On the Equivalence of Location Choice Models: Conditional Logit, Nested Logit and Poisson*

Kurt Schmidheiny‡ — *Universitat Pompeu Fabra*

Marius Brülhart§ — *University of Lausanne*

July 20th, 2010

Published in the

Journal of Urban Economics 69(2), 2011, 214-222

Abstract

It is well understood that the two most popular empirical models of location choice - conditional logit and Poisson - return identical coefficient estimates when the regressors are not individual specific. We show that these two models differ starkly in terms of their implied predictions. The conditional logit model represents a zero-sum world, in which one region's gain is the other regions' loss. In contrast, the Poisson model implies a positive-sum economy, in which one region's gain is no other region's loss. We also show that all intermediate cases can be represented as a nested logit model with a single outside option. The nested logit turns out to be a linear combination of the conditional logit and Poisson models. Conditional logit and Poisson elasticities mark the polar cases and can therefore serve as boundary values in applied research.

JEL Classification: C25, R3, H73

Keywords: firm location, residential choice, conditional logit, nested logit, Poisson count model

---

*We are grateful to Paulo Guimaraes and his coauthors for allowing us to use their data set, and to participants in seminars at Pompeu Fabra and at the University of Barcelona as well as at the 2009 ETSG annual conference in Rome and at the 2009 annual conference of the Urban Economics Association in San Francisco for useful comments. Financial support from the Spanish Ministry of Science and Innovation (Ramon y Cajal and Consolider SEJ2007-64340), the Swiss National Science Foundation (grants 100012-113938 and PDFMP1-123133), and from the EU's Sixth Framework Program ("Micro-Dyn" project) is gratefully acknowledged.

‡Department of Economics and Business, Universitat Pompeu Fabra, Ramon Trias Fargas 25-27, 08005 Barcelona, Spain; e-mail: `kurt.schmidheiny@upf.edu`. Also affiliated with the Barcelona Graduate School of Economics, the Centre for Economic Policy Research (CEPR) and CESifo.

§Département d'économétrie et économie politique, Ecole des HEC, Université de Lausanne, CH - 1015 Lausanne, Switzerland. E-mail: `Marius.Brulhart@unil.ch`. Also affiliated with the Centre for Economic Policy Research (CEPR).

1 Introduction

Location choices by households and firms are of interest to economists for numerous reasons, ranging from the determinants of residential segregation patterns in cities to the design of national tax policy. Given the discrete nature of such choices, they are typically modelled by empirical researchers through McFadden's (1974) conditional logit framework.[1] The appeal of this approach lies in its formal link between the theoretical objective function of a representative location-seeking agent and the likelihood function of the empirical model.

Mostly out of a perception of greater computational ease, researchers have resorted to Poisson count estimation as an alternative approach to the conditional logit.[2] Guimaraes, Figueiredo and Woodward (2003), henceforth GFW, have shown that, with purely location-specific locational determinants or with determinants that are specific to locations and to groups of agents, the conditional logit and Poisson estimators return identical parameter estimates. In this sense, the two estimators are equivalent, and the rigorous link to the theory offered by the conditional logit model therefore applies identically to the Poisson. This useful result has already been applied widely in the location choice literature.[3]

We show that the identical coefficient estimates resulting from the two estimation strategies in fact have fundamentally different economic implications. The conditional logit model implies that the aggregate number of agents is fixed and that differences across locations affect only the distribution of those agents across those locations. Hence, an additional agent attracted to location j means one less agent among the other locations in the relevant set, $i \neq j$. In the Poisson model, however, an additional agent attracted to location j has no impact on the number of agents in the remaining locations and thus raises the aggregate number of agents, summed across i and j, by one. Thus, the conditional logit model and the Poisson model can be viewed as polar cases, with the former representing zero-sum reallocations of firms or households across locations and the latter implying a positive-sum world.[4]

We also show that intermediate cases between these two extremes can be represented by a nested logit model featuring a generic outside option. This approach returns the same parameter estimates as the two other estimators. The nested logit in fact can be written

[1]For studies of firm location choices using the conditional logit approach, see, e.g., Carlton (1983), Bartik (1985), Head, Ries and Swenson (1995), Guimaraes, Figueiredo and Woodward (2000), Head and Mayer (2004), Crozet, Mayer and Mucchielli (2004), and Devereux, Griffith and Simpson (2007). For corresponding studies of household location chocies, see, e.g., Ellickson (1981), Quigley (1985), Nechyba and Strauss (1998), Schmidheiny (2006), and Bayer, Ferreira and McMillan (2007).

[2]See, e.g., Papke (1991), Becker and Henderson (2000), List (2001), Guimaraes, Figueiredo and Woodward (2004), and Holl (2004), for studies of firms' location choices; and Flowerdew and Aitkin (1982) and Börsch-Supan (1990) for studies of individual migration choices.

[3]Studies invoking the GFW equivalence result include Holl (2004), Duranton, Gobillon and Overman (2006), Brülhart, Jametti and Schmidheiny (2007), Arzaghi and Henderson (2008), Davis and Henderson (2008), and Coeurdacier, De Santis and Aviat (2009).

[4]For recent research on the cross-region effects of region-specific policies aimed at attracting firms, see, e.g., Greenstone and Moretti (2003), Chirinko and Wilson (2008), and Wilson (2009).

as a linear combination of the conditional logit and Poisson models, with a single "rivalness parameter" representing the closeness of the nested logit to the conditional logit (and thus the distance from the Poisson). Conditional logit and Poisson elasticities mark the polar cases and can therefore serve as boundary values in applied research.

The paper is structured as follows. In Section 2, we formally derive the commonalities and differences among the conditional logit, Poisson and nested logit models. Empirical implications and an illustration are presented in Section 3. Section 4 concludes.

2 The models

We denote agents with $f = 1, ..., N$ and regions with $j = 1, ..., J$. For simplicity, we shall frame our discussion in terms of corporate location decisions, and therefore relate to f as "firms".

Following GFW, we first assume the determinants of locational attractiveness to be purely region specific, such that they affect all firms symmetrically (case A). The K observable characteristics of each region are given by the $(K \times 1)$ vector x_j. We shall later relax this assumption, and allow locational attractiveness to be region-industry specific (case B).

The random variable n_j represents the count of firms in region j, whereas N_j denotes the number of firms actually observed in region j. Analogously, the random variable n represents the total number of firms, whereas N denotes the observed total number of firms.

2.1 Case A: industry-invariant locational determinants

2.1.1 Conditional logit

Suppose that firm f's profit in region j is determined by the linear model $\pi_{fj} = x_j'\beta + \varepsilon_{fj}$, where β is a $(K \times 1)$ vector of coefficients. Then, the conditional logit model is defined by the assumption that the random term ε_{fj} is independent across f and j and follows an extreme-value type 1 distribution. With this assumption, the probability that a given firm f chooses region j rather than another region is given by

$$P_{j|f} = P_j = \frac{e^{x_j'\beta}}{\sum_{i=1}^{J} e^{x_i'\beta}}, \tag{1}$$

where $\sum_j P_{j|f} = 1$ for all f. Since locational characteristics x_j are assumed here to affect all firms symmetrically, this probability also represents the share of firms that will choose region j.

The parameter β can be estimated by maximum likelihood. We can write the log likelihood as

$$\log L(\beta) = \sum_{j=1}^{J} N_j \log P_j = \sum_{j=1}^{J} N_j x_j'\beta - \sum_{j=1}^{J} [N_j \log(\sum_{i=1}^{J} e^{x_i'\beta})]. \tag{2}$$

The conditional logit model implicitly assumes that the total number of firms n is given and does not depend on the locational characteristics x. The expected number of firms in region j, $E(n_j)$ is therefore simply

$$E(n_j) = nP_j = n\frac{e^{x_j'\beta}}{\sum_{i=1}^{J} e^{x_i'\beta}}. \tag{3}$$

The percentage change in the expected number of firms in region j, $E(n_j)$, with respect to a unit change in the k-th locational characteristic of region j itself is given respectively by the *own-region semi-elasticity*:

$$\epsilon_{jj} = \frac{\partial \log E(n_j)}{\partial x_{jk}} = (1 - P_j)\beta_k. \tag{4}$$

Similarly, the percentage change in the expected number of firms in another region, $E(n_{i\neq j})$, with respect to a unit change in j's k-th locational characteristic is given by the *cross-region semi-elasticity*:

$$\epsilon_{ij} = \frac{\partial \log E(n_i)}{\partial x_{jk}} = -P_j\beta_k. \tag{5}$$

For simplicity, we shall henceforth refer to these and all subsequently presented semi-elasticities as "elasticities". Hence, all "elasticities" derived and calculated in this paper in fact are semi-elasticities.

The own-region elasticity (4) shows that by enhancing its attractiveness a region will increase its expected number of firms, and the cross-region elasticity (5) implies that one region's increased attractiveness to firms will reduce the number of firms choosing other regions: one region's gain is another region's loss. Moreover, a simple comparison of the two elasticities shows that small regions (in terms of P_j, the share of firms they host) are predicted by the conditional logit model to find their own firm counts to be relatively elastic to changes in their own locational characteristics, while not affecting firm counts in other regions as much as large regions.

We now turn from the viewpoint of individual regions to an analysis of what the conditional logit model implies for the total number of firms among the J regions. By definition,

$$E(n) = \sum_{j=1}^{J} E(n_j) = n = N. \tag{6}$$

Hence, the expected total number of firms is equal to the observed total, N, irrespective of regressors and parameters. This again shows the "zero sum" aspect of the conditional logit model, where the implied problem is one of allocating an exogenously fixed number of firms over a set of regions. It also follows logically that changes in the locational attractiveness of individual regions will not affect the total number of firms. Formally, the elasticity of the expected total firm count relative to a change in one of the K locational characteristics of any particular region j is zero:

$$\epsilon_j = \frac{\partial \log E(n)}{\partial x_{jk}} = 0.$$

2.1.2 Poisson

The Poisson estimator is based on the assumption that n_j is independently Poisson distributed with region-specific mean

$$E(n_j) = e^{\alpha + x_j'\beta}. \tag{7}$$

Here too, β can be estimated by maximum likelihood. We can write the concentrated log likelihood as

$$\log L(\beta) = \sum_{j=1}^{J} N_j x_i'\beta - \sum_{j=1}^{J} [N_j \log(\sum_{i=1}^{J} e^{x_i'\beta})] - \sum_{j=1}^{J} \log N_j! - N + N \log N. \tag{8}$$

When comparing this to (2), the point made by GFW is plain to see: the log likelihood functions of the two models are identical up to a constant, and maximum likelihood estimation therefore yields identical parameter estimates $\hat{\beta}$.

In expectations, the share of firms in region j can be written as

$$P_j = \frac{E(n_j)}{\sum_{i=1}^{J} E(n_j)} = \frac{e^{\alpha + x_j'\beta}}{\sum_{i=1}^{J} e^{\alpha + x_i'\beta}} = \frac{e^{x_j'\beta}}{\sum_{i=1}^{J} e^{x_i'\beta}}, \tag{9}$$

which is exactly the same expression as (1), for the conditional logit model. This equivalence lies at the heart of the GFW result:

Observation 1 (Guimaraes, Figueiredo and Woodward, 2003) *The log likelihood functions for the conditional logit and the Poisson model are identical up to a constant, and maximum likelihood estimation therefore yields identical parameter estimates $\hat{\beta}$.*

The elasticity of the expected number of firms in region j, $E(n_j)$, with respect to a change in the k-th locational characteristic of region j itself, and by that of another region $i \neq j$, is given respectively by the *own-region elasticity*

$$\epsilon_{jj} = \frac{\partial \log E(n_j)}{\partial x_{jk}} = \beta_k \tag{10}$$

and by the *cross-region elasticity*

$$\epsilon_{ij} = \frac{\partial \log E(n_i)}{\partial x_{jk}} = 0. \tag{11}$$

Comparing these elasticities to their conditional-logit equivalents (4) and (5), we observe the following differences.

Observation 2 *The Poisson model implies more elastic responses by firm counts to given changes in own-region characteristics than the conditional logit model.*

Observation 3 *Unlike in the conditional logit model, in the Poisson model, one region's change in locational attractiveness has no impact on the number of firms located among any of the $J-1$ other regions.*

Hence, even though the estimated parameters $\hat{\beta}$ will be invariant to the choice of model, their implied predictions differ starkly. The conditional logit model implies a zero-sum allocation process of a fixed number of firms over the J jurisdictions. In contrast, in the Poisson model new firms are non-rivalrous, in the sense that adjustment to one regions's locational characteristics works not through changes in firm numbers among the $J-1$ other regions but from changes either in the supply of local entrepreneurship or in firms attracted from or repelled to somewhere outside the considered set of J regions.

Moving again from the viewpoint of individual regions to an analysis of the model's implications for the total number of firms among the J regions, and using (7), we find that

$$E(n) = \sum_{i=1}^{J} E(n_i) = \sum_{i=1}^{J} e^{\alpha + x_i'\beta} = e^{\alpha} \sum_{i=1}^{J} e^{x_i'\beta}.$$

Comparing this expression with its conditional logit equivalent (6), we note that the expected total number of firms is now not generally equal to the observed total number of firms, N, but depends on the regressors and parameters.[5] The Poisson model thus implies that a change in a region's locational attractiveness will affect the sum of firms active in the J regions. Specifically, the elasticity of the expected total firm count with respect to a change in one of the K locational characteristics of any particular region j is given by[6]

$$\epsilon_j = \frac{\partial \log E(n)}{\partial x_{jk}} = \frac{e^{x_j'\beta}}{\sum_{i=1}^{J} e^{x_i'\beta}} \beta_k = \frac{E(n_j)}{E(n)} \beta_k = P_j \beta_k.$$

[5] Note that the predicted total number of firms *at the estimated coefficients and actual data* corresponds to the observed total of firms in the Poisson model just as it does in the conditional logit model. In symbols, $E(n|\hat{\alpha}, \hat{\beta}) = N$.

[6] We define $P_j \equiv E(n_j)/E(n) \neq E(n_j/n)$ in the Poisson model. Using this definition, $P_j = e^{x_j'\beta} / \sum_{i=1}^{J} e^{x_i'\beta}$ in both the conditional logit and the Poisson model.

Observation 4 *In the Poisson model, an increase (decrease) in one region's locational attractiveness increases (decreases) the total of firms summed across the J regions. In the conditional logit model, a change in one region's locational attractiveness leaves the total of firms summed across the J regions unchanged.*

2.1.3 Nested logit

Observations 2, 3 and 4 show that the two models are the polar cases of a continuum of relative adjustment margins, ranging from reallocations purely *within* the set of alternatives considered (conditional logit) to reallocations purely *between* that set and some outside option (Poisson). We now turn to a micro-founded approach that covers this whole continuum and thus encompasses the polar cases.

Suppose that firms make two sequential choices. At the first stage, they choose between locating in one of the J regions considered (which could stand for "domestic" regions) and an outside option $j = 0$ (which could stand for locating "abroad", or for remaining inactive). If they have chosen to set up in one of the J regions, they pick one of them at the second stage. Like in the conditional logit model, firm f's profit in region $j > 0$ is determined by a linear function of the region-specific characteristics x_j, such that $\pi_{fj} = x_j'\gamma + \nu_{fj}$. Firm f's profit associated with the outside option is given by $\pi_{f0} = \delta + \nu_{f0}$, where δ summarizes the exogenously fixed locational attractiveness of the outside option. The stochastic term ν_{f0} is assumed to follow a generalized extreme value distribution as in McFadden (1978).[7] This leads to a nested logit model with one degenerate "nest" that includes $j = 0$ only and one other "nest" that includes all regions $j > 0$. This two-stage structure assumes independence between ν_{f0} and ν_{fj} for all $j > 0$, and non-negative correlation $\left(1 - \lambda^2\right)$ across ν_{fj} for all $j > 0$; where $0 < \lambda \leq 1$, sometimes called the "log-sum" coefficient, measures the importance of the domestic nest as a whole relative to the outside option.

In this setting, the probability that a particular firm f chooses the outside option $j = 0$ is given by

$$P_0 = \frac{e^{\delta}}{e^{\delta} + (\sum_{j=1}^{J} e^{x_j'\beta})^{\lambda}}, \tag{12}$$

and the probability that it chooses a particular region $j > 0$ among the J regions of interest is

$$P_j = \frac{e^{x_j'\beta}(\sum_{i=1}^{J} e^{x_i'\beta})^{\lambda-1}}{e^{\delta} + (\sum_{i=1}^{J} e^{x_i'\beta})^{\lambda}} = P_{j>0} \cdot P_{j|j>0},$$

where we reparametrize $\beta = \gamma/\lambda$. This implies that, unlike in the conditional logit, the estimated regression parameters $\widehat{\beta}$ are not identical to the structural parameters of the underlying

[7] The specific density function assumed over $\{0, J\}$ is $F(\nu_f) = \exp\left[-(\sum_{j=1}^{J} e^{-\nu_{fj}/\lambda})^{\lambda} - e^{-\nu_{f0}}\right]$.

profit function.

The choice probabilities P_j can be decomposed into (a) the probability of choosing any of the J regions, $P_{j>0} = 1 - P_0$, and (b) the probability of choosing a specific region j given that the firm chooses to set up in one of the J regions,

$$P_{j|j>0} = \frac{e^{x_j'\beta}}{\sum_{i=1}^{J} e^{x_i'\beta}}. \tag{13}$$

The parameter β can again be estimated by maximum likelihood. We can write the concentrated log likelihood as

$$\log L(\beta) = \sum_{j=1}^{J} N_j x_i'\beta - \sum_{j=1}^{J}[N_j \log(\sum_{i=1}^{J} e^{x_i'\beta})] + N_0 \log(N_0) + N \log(N) - (N+N_0)\log(N+N_0),$$

where δ and λ cancel out when we substitute the first-order condition $\partial \log L/\partial\delta = 0$ into $L(\beta, \delta, \lambda)$.[8] N is the number of firms locating in any of the regions $j > 0$, and N_0 is the number of firms choosing the outside option $j = 0$.

Observation 5 *The log likelihood functions for the conditional logit, the Poisson and the nested logit model with a single outside option are identical up to a constant, and maximum likelihood estimation therefore yields identical parameter estimates $\hat{\beta}$.*

The ratio (13) is identical to its equivalents in the conditional logit and Poisson models, (1) and (9). This correspondence among the three models lies at the heart of Observation 5, which extends the GFW result to the nested logit case with a single outside option.

The parameter vector β is identified and can be estimated without observing N_0 because the first-order condition of the concentrated likelihood function, $\partial \log L(\beta)/\partial\beta$, is independent of N_0. The parameters δ and λ, however, are not identified when N_0 is not observed because the first-order conditions $\partial \log L(\beta, \delta, \lambda)/\partial\delta = 0$ and $\partial \log L(\beta, \delta, \lambda)/\partial\lambda = 0$ depend on N_0 (see footnote 8). Even if N_0 were observed, δ and λ would not be identified separately, as the two first-order conditions are identical.

[8] The concentrated likelihood is obtained as follows:

$$\begin{aligned} \log L(\beta, \delta, \lambda) &= N_0 \log P_0 + \sum_{j=1}^{J} N_j \log P_j \\ &= N_0 \log\left(\frac{e^{\delta}}{e^{\delta} + (\sum_{i=1}^{J} e^{\mathbf{x}_j'\beta})^{\lambda}}\right) + \sum_{j=1}^{J}\left[N_j \log\left(\frac{e^{x_j'\beta}(\sum_{i=1}^{J} e^{x_i'\beta})^{\lambda-1}}{e^{\delta} + (\sum_{i=1}^{J} e^{x_i'\beta})^{\lambda}}\right)\right]. \end{aligned}$$

The first-order condition with respect to δ is $\partial \log L/\partial\delta = N_0 - (N+N_0)e^{\delta}/[e^{\delta} + (\sum_{i=1}^{J} e^{x_i'\beta})^{\lambda}] = 0$. The estimated $\hat{\delta}$ can therefore be expressed as a function of the estimated $\hat{\beta}$ and $\hat{\lambda}$: $e^{\hat{\delta}} = N_0/N \cdot (\sum_{i=1}^{J} e^{x_i'\beta})^{\lambda}$. When we substitute $e^{\hat{\delta}}$ in $\log L(\beta, \delta, \lambda)$, λ also cancels out because the first-order condition with respect to λ is automatically satisfied as $\partial \log L/\partial\lambda = \partial \log L/\partial\delta = 0$

The expected number of firms in region $j > 0$ is

$$E(n_j) = (n + n_0)P_j = (n + n_0)\frac{e^{x'_j\beta}(\sum_{i=1}^J e^{x'_i\beta})^{\lambda-1}}{e^{\delta} + (\sum_{i=1}^J e^{x'_i\beta})^{\lambda}}. \tag{14}$$

The *own-region elasticity* of the expected number of firms, $E(n_j)$, relative to locational characteristics is given by

$$\epsilon_{jj} = \frac{\partial \log E(n_j)}{\partial x_{jk}} = [1 - P_{j|j>0}(1 - \lambda P_0)]\beta_k, \tag{15}$$

and the *cross-region elasticity* is given by

$$\epsilon_{ij} = \frac{\partial \log E(n_i)}{\partial x_{jk}} = -P_{j|j>0}(1 - \lambda P_0)\beta_k. \tag{16}$$

We can now compare the own- and cross-region elasticities of the three models. Simple inspection of elasticities (4), (5), (10), (11), (15) and (16) leads to the following observation.

Observation 6 *The nested logit own-region and cross-region elasticities lie between their conditional logit and Poisson counterparts.*

Once more, we now move from the analysis of firm counts in individual regions to the total number of firms that are active in the J regions. Using (14) and (12), we find that

$$E(n) = (n + n_0)\frac{(\sum_{j=1}^J e^{x'_j\beta})^{\lambda}}{e^{\delta} + (\sum_{j=1}^J e^{x'_j\beta})^{\lambda}} = (n + n_0)P_{j>0}.$$

The expected total number of firms active in the J regions is simply given by the share of potential firms that decide to become active in one of those regions. As in the Poisson model, the expected total number of firms is not generally equal to the observed total number of firms, N, but depends on the regressors and parameters, including those for the outside option.[9] The elasticity of the expected total firm count relative to a change in one of the K locational characteristics of any particular region j is given by

$$\epsilon_j = \frac{\partial \log E(n)}{\partial x_{jk}} = \frac{\lambda e^{\delta} e^{x'_j\beta}(\sum_{i=1}^J e^{x'_i\beta})^{-1}\beta_k}{e^{\delta} + (\sum_{i=1}^J e^{x'_i\beta})^{\lambda}} = \lambda P_0 P_{j|j>0}\beta_k.$$

Observation 7 *Like the Poisson, the nested logit model implies that a change in a region's locational attractiveness will affect the total of firms summed across the J regions.*

Here, the responsiveness of the aggregate firm number is due to the effect on the decisions

[9] As in the Poisson and conditional logit models, the predicted total number of firms among the J regions *at the estimated coefficients and actual data* corresponds to the observed total: $E(n|\hat{\beta}, \hat{\delta}, \hat{\lambda}) = N$.

Table 1: Comparing implied elasticities (case A)

	Conditional Logit	Nested Logit	Poisson
$\epsilon_{jj} = \frac{\partial \log E(n_j)}{\partial x_{jk}}$	$(1 - P_{j\|j>0})\beta_k$	$[1 - P_{j\|j>0}(1 - \lambda P_0)]\beta_k$	β_k
$\epsilon_{ij} = \frac{\partial \log E(n_i)}{\partial x_{jk}}$	$-P_{j\|j>0}\beta_k$	$-P_{j\|j>0}(1 - \lambda P_0)\beta_k$	0
$\epsilon_j = \frac{\partial \log E(n)}{\partial x_{jk}}$	0	$\lambda P_0 P_{j\|j>0}\beta_k$	$P_{j\|j>0}\beta_k$

Notes: $P_{j|j>0} = E(n_j)/E(n)$, $P_0 = E(n_0)/E(n + n_0)$

taken by firms that would have chosen the outside option in the absence of a change in regional attractiveness.

2.1.4 A synthesis of the three models

We can now pull together the salient features of the three models. First, we consider the impact of a change in the attractiveness of an individual region on the number of firms in that region and across the $J - 1$ remaining regions. Table 1 gathers the own-region, cross-region and aggregate elasticities implied by the three models.

In order to compare these elasticities, we define $\rho = 1 - \lambda P_0$ which satisfies $0 \geq \rho \geq 1$ under the standard nested logit assumption $0 < \lambda \leq 1$. We call ρ the *rivalness parameter*. It allows us to write the nested logit elasticities as a linear combination of their conditional logit and Poisson equivalents: $\epsilon_{jj}^{\text{nlogit}} = \rho\epsilon_{jj}^{\text{clogit}} + (1 - \rho)\epsilon_{jj}^{\text{Poisson}}$, $\epsilon_{ij}^{\text{nlogit}} = \rho\epsilon_{ij}^{\text{clogit}}$ and $\epsilon_{j}^{\text{nlogit}} = (1-\rho)\epsilon_{j}^{\text{Poisson}}$. The rivalness parameter therefore acts as a summary measure of the position of the data generating process between the two polar cases, conditional logit ($\rho = 1$) and Poisson ($\rho = 0$). One may think of ρ as capturing of the relative importance of the outside option: as $\rho \to 0$, competition among the J regions becomes unimportant relative to the weight of the outside option, while with $\rho \to 1$, the outside option becomes negligible and any reallocations have to occur within the set of the J regions.

We can also establish rankings of the elasticities implied by the three models. Provided that $\beta \neq 0$, the ranking of own-region elasticities is (c.f. Observations 2 and 6)

$$|\epsilon_{jj}^{\text{Poisson}}| > |\epsilon_{jj}^{\text{nlogit}}| > |\epsilon_{jj}^{\text{clogit}}| > 0,$$

while the ranking of cross-regions elasticities is just the reverse (c.f. Observations 3 and 6),

$$|\epsilon_{ij}^{\text{clogit}}| > |\epsilon_{ij}^{\text{nlogit}}| > |\epsilon_{ij}^{\text{Poisson}}| = 0.$$

and the ranking of aggregate elasticities is again (c.f. Observations 4 and 7)

$$|\epsilon_j^{\text{Poisson}}| > |\epsilon_j^{\text{nlogit}}| > |\epsilon_j^{\text{clogit}}| > 0.$$

An alternative way of comparing the three models is to inspect the predicted counts.[10] In all three models the expected number of firms in region j can be written as

$$E(n_j) = m \cdot e^{x'_j \beta}$$

where the multiplier m is given by

$$\begin{aligned}
m &= \frac{n}{\sum_{i=1}^{J} e^{x'_i \beta}} && \text{in the conditional logit,} \\
m &= e^{\alpha} = constant && \text{in the Poisson,} \\
m &= \frac{n + n_0}{\sum_{i=1}^{J} e^{x'_i \beta}} \cdot \frac{(\sum_{i=1}^{J} e^{x'_i \beta})^{\lambda}}{e^{\delta} + (\sum_{i=1}^{J} e^{x'_i \beta})^{\lambda}} && \text{in the nested logit.}
\end{aligned}$$

Nevertheless, given the data, the three models predict the same number of firms, as the multiplier is estimated as $\hat{m} = N / \sum_{i=1}^{J} e^{x'_i \hat{\beta}}$ in all three cases.[11] This reflects the GFW equivalence in Observation 1 and our equivalence in Observation 5. However, the multiplier m reacts differently to changes in locational characteristics depending on the estimator chosen, in line with our Observations 2, 3, 4, 6 and 7.

2.2 Case B: industry-specific locational determinants

Consider now that we observe K characteristics x_{sj} for every region j and industry s. Hence, we again do not observe firm-specific regional attributes, but we now allow for these attributes to differ across *groups* of firms, best thought of as industries. We maintain the notation x_j for the subset of locational determinants that are constant across industries. Furthermore, n_{js} is the number of firms in region j and industry s, n_s is the observed number of industry-s firms across all regions, n is the total number of firms, and N stands for the corresponding *observed* firm count in the sample.

The *grouped conditional logit* model is given by the probability that a given firm f of industry s chooses region j rather than another region:

$$P_{j|f} = P_{j|s} = \frac{e^{x'_{sj}\beta}}{\sum_{i=1}^{J} e^{x'_{si}\beta}},$$

[10]We thank an anonymous referee for suggesting this elegant alternative approach.

[11]This is easily verified by plugging in the respective first-order conditions of the unconcentrated likelihood functions: $e^{\hat{\alpha}} = N / \sum_{i=1}^{J} e^{x'_i \hat{\beta}}$ and $e^{\hat{\delta}} = N_0 / N (\sum_{i=1}^{J} e^{x'_i \hat{\beta}})^{\hat{\lambda}}$.

where $\sum_j P_{j|f} = 1$. $P_{j|s}$ is the probability for a particular firm to choose region j given that the firm belongs to industry s.

The *grouped Poisson* model is given by

$$E(n_{sj}) = e^{\alpha_s + x'_{sj}\beta},$$

where α_s is an industry-specific constant.

Finally, the *grouped nested logit* model is given by the probability that a given firm f of industry s chooses either the outside option $j = 0$,

$$P_{0|s} = \frac{e^{\delta_s}}{e^{\delta_s} + (\sum_{j=1}^{J} e^{x'_{sj}\gamma/\lambda})^{\lambda}} = \frac{e^{\delta_s}}{e^{\delta_s} + (\sum_{j=1}^{J} e^{x'_{sj}\beta})^{\lambda}},$$

or a particular domestic region $j > 0$,

$$P_{j|s} = \frac{e^{x'_{sj}\beta}(\sum_{i=1}^{J} e^{x'_{si}\beta})^{\lambda-1}}{e^{\delta_s} + (\sum_{i=1}^{J} e^{x'_{si}\beta})^{\lambda}} = P_{j>0|s} \cdot P_{j|j>0,s} = (1 - P_{0|s})P_{j|j>0,s},$$

where δ_s is an industry-specific constant, and $\beta = \gamma/\lambda$. $P_{j>0|s} = 1 - P_{0|s}$ is the probability that a given industry-s firm chooses any domestic region $j > 0$, and

$$P_{j|j>0,s} = \frac{e^{x'_{sj}\beta}}{\sum_{i=1}^{J} e^{x'_{si}\beta}} \tag{17}$$

is the probability that such a firm chooses a particular domestic region conditional on not choosing the outside option.

As in case A, the three models are observationally equivalent in a cross section of domestic firm choices and yield identical estimates for the parameter vector β. This has been shown by GFW for the grouped conditional logit and the grouped Poisson models, and we show it in the Appendix for the grouped nested logit model.

Table 2 summarizes the implied elasticities in the three grouped models (see the Appendix for derivations). As in case A, the elasticities in the grouped nested logit model are (industry-specific) linear combinations of their conditional logit and Poisson equivalents: $\epsilon_{..}^{\text{nlogit}} = \rho_s \epsilon_{..}^{\text{clogit}} + (1 - \rho_s)\epsilon_{..}^{\text{Poisson}}$ where $\rho_s = 1 - \lambda P_{0|s}$.

3 Estimation

3.1 Elasticity bounds

We have shown that estimation of any of the three models will yield identical parameter estimates $\widehat{\beta}$. The additional parameters λ and δ in the nested logit model are not identified but

Table 2: Comparing implied elasticities (case B)

		Conditional Logit	Nested Logit	Poisson
Region-industry specific regressor x_{sjk}:				
(a)	$\frac{\partial \log E(n_{sj})}{\partial x_{sjk}}$	$\beta_k(1-P_{j\|j>0,s})$	$\beta_k[1-P_{j\|j>0,s}(1-\lambda P_{0\|s})]$	β_k
(b)	$\frac{\partial \log E(n_{si})}{\partial x_{sjk}}$	$-\beta_k P_{j\|j>0,s}$	$-\beta_k P_{j\|j>0,s}(1-\lambda P_{0\|s})$	0
(c)	$\frac{\partial \log E(n_{s})}{\partial x_{sjk}}$	0	$\beta_k P_{j\|j>0,s}\lambda P_{0\|s}$	$\beta_k P_{j\|j>0,s}$
(d)	$\frac{\partial \log E(n_{j})}{\partial x_{sjk}}$	$\beta_k(1-P_{j\|j>0,s})P_{s\|j}$	$\beta_k[1-P_{j\|j>0,s}(1-\lambda P_{0\|s})]P_{s\|j}$	$\beta_k P_{s\|j}$
(e)	$\frac{\partial \log E(n_{i})}{\partial x_{sjk}}$	$-\beta_k P_{j\|j>0,s}P_{s\|i}$	$-\beta_k P_{j\|j>0,s}(1-\lambda P_{0\|s})P_{s\|i}$	0
(f)	$\frac{\partial \log E(n)}{\partial x_{sjk}}$	0	$\beta_k P_{s\|j}P_j\lambda P_{0\|s}$	$\beta_k P_{s\|j}P_{j\|j>0}$
Region specific regressor x_{jk}:				
(g)	$\frac{\partial \log E(n_{j})}{\partial x_{jk}}$	$\beta_k \sum_{s=1}^{S}(1-P_{j\|j>0,s})P_{s\|j}$	$\beta_k \sum_{s=1}^{S}[1-P_{j\|j>0,s}(1-\lambda P_{0\|s})]P_{s\|j}$	β_k
(h)	$\frac{\partial \log E(n_{i})}{\partial x_{jk}}$	$-\beta_k \sum_{s=1}^{S} P_{j\|j>0,s}P_{s\|i}$	$-\beta_k \sum_{s=1}^{S} P_{j\|j>0,s}(1-\lambda P_{0\|s})P_{s\|i}$	0
(i)	$\frac{\partial \log E(n)}{\partial x_{jk}}$	0	$\beta_k P_j \sum_{s=1}^{S}(\lambda P_{0\|s}P_{s\|j})$	$\beta_k P_{j\|j>0}$

Notes: $P_{j|j>0,s} = E(n_{sj})/E(n_s)$, $P_{0|s} = E(n_{s0})/E(n_s + n_{s0})$, $P_{s|j} = E(n_{sj})/E(n_j)$. $P_{s|j}$ is the fraction of firms in industry s in a given region j.

irrelevant for the estimation of β. Hence, it is impossible to discriminate formally between these three model based on cross-section data. And yet, the implied elasticities differ substantially. In previous research, reported elasticities were based either on the conditional logit model or the Poisson model, without justification of the particular choice made or, mistakenly in this respect, by referring to the equivalence of the two models as established by GFW.

What can researchers do if they are not willing to make this choice by assumption but rely on cross-sectional data? We propose in this situation that one calculate the elasticities of both the conditional logit and the Poisson model and report these predictions as bounds for the true effects. As shown in Observation 6, intermediate values can be rationalized by a nested logit model.

The computation of both conditional logit and Poisson elasticities requires that one calculate predicted probabilities. In terms of case A (Table 1), the predicted probability is obtained as follows:

$$\hat{P}_{j|j>0} = \frac{e^{x_j'\hat{\beta}}}{\sum_{i=1}^{J} e^{x_i'\hat{\beta}}}\,, \tag{18}$$

while for case B (Table 2), three predicted probabilities are to be computed:

$$
\begin{aligned}
\hat{P}_{j|j>0,s} &= \frac{e^{x'_{sj}\hat{\beta}}}{\sum_{i=1}^{J} e^{x'_{si}\hat{\beta}}}\,, && (19)\\
\hat{P}_{s|j} &= \frac{e^{x'_{sj}\hat{\beta}+\hat{\alpha}_s}}{\sum_{r=1}^{S} e^{x'_{rj}\hat{\beta}+\hat{\alpha}_r}} = \frac{N_s\hat{P}_{j|s}}{N\hat{P}_{j|j>0}}\,, && (20)\\
\hat{P}_{j|j>0} &= \frac{\sum_{s=1}^{S} e^{x'_{sj}\hat{\beta}+\hat{\alpha}_s}}{\sum_{s=1}^{S}\sum_{i=1}^{J} e^{x'_{si}\hat{\beta}+\hat{\alpha}_s}} = \frac{\sum_{s=1}^{S} N_s\hat{P}_{j|s}}{N}. && (21)
\end{aligned}
$$

where $\hat{\alpha}_s = \log[N_s/(\sum_j e^{x'_{js}\hat{\beta}})]$ from the first-order condition of the unconcentrated Poisson likelihood function.

3.2 An example

By way of an illustration, we take the data on location choices in Portugal by foreign-owned plants used in Guimaraes *et al.* (2000, 2003), and we report the elasticities implied by the coefficients of their regression model. The data cover a cross section of 758 location choices among 275 Portuguese regions by firms belonging to one of 151 industries.[12] Their region-industry level regressor of main interest, x_{sjk}, is "industry-specific agglomeration", defined as the share of regional employment in the same industry as the relevant firm. Their region level regressor of main interest, x_{jk}, is "total manufacturing agglomeration", defined as the log of aggregate manufacturing employment per square kilometer.

Taking their estimated parameters and computing the empirical probabilities (18)-(21), we can calculate all the implied elasticities of Table 2. Since the probabilities (18)-(21) vary by region and industry, we need to select specific cases for the computation of elasticities. We provide illustrations for two base regions j: Lisbon, the largest region in terms of $\hat{P}_{j|j>0}$, and Oleiros, the smallest region in terms of $\hat{P}_{j|j>0}$ that still had non-zero firm counts in the larger industry considered.[13]

Table 3 shows the implied elasticities for changes in a region-industry specific regressor and in a region specific regressor. We can take these estimates to illustrate Observations 2 to 4.[14]

[12]We follow GFW by referring to industry-year pairs as "industries". The 151 industries in their data set are combinations of 27 three-digit manufacturing sectors and seven sample years, ranging from 1985 to 1991.

[13]Where a comparison region i needs to be specified for the computation of cross elasticities, we choose Porto, the second largest region in the data set. Where an industry s needs to be specified, we choose Industrial Chemicals (ISIC 351) in 1989, the largest sector-year pair in the dataset (31 observed choices, i.e. 4 percent of the total of 758 choices).

[14]A note on the estimation of standard errors. In the Poisson model with group (industry) fixed effects, large sample properties are usually derived assuming a large number of groups, $S \to \infty$. In the conditional logit model, one typically assumes a large number of individuals, $N \to \infty$. The conventional standard errors will therefore in practice differ between the two models. Clustering at the group level, however, will produce identical standard errors. Such robust standard errors can either be estimated using asymptotic theory (cluster generalization of Eicker-Huber-White) or through block-wise bootstrapping.

Table 3: Comparing implied elasticities in an example of case B

		large region j		small region j	
		CL	Poisson	CL	Poisson
Region-industry specific regressor x_{sjk}:					
(a)	$\frac{\partial \log E(n_{sj})}{\partial x_{sjk}}$	2.9290	3.1883	3.18778	3.18832
		(0.1978)	(0.2364)	(0.23632)	(0.23640)
(b)	$\frac{\partial \log E(n_{si})}{\partial x_{sjk}}$	-0.2593	0	-0.00054	0
		(0.0815)	–	(0.00014)	–
(c)	$\frac{\partial \log E(n_{s})}{\partial x_{sjk}}$	0	0.2593	0	0.00054
		–	(0.0815)	–	(0.00014)
(d)	$\frac{\partial \log E(n_{j})}{\partial x_{sjk}}$	0.1110	0.1208	0.09624	0.09626
		(0.0110)	(0.0127)	(0.01137)	(0.01137)
(e)	$\frac{\partial \log E(n_{i})}{\partial x_{sjk}}$	-0.0095	0	-0.00002	0
		(0.0031)	–	(0.00001)	–
(f)	$\frac{\partial \log E(n)}{\partial x_{sjk}}$	0	0.0106	0	0.00002
		–	(0.0019)	–	(0.00000)
Region specific regressor x_{jk}:					
(g)	$\frac{\partial \log E(n_{j})}{\partial x_{jk}}$	0.3873	0.42577	0.42548	0.42577
		(0.0475)	(0.0508)	(0.05075)	(0.05076)
(h)	$\frac{\partial \log E(n_{i})}{\partial x_{jk}}$	-0.0378	0	-0.00004	0
		(0.0048)	–	(0.00001)	–
(i)	$\frac{\partial \log E(n)}{\partial x_{jk}}$	0	0.0374	0	0.00010
		–	(0.0066)	–	(0.00001)

Notes: large region: j = Lisbon; small region: j = Oleiros; i = Porto in rows (e) and (h); k = "industry-specific agglomeration" in rows (a) to (f), k = "total manufacturing agglomeration" in rows (g) to (i); industry: s = ISIC 351 (Industrial Chemicals) in 1989 in rows (a) to (f); any industry in rows (g) to (i). Bootstrapped robust standard errors in parantheses, 300 replications, clustered by industries.

- Observation 2: Own-region elasticities are larger in the Poisson model than in the conditional logit (rows (a), (d) and (g)). We can see that the difference between implied own-region elasticities is non-trivial for the large region (some 10 percent) but very small for the small region (less than 0.1 percent). This illustrates that the difference between implied own-region elasticities of the two models vanishes as the number of regions grows large and individual regions therefore become small.

- Observation 3: All Poisson cross-region elasticities are zero (rows (b), (e) and (h)).

- Observation 4: In the conditional logit model, the total number of firms (across all of Portugal) is invariant to changes in the values of x_{sjk} or x_{jk} whereas in the Poisson model the total changes with x_{sjk} or x_{jk} (rows (c), (f) and (i)). The effect on the total number of firms of a given change in x_{sjk} is stronger if the change occurs in a large region.

These computations illustrate the qualitatively different predictions implied by the conditional logit and Poisson models. With $J = 275$ spatial alternatives and $s = 151$ industries, the underlying data set is highly disaggregated, implying relatively modest quantitative differences between implied elasticities. Nonetheless, even here some of the differences are far from negligible. Perhaps the most striking difference appears in row (c) of Table 3. A one-unit change in x_{sjk} of Lisbon leaves the number of Portuguese plants in industry s unchanged in the conditional logit framework, while it increases by up to 29 percent in the Poisson model. Policy makers ought not to ignore a difference of such magnitude.

4 Conclusions

We show that the three standard location choice models - conditional logit, nested logit and Poisson - are observationally equivalent in terms of cross-section estimation yet imply starkly different predictions.

Take a corporate tax cut in a particular region. Provided that this is perceived by firms as making that region more attractive, all three models imply that the region itself will see an increase in its number of firms. We show that the magnitude of the implied increase differs: it is largest if the world is properly represented by the Poisson model, smallest if the world conforms with the conditional logit, and somewhere in-between if the world is nested logit. In a Poisson world, the tax cut will have no impact on firm counts in any other of regions within the data set. It will, however, pull firms away from other regions in the conditional logit and the nested logit cases. As the total number of firms is fixed in the conditional logit, the sum of the firms pulled away from the other regions is the same as the increase in the number of firms in the tax-cutting region itself. The nested logit again represents an intermediate case, with some of the attracted firms relocating from elsewhere within the data set, implying that regional corporate tax bases are "rival"; and some firms appearing from outside that set, implying a "non-rival" tax base. The same logic can be applied to residential choices of private households with respect, for instance, to changes in local property tax rates.

Empirical researchers should be aware of the interpretational ambiguity affecting estimated parameters in standard location choice models, particularly if the number of locations and industries distinguished in the data is small. It can therefore be useful to report both conditional logit and Poisson elasticity estimates as bounds on the effects implied by the estimated parameters.

References

[1] Arzaghi, Mohammad and J. Vernon Henderson (2008) Networking off Madison Avenue. *Review of Economic Studies*, 75(4): 1011-1038.

[2] Bartik, Timothy J. (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.

[3] Bayer, Patrick; Fernando Ferreira and Robert McMillan (2007) A Unified Framework for Measuring Preferences for Schools and Neighborhoods. *Journal of Political Economy*, 115(4): 588-638.

[4] Becker, Randy and Vernon Henderson (2000) Effects of Air Quality Regulations on Polluting Industries. *Journal of Political Economy*, 108(2): 379-421.

[5] Börsch-Supan, Axel (1990) Education and its Double-Edged Impact on Mobility. *Economics of Education Review*, 9(1): 39-53.

[6] Brülhart, Marius; Mario Jametti and Kurt Schmidheiny (2007) Do Agglomeration Economies Reduce the Sensitivity of Firm Location to Tax Differentials? *CEPR Discussion Paper* #6609.

[7] Carlton, Dennis W. (1983) The Location and Employment Choices of New Firms: An Econometric Model with Discrete and Continuous Endogenous Variables. *Review of Economics and Statistics*, 65: 440-449.

[8] Chirinko, Robert S. and Daniel J. Wilson (2008) State Investment Tax Incentives: A Zero-Sum Game? *Journal of Public Economics*, 92(12): 2362-2384.

[9] Coeurdacier, Nicolas; Roberto A. De Santis and Antonin Aviat (2009) Cross-Border Mergers and Acquisitions and European Integration. *Economic Policy*, 24(57): 55-106.

[10] Crozet, Matthieu; Thierry Mayer and Jean-Louis Mucchielli (2004) How Do Firms Agglomerate? A Study of FDI in France. *Regional Science and Urban Economics*, 34(1): 27-54.

[11] Davis, James C. and J. Vernon Henderson (2008) The Agglomeration of Headquarters. *Regional Science and Urban Economics*, 38(5): 445-460.

[12] Devereux, Michael P.; Rachel Griffith and Helen Simpson (2007) Firm Location Decisions, Regional Grants and Agglomeration Externalities. *Journal of Public Economics*, 91(3-4): 413-435.

[13] Duranton, Gilles; Laurent Gobillon and Henry G. Overman (2006) Assessing the Effects of Local Taxation Using Microgeographic Data. *CEPR Discussion Paper* #5856.

[14] Ellickson, Bryan (1981) An Alternative Test of the Hedonic Theory of Housing Markets. *Journal of Urban Economics*, 9(1): 56-79.

[15] Flowerdew, Robin and Murray Aitkin (1982) A Method of Fitting the Gravity Model Based on the Poisson Distribution. *Journal of Regional Science*, 22(2): 191-202.

[16] Greenstone, Michael and Enrico Moretti (2003) Bidding for Industrial Plants: Does Winning a 'Million Dollar Plant' Increase Welfare? *NBER Working Paper* #9844.

[17] Guimaraes, Paulo; Octávio Figueiredo and Douglas Woodward (2000) Agglomeration and the Location of Foreign Direct Investment in Portugal. *Journal of Urban Economics*, 47(1): 115-135.

[18] Guimaraes, Paulo; Octávio Figueiredo and Douglas Woodward (2003) A Tractable Approach to the Firm Location Decision Problem. *Review of Economics and Statistics*, 85(1): 201-204.

[19] Guimaraes, Paulo; Octávio Figueiredo and Douglas Woodward (2004) Industrial Location Modelling: Extending the Random Utility Framework. *Journal of Regional Science*, 44(1): 1-20.

[20] Head, Keith and Thierry Mayer (2004) Market Potential and the Location of Japanese Investment in the European Union. *Review of Economics and Statistics*, 86(2): 959-972.

[21] Head, Keith; John Ries and Deborah L. Swenson (1995) Agglomeration Benefits and Location Choice: Evidence from Japanese Manufacturing Investment in the United States. *Journal of International Economics*, 38(3-4): 223-247.

[22] Holl, Adelheid (2004) Manufacturing Location and Impacts of Road Transport Infrastructure: Empirical Evidence from Spain. *Regional Science and Urban Economics*, 34(3): 341-363.

[23] List, John A. (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.

[24] McFadden, Daniel (1974) Conditional Logit Analysis of Qualitative Choice Behavior. In P. Zarembka (ed.) *Frontiers in Econometrics*, New York: Academic Press, 105-142.

[25] McFadden, Daniel (1978), Modelling the Choice of Residential Location. In A. Karlqvist et al. (ed.) *Spatial Interaction Theory and Planning Models*, Amsterdam: North-Holland.

[26] Nechyba, Thomas J. and Robert P. Strauss (1998) Community Choice and Local Public Services: A Discrete Choice Approach. *Regional Science and Urban Economics*, 28(1): 51-73.

[27] Papke, Leslie E. (1991) Interstate Business Tax Differentials and New Firm Location: Evidence from Panel Data. *Journal of Public Economics*, 45(1): 47-68.

[28] Quigley, John M. (1985) Consumer Choice of Dwelling, Neighborhood and Public Services. *Regional Science and Urban Economics*, 15(1): 41-63.

[29] Schmidheiny, Kurt (2006) Income Segregation and Local Progressive Taxation: Empirical Evidence from Switzerland. *Journal of Public Economics*, 90(3): 429-458.

[30] Wilson, Daniel J. (2009) Beggar thy Neighbor? The In-State, Out-of-State, and Aggregate Effects of R&D Tax Credits. *Review of Economics and Statistics*, 91(2): 431-436.

Appendix: Derivations for case B

Grouped conditional logit

The conditional logit model for *grouped data* is given by the probability that a given firm f of industry s chooses region j

$$P_{j|f} = P_{j|s} = \frac{e^{x'_{sj}\beta}}{\sum_{i=1}^{J} e^{x'_{si}\beta}}.$$

The log likelihood function is

$$\log L(\beta) = \sum_{s=1}^{S}\sum_{j=1}^{J} N_{sj} P_{j|s} = \sum_{s=1}^{S}\left\{\sum_{j=1}^{J} N_{sj} x'_{sj}\beta - \sum_{j=1}^{J}[N_{sj}\log\sum_{i=1}^{J} e^{x'_{si}\beta}]\right\}.$$

The expected number of firms in region j and industry s is

$$E(n_{sj}) = n_s P_{j|s} = \frac{n_s e^{x'_{sj}\beta}}{\sum_{i=1}^{J} e^{x'_{si}\beta}}.$$

and the corresponding own-region and cross-region elasticities within industry s are, respectively,

$$\frac{\partial \log E(n_{sj})}{\partial x_{sjk}} = (1 - P_{j|s})\beta_k,$$

$$\frac{\partial \log E(n_{si})}{\partial x_{sjk}} = -P_{j|s}\beta_k.$$

The expected number of firms in industry s is

$$E(n_s) = \sum_{j=1}^{J} E(n_{sj}) = n_s = N_s,$$

and the corresponding elasticity within industry s is

$$\frac{\partial \log E(n_s)}{\partial x_{sjk}} = 0.$$

The expected number of firms in region j is

$$E(n_j) = \sum_{s=1}^{S} E(n_{sj}) = \sum_{s=1}^{S} n_s P_{j|s} = \sum_{s=1}^{S} \frac{n_s e^{x'_{sj}\beta}}{\sum_{i=1}^{J} e^{x'_{si}\beta}}.$$

The corresponding own-region and cross-region elasticities are for a region-industry specific shock x_{sjk} are

$$\frac{\partial \log E(n_j)}{\partial x_{sjk}} = \frac{\partial \log E(n_{sj})}{\partial x_{sjk}} \cdot \frac{E(n_{sj})}{E(n_j)} = (1 - P_{j|s})P_{s|j}\beta_k,$$

$$\frac{\partial \log E(n_i)}{\partial x_{sjk}} = \frac{\partial \log E(n_{si})}{\partial x_{sjk}} \cdot \frac{E(n_{si})}{E(n_i)} = -P_{j|s}P_{s|i}\beta_k,$$

where $P_{s|j} = E(n_{sj})/E(n_j)$.
The own-region and cross-region elasticities for a region-specific shock x_{jk} are

$$\frac{\partial \log E(n_j)}{\partial x_{jk}} = \sum_{s=1}^{S}\left[\frac{\partial \log E(n_{sj})}{\partial x_{jk}} \cdot \frac{E(n_{sj})}{E(n_j)}\right] = \beta_k \sum_{s=1}^{S}(1 - P_{j|s})P_{s|j},$$

$$\frac{\partial \log E(n_i)}{\partial x_{jk}} = \sum_{s=1}^{S} \left[\frac{\partial \log E(n_{si})}{\partial x_{jk}} \cdot \frac{E(n_{si})}{E(n_i)} \right] = -\beta_k \sum_{s=1}^{S} P_{j|s} P_{s|i}.$$

The expected total number of firms in all regions and industries is

$$E(n) = \sum_{s=1}^{S} \sum_{j=1}^{J} E(n_{js}) = \sum_{s=1}^{S} E(n_s) = n,$$

and the corresponding elasticities for a region-industry specific shock x_{sjk} and a region-specific shock x_{jk} are, respectively,

$$\frac{\partial \log E(n)}{\partial x_{sjk}} = 0,$$

$$\frac{\partial \log E(n)}{\partial x_{jk}} = 0.$$

Grouped Poisson

The Poisson model for grouped data is given as

$$E(n_{sj}) = \lambda_{sj} = e^{\alpha_s + x'_{sj}\beta},$$

where α_s is an industry-specific constant. The concentrated log likelihood function is

$$\log L(\beta) = \sum_{s=1}^{S} \left\{ \sum_{j=1}^{J} N_{sj} x'_i \beta - \sum_{j=1}^{J} [N_{sj} \log(\sum_{i=1}^{J} e^{x'_{si}\beta})] - \sum_{j=1}^{J} \log N_{sj}! + N_s \log N_s \right\} - N.$$

In expecations, the share of firms in region j for any given industry s is given by

$$P_{j|s} = \frac{E(n_{sj})}{\sum_{i=1}^{J} E(n_{sj})} = \frac{e^{\alpha_s + x'_{sj}\beta}}{\sum_{i=1}^{J} e^{\alpha_s + x'_{si}\beta}} = \frac{e^{x'_{sj}\beta}}{\sum_{i=1}^{J} e^{x'_{si}\beta}}.$$

The own-region and cross-region elasticities within industry s are, respectively,

$$\frac{\partial \log E(n_{sj})}{\partial x_{sjk}} = \beta_k,$$

$$\frac{\partial \log E(n_{si})}{\partial x_{sjk}} = 0.$$

The expected number of firms in industry s is

$$E(n_s) = \sum_{i=1}^{J} E(n_{si}) = \sum_{i=1}^{J} e^{\alpha_s + x'_{si}\beta} = \sum_{i=1}^{J} E(n_{si}) = \sum_{i=1}^{J} e^{\alpha_s + x'_{si}\beta} = e^{\alpha_s} \sum_{i=1}^{J} e^{x'_{si}\beta},$$

and the corresponding elasticity within industry s is

$$\frac{\partial \log E(n_s)}{\partial x_{sjk}} = \frac{e^{x'_{sj}}}{\sum_{i=1}^{J} e^{x'_{si}\beta}} \beta_k = P_{j|s} \beta_k.$$

The expected number of firms in region j is

$$E(n_j) = \sum_{s=1}^{S} E(n_{sj}) = \sum_{s=1}^{S} e^{\alpha_s + x'_{si}\beta}.$$

The corresponding own-region and cross-region elasticities for a region-industry specific shock

x_{sjk} are

$$\frac{\partial \log E(n_j)}{\partial x_{sjk}} = \frac{\partial \log E(n_{sj})}{\partial x_{sjk}} \cdot \frac{E(n_{sj})}{E(n_j)} = P_{s|j}\beta_k,$$

$$\frac{\partial \log E(n_i)}{\partial x_{sjk}} = \frac{\partial \log E(n_{si})}{\partial x_{sjk}} \cdot \frac{E(n_{si})}{E(n_i)} = 0,$$

where $P_{s|j} = E(n_{sj})/E(n_j)$.
The own-region and cross-region elasticities for a region-specific shock x_{jk} are

$$\frac{\partial \log E(n_j)}{\partial x_{jk}} = \sum_{s=1}^{S} \left[\frac{\partial \log E(n_{sj})}{\partial x_{jk}} \cdot \frac{E(n_{sj})}{E(n_j)}\right] = \beta_k,$$

$$\frac{\partial \log E(n_i)}{\partial x_{jk}} = \sum_{s=1}^{S} \left[\frac{\partial \log E(n_{si})}{\partial x_{jk}} \cdot \frac{E(n_{si})}{E(n_i)}\right] = 0.$$

The expected total number of firms in all regions and industries is

$$E(n) = \sum_{s=1}^{S}\sum_{j=1}^{J} E(n_{js}) = \sum_{s=1}^{S} E(n_s) = \sum_{s=1}^{S}\left[e^{\alpha_s}\sum_{i=1}^{J} e^{x'_{si}\beta}\right],$$

and the corresponding elasticities for a region-industry specific shock x_{sjk} and a region-specific shock x_{jk} are, respectively,

$$\frac{\partial \log E(n)}{\partial x_{sjk}} = \frac{\partial \log E(n_{sj})}{\partial x_{sjk}} \cdot \frac{E(n_{sj})}{E(n)} = P_{s|j}P_j\beta_k,$$

$$\frac{\partial \log E(n)}{\partial x_{jk}} = \frac{\partial \log E(n_j)}{\partial x_{jk}} \cdot \frac{E(n_j)}{E(n)} = P_j\beta_k,$$

where $P_j = E(n_j)/E(n)$.

Grouped nested logit

The nested logit model for *grouped data* is given by the probability that firm f of industry s chooses the outside option $j = 0$ or region $j > 0$:

$$P_{0|s} = \frac{e^{\delta_s}}{e^{\delta_s} + (\sum_{j=1}^{J} e^{x'_{sj}\gamma/\lambda})^{\lambda}} = \frac{e^{\delta_s}}{e^{\delta_s} + (\sum_{j=1}^{J} e^{x'_{sj}\beta})^{\lambda}},$$

$$P_{j|s} = \frac{e^{x'_{sj}\beta}(\sum_{i=1}^{J} e^{x'_{si}\beta})^{\lambda-1}}{e^{\delta_s} + (\sum_{i=1}^{J} e^{x'_{si}\beta})^{\lambda}} = P_{j>0|s}\cdot P_{j|j>0,s} = (1-P_{0|s})P_{j|j>0,s},$$

where δ_s is an industry-specific constant, $\beta = \gamma/\lambda$ and

$$P_{j|j>0,s} = \frac{e^{x'_{sj}\beta}}{\sum_{i=1}^{J} e^{x'_{si}\beta}}.$$

The concentrated log likelihood function is

$$\begin{aligned}\log L(\beta) \quad &= \quad \sum_{s=1}^{S}\left\{\sum_{j=1}^{J} N_{sj}x'_{sj}\beta - \sum_{j=1}^{J}[N_{sj}\log\sum_{i=1}^{J} e^{x'_{si}\beta}]\right.\\ &+ \quad N_{s0}\log(N_{s0}) + N_s\log(N_s) - (N_s+N_{s0})\log(N_s+N_{s0})\}\,.\end{aligned}$$

The expected number of firms in domestic region $j > 0$ and industry s is

$$E(n_{sj}) = (n_s + n_{s0})P_{j|s} = (n_s + n_{s0})(1-P_{0|s})P_{j|j>0,s},$$

and the corresponding own-region and cross-region elasticities within industry s are, respectively,

$$\frac{\partial \log E(n_{sj})}{\partial x_{sjk}} = [1 - P_{j|j>0,s}(1 - \lambda P_{0|s})]\beta_k,$$
$$\frac{\partial \log E(n_{si})}{\partial x_{sjk}} = -P_{j|j>0,s}(1 - \lambda P_{0|s})\beta_k.$$

The expected number of all domestic firms in industry s is

$$E(n_s) = \sum_{j=1}^{J} E(n_{sj}) = (n_s + n_{s0})(1 - P_{0|s}),$$

and the corresponding elasticity within industry s is

$$\frac{\partial \log E(n_s)}{\partial x_{sjk}} = \lambda P_{0|s} P_{j|j>0,s}\beta_k.$$

The expected number of firms in domestic region j is

$$E(n_j) = \sum_{s=1}^{S} E(n_{sj}) = \sum_{s=1}^{S} (n_s + n_{s0}) P_{j|s}.$$

The corresponding own-region and cross-region elasticities for a region-industry specific shock x_{sjk} are

$$\frac{\partial \log E(n_j)}{\partial x_{sjk}} = \frac{\partial \log E(n_{sj})}{\partial x_{sjk}} \cdot \frac{E(n_{sj})}{E(n_j)} = [1 - P_{j|j>0,s}(1 - \lambda P_{0|s})] P_{s|j}\beta_k,$$
$$\frac{\partial \log E(n_i)}{\partial x_{sjk}} = \frac{\partial \log E(n_{si})}{\partial x_{sjk}} \cdot \frac{E(n_{si})}{E(n_i)} = -P_{j|j>0,s}(1 - \lambda P_{0|s}) P_{s|i}\beta_k,$$

where $P_{s|j} = E(n_{sj})/E(n_j)$.
The own-region and cross-region elasticities for a region-specific shock x_{jk} are

$$\frac{\partial \log E(n_j)}{\partial x_{jk}} = \sum_{s=1}^{S} \left[\frac{\partial \log E(n_{sj})}{\partial x_{jk}} \cdot \frac{E(n_{sj})}{E(n_j)} \right] = \beta_k \sum_{s=1}^{S} [1 - P_{j|j>0,s}(1 - \lambda P_{0|s})] P_{s|j},$$

$$\frac{\partial \log E(n_i)}{\partial x_{jk}} = \sum_{s=1}^{S} \left[\frac{\partial \log E(n_{si})}{\partial x_{jk}} \cdot \frac{E(n_{si})}{E(n_i)} \right] = -\beta_k \sum_{s=1}^{S} P_{j|j>0,s}(1 - \lambda P_{0|s}) P_{s|i}.$$

The expected total number of firms in all domestic regions and industries is

$$E(n) = \sum_{s=1}^{S} \sum_{j=1}^{J} E(n_{js}) = \sum_{j=1}^{J} E(n_j),$$

and the corresponding elasticities for a region-industry specific shock x_{sjk} and a region-specific shock x_{jk} are, respectively,

$$\frac{\partial \log E(n)}{\partial x_{sjk}} = \left[\frac{\partial \log E(n_{sj})}{\partial x_{sjk}} + \sum_{i \neq j, i>0} \frac{\partial \log E(n_{si})}{\partial x_{sjk}} \right] \cdot \frac{E(n_{sj})}{E(n)} = \beta_k \lambda P_{0|s} P_{s|j} P_j,$$

$$\frac{\partial \log E(n)}{\partial x_{jk}} = \left[\frac{\partial \log E(n_j)}{\partial x_{jk}} + \sum_{i \neq j, i>0} \frac{\partial \log E(n_i)}{\partial x_{jk}} \right] \cdot \frac{E(n_j)}{E(n)} = \beta_k P_j \sum_{s=1}^{S} (\lambda P_{0|s} P_{s|j}),$$

where $P_j = E(n_j)/E(n)$.