issue. While it is often argued that local tax policy is relevant for location decisions, empirical studies have failed to produce strong, consistent evidence. The property tax was tested in various studies of location by foreign investors [Woodward (1992), Coughlin & Segev (2000)and List (2001)] but these studies were unable to found a significant relationship. Carlton (1983)included local taxes in his seminal CLM location model (an interurban choice model without fixed effects), but was unable to demonstrate the relevance of property taxes.

Our results for factor costs (land and labor) are not as clear. The same is true for local market size. While these variables are shown to be statistically significant in the model with random effects, the same is not true for the fixed-effects model. With the exception of Papke (1991)and Figueiredo, Guimarães & Woodward (2002), previous empirical research on domestic decisions failed to demonstrate the relevance of land costs [Bartik (1985)and Hansen (1987)]. Evidence for the negative impact of labor costs on domestic location decisions was found by Bartik (1985)and Figueiredo, Guimarães

& Woodward (2002), a result not corroborated by other studies [Carlton (1983), Hansen (1987)and Levinson (1996)]. None of the above domestic studies tested market size.

5 Conclusion

As one of the central concerns of regional analysis, location studies require a sound empirical and theoretical foundation. Given its microfoundation, the CLM has been the most promising econometric approach for modeling industrial location decisions under profit maximization. This CLM established a solid methodological basis for applied location research. However, related research on this topic has been unable to fully accommodate the problem posed by the IIA assumption. This assumption becomes even more problematic when dealing with complex choice scenarios where the decision-maker confronts a large number of narrowly defined spatial alternatives.

In this paper we show that by taking advantage of the equivalence relation between the log-likelihood functions of the CLM and the Poisson regression one can more effectively control for the potential IIA violation resulting from omitted attribute characteristics. Both the random and the fixed effects versions of the Poisson regression can be used to introduce an additional effect specific to each spatial alternative. The introduction of these specific effects should absorb all the unaccounted for factors affecting the firm location decision and thus provide a control for the potential IIA violation. Meanwhile, the implementation of the fixed-effects version of the Poisson regression requires time series data exhibiting sufficient temporal

variation. Fortunately, reliable micro data sets like the one obtained in this paper to test our propositions are becoming increasingly available for longer time periods.

As stressed in this paper, our approach to the IIA problem is compliant with the Random Profit Maximization framework. Estimating a Poisson regression model with random effects is equivalent to estimating a particular case of the mixed logit model. Equivalently, the results of a Poisson regression with fixed-effects are the same as those obtained from a CLM with an alternative specific constant. Hence, this paper also shows that there is a theoretical foundation for a recent branch of the location literature that relies on the Poisson model and its extensions.

APPENDIX A

To simplify matters, let us admit that the probability of locating in a particular site is only a function of area characteristics ($\mathbf{y}_j$), as in Bartik (1985), Woodward (1992)and Levinson (1996). Replacing the j index by an index for state, s, and for county, c, we obtain,

$$p_{sc} = \frac{\exp(\boldsymbol{\alpha}_s + \boldsymbol{\theta}'\mathbf{y}_{sc})}{\sum_{s=1}^{S}\sum_{c=1}^{C_s}\exp(\boldsymbol{\alpha}_s + \boldsymbol{\theta}'\mathbf{y}_{sc})}, \tag{9}$$

where C_s is the number of counties in state s. Thus, the log-likelihood for the discrete choice problem is:

$$\log L = \sum_{s=1}^{S}\sum_{c=1}^{C_s} n_{sc}\log p_{sc}.$$

If we compute the first order condition with respect to any one of the state "dummy variables" we get,

$$n_s - n\sum_{c=1}^{C_s} p_{sc} = 0,$$

and thus,

$$\exp(\alpha_s) = \frac{n_s}{n}\frac{\sum_{s=1}^{S}\sum_{c=1}^{C_s}\exp(\boldsymbol{\alpha}_s + \boldsymbol{\theta}'\mathbf{y}_{sc})}{\sum_{c=1}^{C_s}\exp(\boldsymbol{\theta}'\mathbf{y}_{sc})}.$$

If we now plug this back into the log-likelihood function we obtain the

following concentrated likelihood function,

$$\log L = \sum_{s=1}^{S} \sum_{c=1}^{C_s} n_{sc} \log \left(\frac{n_s}{n} \frac{\exp(\boldsymbol{\theta}' \mathbf{y}_{sc})}{\sum_{c=1}^{C_s} \exp(\boldsymbol{\theta}' \mathbf{y}_{sc})} \right),$$

$$\log L = \sum_{s=1}^{S} \sum_{c=1}^{C_s} n_{sc} \log(\frac{n_s}{n}) + \sum_{s=1}^{S} \sum_{c=1}^{C_s} n_{sc} p_{c/s}. \tag{A1}$$

where,

$$p_{c/s} = \frac{\exp(\boldsymbol{\beta}' \mathbf{z}_{sc})}{\sum_{c=1}^{C_s} \exp(\boldsymbol{\beta}' \mathbf{z}_{sc})}.$$

is the probability of an investor locating in a particular county, conditional on the chosen state. The first term in expression (A1) is a constant. The second term is the log-likelihood for a discrete choice problem where the choice sets are restricted to the states where the investments were observed.

APPENDIX B

The log-likelihood for this conditional logit problem is,

$$\log L_{cl} = \sum_{k=1}^{K} \sum_{j=1}^{J} n_{jk} \log p_{j/k}.$$

From the first order condition for maximization with respect to one of the fixed effects we obtain,

$$\frac{\partial \log L_{cl}}{\partial \gamma_j} = \sum_{k=1}^{K} \left[n_{jk} - p_{j/k} n_k \right] = 0$$

Solving the first order condition with respect to γ_j we arrive at,

$$n_j = \exp(\gamma_j) \sum_{k=1}^{K} \frac{\exp(\boldsymbol{\beta}' \mathbf{z}_{jk})}{\sum_{j=1}^{J} \exp(\boldsymbol{\beta}' \mathbf{z}_{jk} + \gamma_j)} n_k$$

Now, if we let,

$$I_k = \log \left(\frac{n_k}{\sum_{j=1}^{J} \exp(\boldsymbol{\beta}' \mathbf{z}_{jk} + \gamma_j)} \right)$$

we can express the γ_js as,

$$\exp(\gamma_j) = \frac{n_j}{\sum_{k=1}^{K} \exp(\boldsymbol{\beta}' \mathbf{z}_{jk} + I_k)}$$

If we plug the γ_js back into the expression for $p_{j/k}$, we obtain,

$$\begin{aligned} p_{j/k} &= \frac{\exp(\gamma_j)\exp(\boldsymbol{\beta}'\mathbf{z}_{jk})}{\sum_{j=1}^{J}\exp(\boldsymbol{\beta}'\mathbf{z}_{jk}+\gamma_j)} \\ &= \frac{n_j}{\sum_{k=1}^{K}\exp(\boldsymbol{\beta}'\mathbf{z}_{jk}+I_k)}\frac{\exp(\boldsymbol{\beta}'\mathbf{z}_{jt})}{\sum_{j=1}^{J}\exp(\alpha_j+\boldsymbol{\beta}'\mathbf{z}_{jk})} \\ &= \frac{\exp(\boldsymbol{\beta}'\mathbf{z}_{jk}+I_k)}{\sum_{t=1}^{T}\exp(\boldsymbol{\beta}'\mathbf{z}_{jk}+I_k)}, \end{aligned}$$

and the concentrated log-likelihood is that of a logit model where the choices are now the sectors with an alternative specific constant added to the model. This log-likelihood is equivalent to that of a Poisson regression with fixed-effects (see, for example, Cameron & Trivedi (1998)].

References

Anderson, John & 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. Brian. 1994. *Increasing Returns and Path Dependence in the Economy.* Ann Arbor, MI: University of Michigan Press.

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, MI: W. E. Upjohn Institute for Employment Research.

Cameron, A. Colin & Pravin K. Trivedi. 1998. *Regression Analysis of Count Data.* Econometric Society Monographs Cambridge, UK: Cambridge University Press.

Carlton, Dennis. 1979. *Why New Firms Locate Where They Do: An Econometric Model.* 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, UK: Cambridge University Press.

Chen, Zhen & Lynn Kuo. 2001. "A Note on the Estimation of the Multinomial Logit Model with Random Effects." *The American Statistician* 55:89–95.

Coughlin, Cletus & Eran Segev. 2000. "Location Determinants of New Foreign-Owned Manufacturing Plants." *Journal of Regional Science* 40:323–351.

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.

Figueiredo, Octávio, Paulo Guimarães & Douglas Woodward. 2002. "Home Field Advantage: Location Decisions of Portuguese Entrepreneurs." *Journal of Urban Economics* . forthcoming.

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 & Doug Woodward. 2002. "A Tractable Approach to the Firm Location Decision Problem." *Review of Economic and Statistics* . forthcoming.

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, Gordon H. 1996. "Localization Econonies, Vertical Organization, and Trade." *American Economic Review* 86:1266–1278.

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, C. Keith, John C. Ries & Deborah L. Swenson. 1999. "Attracting Foreign Manufacturing: Investment Promotion and Agglomeration." *Regional Science and Urban Economics* 29:197–208.

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.

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

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

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

Levinson, Arik. 1996. "Environmental Regulations and Manufacturers' Location Choice: Evidence from the Census of Manufactures." *Journal of Public Economics* 62:5–29.

List, John. 2001. "US County-Level Determinants of Inbound FDI: Evidence from a Two-Step Modified Count Data Model." *International Journal of Industrial Organization* 19:953–973.

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." *Cambrige Journal of Economics* 23:65–91.

McFadden, Daniel. 1974. *Conditional Logit Analysis of Qualitative Choice Behavior*. NY: Academic Press chapter 4, pp. 105–142.

McFadden, Daniel. 1978. *Modelling the Choice of Residential Location.* Amsterdam: North-Holland chapter 3, pp. 75–96.

McFadden, Daniel & Kenneth Train. 2000. "Mixed MNL Models for Discrete Response." *Journal of Applied Econometrics* 15:447–470.

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

Papke, L. 1991. "Interstate Business Tax Differentials and New Firm Location." *Journal of Public Economics* 45:47–68.

Porter, Michael. 1994. "The Role of Location in Competition." *Journal of the Economics of the Business* 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.

Venables, Anthony. 1996. "Localization of Industry and Trade Performance." *Oxford Review of Economic Policy* 12:52–60.

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

Wu, Fulong. 1999. "Intrametropolitan FDI Firm Location in Guangzhou, China: A Poisson and Negative Binomial Analysis." *The Annals of Regional Science* 33:535–555.

Notes

[1]The international location literature chiefly concerns the country-level decisions of foreign direct investors while the intraurban/regional literature focuses on the subnational location decisions of domestic and foreign investors. A succinct review of the foreign direct investment literature can be found in Caves (1996).

[2]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).

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

[4]Note that Carlton (1983) avoided this problem by restricting 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.

[5]It is also conceivable that unobserved characteristics of the choosers

might make some choices closer substitutes for certain investors. In this paper we do not address this problem.

[6]For that reason, one would, for example, expect two adjacent counties to be closer substitutes than two adjacent states.

[7]Usually, by including Census Divisions dummies in studies dealing with choices across the U.S. states.

[8]Note that what were choices in the CLM are now observations.

[9]Note that the sector specific characteristics drop out of the next expression.

[10]For example, in a state choice set analysis, introducing dummy variables for the nine Census Divisions is equivalent to admitting that each investor restricts his choice set to the particular Census Division where the investment was observed. The demonstration is provided in Appendix A. Note also that, in light of this relation, introducing dummies variables for groups of elemental alternatives is equivalent to estimating the lower levels of a two-step limited information nested logit.

[11]This model is extensively reviewed in McFadden & Train (2000).

[12]See Chen & Kuo (2001)for a proof of this result.

[13]Note that, since we now have an additional time dimension, the compatibility between the CLM and Poisson approaches requires the inclusion of dummies for each combination of time period and sector.

[14]A number of counties in Virginia are merged with independent cities. This is because the data for some independent variables (those obtained from the *Regional Economic Information System* database) are reported in this manner.

[15]Note also that the correlation between the spatial distribution of new plants in 1989 and 1997 is 99,7%. The distribution of these plants by the 2-digits SIC sectors in the two considered years also exhibit a strong correlation (96,4%).

[16]While industry-level wages would be preferable, these data present a high number of missing values at the county level.

[17]In the definition of this variable we include SICs 20 to 39 (Manufacturing), SICs 50 and 51 (Wholesale), SICs 52 to 59 (Retail), SICs 60 to 67 (Finance, Insurance and Real State), and SICs 70 to 89 (Services Industries).

[18]All variables were introduced in logarithmic form. Wages and salary earnings per job, personal income and population were taken from the *Re-*

gional Economic Information System (REIS) database published by the Bureau of Economic Analysis (Table CA30 and Table CA05). The number of establishments at the 2-digit SIC level was obtained from the U.S. Bureau of Census, *County Business Patterns*. The source for *per capita* property tax is also the U.S. Bureau of Census, *Census of Government*. Land area is from the *Census Geographic Coding Scheme* (GICS).

[19]The inclusion of time-sectorial dummies in the Poisson model imposes the restriction that,

$$n_{kt} = \sum_{j=1}^{J} \exp(\boldsymbol{\alpha}_{kt} + \boldsymbol{\theta}'\mathbf{y}_{jt} + \boldsymbol{\beta}'\mathbf{z}_{jkt}),$$

where n_{kt} is the total number of investments in sector k at time t. Given that $n_{jkt} = p_{jkt}.n_{kt}$ we may compute the marginal effects in terms of their impact on n_{jkt} or p_{jkt}. Our explanatory variables are all in logarithmic form what means that the estimated coefficients can be directly interpreted as elasticities if we measure the impact on n_{jkt}. To obtain the elasticities in terms of p_{jkt} one should multiply the estimated coefficients by $(1 - p_{jkt})$.

Table 1: Plant Births by Sector

SIC Code	Industry	Plant Births 1989 Number	%	1997 Number	%
20	Food and Kindred Products	1756	5.0	1513	5.1
21	Tobacco Products	15	0.0	15	0.1
22	Textile Mill Products	572	1.6	486	1.6
23	Apparel and Others Textile Products	3319	9.4	2704	9.1
24	Lumber and Wood Products	4058	11.5	3461	11.6
25	Furniture and Fixtures	1481	4.2	953	3.2
26	Paper and Allied Products	397	1.1	346	1.2
27	Printing and Publishing	6273	17.8	4375	14.7
28	Chemicals and Allied Products	1048	3.0	997	3.3
29	Petroleum and Coal Products	206	0.6	287	1.0
30	Rubber and Miscellaneous Plastic Products	1360	3.8	1049	3.5
31	Leather and Leather Products	167	0.5	147	0.5
32	Stone, Clay, and Glass Products	1400	4.0	1110	3.7
33	Primary Metal Industries	526	1.5	414	1.4
34	Fabricated Metal Products	3096	8.8	2264	7.6
35	Industrial Machinery and Equipment	3696	10.5	4258	14.3
36	Electronic and Other Electrical Equipment	1845	5.2	1598	5.4
37	Transportation Equipment	1128	3.2	1193	4.0
38	Instruments and Related Products	808	2.3	802	2.7
39	Miscellaneous Manufacturing Industries	2186	6.2	1849	6.2
Total		35337	100.0	29821	100.0

Table 2: Plant Births by State

State	Plant Births			
	1989		1997	
	Number	%	Number	%
Alabama	702	2.0	544	1.8
Arizona	559	1.6	498	1.7
Arkansas	414	1.2	287	1.0
California	5680	16.1	4611	15.5
Colorado	617	1.7	557	1.9
Connecticut	424	1.2	332	1.1
Delaware	59	0.2	57	0.2
District of Columbia.	45	0.1	31	0.1
Florida	1975	5.6	1651	5.5
Georgia	1001	2.8	860	2.9
Idaho	227	0.6	202	0.7
Illinois	1328	3.8	1115	3.7
Indiana	750	2.1	598	2.0
Iowa	287	0.8	242	0.8
Kansas	278	0.8	269	0.9
Kentucky	406	1.1	339	1.1
Louisiana	425	1.2	362	1.2
Maine	211	0.6	226	0.8
Maryland	329	0.9	287	1.0
Massachusetts	716	2.0	594	2.0
Michigan	1112	3.1	1050	3.5
Minnesota	677	1.9	616	2.1
Mississippi	442	1.3	313	1.0
Missouri	670	1.9	470	1.6
Montana	185	0.5	178	0.6
Nebraska	148	0.4	116	0.4
Nevada	167	0.5	194	0.7
New Hampshire	222	0.6	203	0.7
New Jersey	999	2.8	785	2.6
New Mexico	208	0.6	173	0.6
New York	2476	7.0	2154	7.2
North Carolina	1079	3.1	865	2.9
North Dakota	60	0.2	70	0.2
Ohio	1297	3.7	1064	3.6
Oklahoma	394	1.1	348	1.2
Oregon	706	2.0	569	1.9
Pennsylvania	1333	3.8	1068	3.6
Rhode Island	221	0.6	151	0.5
South Carolina	456	1.3	439	1.5
South Dakota	101	0.3	71	0.2
Tennessee	710	2.0	561	1.9
Texas	2293	6.5	2164	7.3
Utah	277	0.8	294	1.0
Vermont	145	0.4	140	0.5
Virginia	461	1.3	449	1.5
Washington	1013	2.9	750	2.5
West Virginia	189	0.5	156	0.5
Wisconsin	789	2.2	687	2.3
Wyoming	74	0.2	61	0.2
Total	35337	100.0	29821	100.0

Table 3: Location Determinants of Manufacturing Plants Births in the US Counties

Variables	Conditional Logit Model / Poisson Regression				
	Without County Effects		With County Effects		
			Random Effects		Fixed Effects
Labor Costs	-0.4643^{*}	-0.4393^{*}	-0.2225^{*}	-0.2504^{*}	-0.1308
	$[-13.818]$	$[-10.938]$	$[-3.047]$	$[-3.336]$	$[-0.820]$
Land Costs	-0.8061^{*}	-0.7058^{*}	-0.4989^{*}	-0.4199^{*}	0.2557
	$[-46.735]$	$[-35.084]$	$[-13.198]$	$[-10.077]$	$[1.635]$
Taxes	-0.2606^{*}	-0.3395^{*}	-0.1602^{*}	-0.1130^{*}	-0.1106^{*}
	$[-28.880]$	$[-22.683]$	$[-8.993]$	$[-4.575]$	$[-3.166]$
Market Size	1.0129^{*}	0.9266^{*}	0.8417^{*}	0.7496^{*}	-0.0616
	$[227.407]$	$[145.482]$	$[52.052]$	$[36.949]$	$[-0.463]$
Localization Economies	0.1345^{*}	0.1283^{*}	0.1199^{*}	0.1192^{*}	0.1172^{*}
	$[52.697]$	$[52.494]$	$[57.512]$	$[57.126]$	$[57.768]$
Urbanization Economies	0.5563^{*}	0.5444^{*}	0.2816^{*}	0.3027^{*}	0.2006^{***}
	$[34.571]$	$[30.242]$	$[7.557]$	$[7.921]$	$[1.881]$
State Dummies	No	Yes	No	Yes	-
Log-Likelihood	-73696.38	-72543.9	-70680.7	-70472.7	-62211.92
$N_{obs} = J \times K \times T$	122520	122520	122520	122520	109560
δ (χ^2)	-	-	0.1663 (6031.4)	0.1308 (4142.4)	-