
ELSEVIER
economics letters

Economics Letters 120 (2013) 288–291

Contents lists available at SciVerse ScienceDirect

Economics Letters

journal homepage: www.elsevier.com/locate/ecolet

On discrete location choice models[☆]

Nils Herger[a,*], Steve McCorriston[b]

[a] Study Center Gerzensee, Dorfstrasse 2, P.O. Box 21, 3115 Gerzensee, Switzerland

[b] Department of Economics, University of Exeter Business School, Streatham Court, Rennes Drive, Exeter EX4 4PU, England, UK

HIGHLIGHTS

- Poisson and conditional logit regressions are polar location choice models.
- A dissimilarity parameter λ covers the continuum between these models.
- The dissimilarity parameter is not identified in Schmidheiny and Brülhart (2011).
- We show that a choice consistent normalisation identifies λ.
- With panel data, a Poisson regression approach facilitates the estimation of λ.

ARTICLE INFO

Article history:
Received 24 January 2013
Received in revised form
5 April 2013
Accepted 8 April 2013
Available online 6 May 2013

JEL classification:
C2

Keywords:
Conditional logit model
Nested logit model
Poisson regression

ABSTRACT

When estimating location choices, Poisson regressions and conditional logit models yield identical coefficient estimates (Guimarães et al., 2003). These econometric models involve polar assumptions as regards the similarity of the different locations. Schmidheiny and Brülhart (2011) reconcile these polar cases by introducing a fixed outside option transforming the conditional logit into a nested logit framework. This gives rise to a dissimilarity parameter ($\lambda \in [0; 1]$) equalling 1 in Poisson regressions (with completely dissimilar locations) and 0 in conditional logit models (with completely similar locations). The dissimilarity parameter is not identified in Schmidheiny and Brülhart (2011). We show that a choice consistent normalisation identifies λ and that, with panel data, its estimation is facilitated by adopting a Poisson regression approach.

1. Introduction

This paper extends recent firm location choice models of Schmidheiny and Brülhart (2011) – henceforth SB – to identify a dissimilarity parameter (λ) between alternative locations by using panel Poisson regressions.

Let the firms undertaking a location choice be indexed with $i = 1, \dots, N$. Source countries are indexed with $s = 1, \dots, S$. The choice set includes host locations indexed with $h = 1, \dots, H$. A location choice denoted by $l_{i,sh}$ reveals that a host h with the profit opportunity $E[\Pi_{i,sh}]$ outperforms the other locations h' that could have been chosen instead, that is

$$l_{i,sh} = \begin{cases} 1 & E[\Pi_{i,sh}] > E[\Pi_{i,sh'}]\ \forall\, h \neq h' \\ 0 & \text{otherwise.} \end{cases} \quad (1)$$

A conditional logit model employs (1) as the dependent variable. Thereby, choice-specific variables x_{sh} (reported in logarithms) linearly affect profit expectations $E[\Pi_{i,sh}]$ via

$$E[\Pi_{i,sh}] = \delta_s + x'_{sh}\beta + \epsilon_{i,sh}, \quad (2)$$

where δ_s absorbs source-specific factors. Furthermore, β are coefficients to be estimated. The stochastic component $\epsilon_{i,sh}$ follows a Gumbel distribution with location and scale parameter normalised to, respectively, 0 and 1. The probability that a firm of s chooses h equals

$$P_{sh} = \frac{\exp\left(x'_{sh}\beta\right)}{\sum_{s=1}^{S}\sum_{h=1}^{H}\exp\left(x'_{sh}\beta\right)} = \frac{E[n^{cl}_{sh}]}{E[\bar{N}]}. \quad (3)$$

[☆] This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided that the original author and source are credited.

[*] Corresponding author. Tel.: +41 317803114.
E-mail addresses: nils.herger@szgerzensee.ch (N. Herger), s.mccorriston@exeter.ac.uk (S. McCorriston).

0165-1765/$ – see front matter © 2013 The Authors. Published by Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.econlet.2013.04.015

The log-likelihood function equals

$$\ln L^{cl}(\beta) = \sum_{s=1}^{S}\sum_{h=1}^{H} n_{sh}\ln(P_{sh}) = \sum_{s=1}^{S}\left\{\sum_{h=1}^{H} n_{sh}x'_{sh}\beta - \sum_{h=1}^{H}\left[n_{sh}\ln\sum_{h=1}^{H}\exp\left(x'_{sh}\beta\right)\right]\right\} \tag{4}$$

and permits us to estimate β. Guimarães et al. (2003) show that a count regression onto n_{sh} (the number of location choices) provides an alternative to estimate β. To see this, multiply (3) with the denominator, which yields the (panel) Poisson regression

$$E[\widetilde{n}^{pc}_{sh}] = \exp\left(\delta_s + x'_{sh}\beta\right) = \alpha_s E[n^{pc}_{sh}], \tag{5}$$

where $\alpha_s = \ln(\delta_s)$ and $E[n^{pc}_{sh}] = \exp(x'_{sh}\beta)$. Assuming that $\widetilde{n}^{pc}_{sh}$ is Poisson distributed with the conditional mean function $\exp(\delta_s + x'_{sh}\beta)$ of (5) yields a log-likelihood contribution of s given by

$$\ln L^{pc}_s(\alpha_s,\beta) = -\alpha_s\sum_{h=1}^{H}\exp(x'_{sh}\beta) + \ln\alpha_s\sum_{h=1}^{H} n_{sh} + \sum_{h=1}^{H} n_{sh}x'_{sh}\beta - \sum_{h=1}^{H}\ln n_{sh}! \tag{6}$$

Equating the first derivative with respect to α_s with 0, and solving for α_s yields the maximum likelihood estimator of

$$\alpha_s = \frac{\sum_{h=1}^{H} n_{sh}}{\sum_{h=1}^{H}\exp\left(x'_{sh}\beta\right)} = \frac{\overline{n}_s}{E[\overline{n}_s]}. \tag{7}$$

Hence, α_s absorbs the discrepancy between the observed number of location choices $\overline{n}_s$ and the number $E[\overline{n}_s]$ expected from a Poisson distribution. Thereby, $0 < \alpha_s < 1$ implies that the observed number of location choices is "underreported". Substituting (7) into (6) and summing over S yields the log-likelihood function of the fixed effects Poisson regression

$$\ln L^{pc}(\beta) = \sum_{s=1}^{S}\left\{\sum_{h=1}^{H} n_{sh}x'_{sh}\beta - \sum_{h=1}^{H}\left[n_{sh}\ln\sum_{h=1}^{H}\exp\left(x'_{sh}\beta\right)\right]\right\} + \text{constant}, \tag{8}$$

which looks like a multinomial logit model (Hausman et al., 1984, p. 919).[1] Specifically, since (8) differs from (4) only by a constant, the corresponding estimates for β are identical!

SB observe that the elasticity of the Poisson regression, given by

$$\eta^{pc}_k = \frac{\partial E[\widetilde{n}^{pc}_{sh}]}{\partial x_{sh,k}}\frac{x_{sh,k}}{E[\widetilde{n}^{pc}_{sh}]} = \beta_k, \tag{9}$$

differs from the conditional logit model, given by

$$\eta^{cl}_{sh,k} = \frac{\partial E[n^{cl}_{sh}]}{\partial x_{sh,k}}\frac{x_{sh,k}}{E[n_{sh}]} = (1-P_{sh})\beta_k, \tag{10}$$

whereby β_k denotes the coefficient pertaining to $x_{sh,k}$. This reflects that Poisson regressions deem the locations to be completely dissimilar. Hence, a change of $x_{sh,k}$ affects the number of location choices with h, but not with h'. SB refer to this as a "positive sum world". Conversely, the conditional logit model is a "zero sum world" where the locations represent completely similar options. Hence, when more firms choose h, this triggers an equivalent reduction elsewhere.

SB show that the introduction of an outside option transforms the conditional logit into a nested logit model covering the continuum between the zero and positive sum world. The outside option $h = 0$ is independent of x_{sh}. The corresponding profit equals

$$E[\Pi_{i,s0}] = \delta_s + \epsilon_{i,s0}. \tag{11}$$

Since the outside option contains only one alternative, this nested logit model, depicted in Fig. 1, involves the partial degeneracy discussed in Hunt (2000). The probability P_{sh} depends now on the probability $P_{øs}$ of *not* choosing the outside option and the (conditional) probability $P_{sh|ø}$ to locate in $h > 0$, that is

$$P_{sh} = P_{øs}\cdot P_{sh|ø} = \frac{\left[\sum_{s=1}^{S}\sum_{h=1}^{H}\exp\left(x'_{sh}\beta\varsigma^{ø}_s\right)\right]^{\frac{\lambda^ø_s}{\varsigma^ø_s}}}{\left[\exp\left(\delta_s\varsigma^o_s\right)\right]^{\lambda^o_s} + \left[\sum_{s=1}^{S}\sum_{h=1}^{H}\exp\left(x'_{sh}\beta\varsigma^ø_s\right)\right]^{\frac{\lambda^ø_s}{\varsigma^ø_s}}} \cdot \frac{\exp\left(x'_{sh}\beta\varsigma^ø_s\right)}{\sum_{s=1}^{S}\sum_{h=1}^{H}\exp\left(x'_{sh}\beta\varsigma^ø_s\right)} \tag{12}$$

$$= \frac{\exp\left(x'_{sh}\beta\varsigma^ø_s\right)\left[\sum_{s=1}^{S}\sum_{h=1}^{H}\exp\left(x'_{sh}\beta\varsigma^ø_s\right)\right]^{\left(\frac{\lambda^ø_s}{\varsigma^ø_s}-1\right)}}{\left[\exp\left(\delta_s\varsigma^o_s\right)\right]^{\lambda^o_s} + \left[\sum_{s=1}^{S}\sum_{h=1}^{H}\exp\left(x'_{sh}\beta\varsigma^ø_s\right)\right]^{\frac{\lambda^ø_s}{\varsigma^ø_s}}}. \tag{13}$$

The inclusive value parameter $(\lambda^ø_s/\varsigma^ø_s) \in [0, 1]$ measures the dissimilarity between the locations $h > 0$. Specifically,

$$(\lambda^ø_s/\varsigma^ø_s) = \sqrt{1-\rho^ø_s} \tag{14}$$

where $\rho^ø_s \in [0, 1]$ is the correlation between the stochastic profit components $\epsilon_{i,sh|ø}$ of investing in different locations. Consider the outside option o offering only the basic "choice" of $h = 0$. Hunt (2000) observes that the distinction between unconditional and conditional probabilities is here obsolete, as $P_{s0|o} = 1$ and $P_{s0} = P_{os}\times P_{s0|o}$. The probability of choosing $h = 0$ equals

$$P_{os} = P_{s0} = (1-P_{øs}) = \frac{\exp\left(\delta_s\varsigma^o_s\right)^{\lambda^o_s}}{\left[\exp\left(\delta_s\varsigma^o_s\right)\right]^{\lambda^o_s} + \left[\sum_{s=1}^{S}\sum_{h=1}^{H}\exp\left(x'_{sh}\beta\varsigma^ø_s\right)\right]^{\frac{\lambda^ø_s}{\varsigma^ø_s}}}. \tag{15}$$

The coefficients β can be estimated by maximum likelihood from (13) and (15). However, empirically, only the correlation $\rho^ø_s$, but not the scale parameters $\lambda^ø_s$ and $\varsigma^ø_s$, can be estimated from the data (Hunt, 2000). This over-identification problem necessitates some normalisation. SB (p. 217) set $\varsigma^ø_s = 1$, $\varsigma^o_s = 1$, and $\lambda^o_s = 1$ wherefore (13) and (15) become

$$P_{sh} = \frac{\exp\left(x'_{sh}\beta\right)\left[\sum_{s=1}^{S}\sum_{h=1}^{H}\exp\left(x'_{sh}\beta\right)\right]^{(\lambda^ø_s-1)}}{\exp\left(\delta_s\right) + \left[\sum_{s=1}^{S}\sum_{h=1}^{H}\exp\left(x'_{sh}\beta\right)\right]^{\lambda^ø_s}} = \frac{\exp(x'_{sh}\beta)(E[N^ø])^{(\lambda^ø_s-1)}}{\exp(\delta_s) + (E[N^ø])^{\lambda^ø_s}} \tag{16}$$

[1] For a textbook discussion of the fixed effects Poisson regression, see Cameron and Trivedi (1998, ch. 9.3).

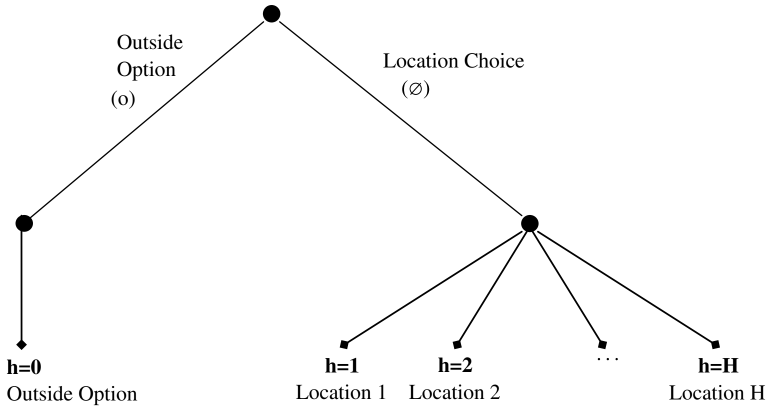

Fig. 1. Location choice with outside option.

$$P_{0s} = \frac{\exp(\delta_s)}{\exp(\delta_s) + \left[\sum_{s=1}^{S}\sum_{h=1}^{H}\exp(x'_{sh}\beta)\right]^{\lambda_s^{\o}}} = \frac{\exp(\delta_s)}{\exp(\delta_s) + (E[N^{\o}])^{\lambda_s^{\o}}}. \quad (17)$$

The nested logit log-likelihood function coincides with (4). Again, the estimates for β are identical and the normalisation of the scale parameters does not affect β.

The nested logit elasticity equals

$$\eta_{sh,k}^{nl} = \frac{\partial E[n_{sh}^{nl}]}{\partial x_{sh,k}} = [1 - P_{sh|\o}(1 - \lambda_s^{\o} P_{s0})]\beta_k, \quad (18)$$

and maps mathematically into the conditional logit model when $\lambda_s^{\o} = 0$ and into the Poisson regression when $\lambda_s^{\o} = 1$ and $P_{s0} = 1$. Hence, though the conditional logit model and the Poisson regression may have different econometric underpinnings, they share a functional form that connects them with the same, suitably parameterised nested logit model.

2. Choice consistent normalisation

The SB approach has two shortcomings. Firstly, the normalisation of λ and ς leads to different nested logit models with different elasticities (Hensher and Greene, 2002). Koppelman and Wen (1998) suggest that any normalisation should at least be invariant to adding a constant Δ to all profits of (2) and (11) since this leaves the ranking of the final options unchanged. Appendix A shows that (16) and (17) do not fulfil this property. Secondly, δ_s and $\lambda_s^{\o}$ cannot be separately identified (SB, 2011, p. 217) since they appear in the same first order condition

$$\exp(\delta_s) = \frac{n_{os}}{n_{\o s}}\left[\sum_{h=1}^{H}\exp(x'_{sh}\beta)\right]^{\lambda_s^{\o}}, \quad (19)$$

where n_{os} and $n_{\o s}$ denote, respectively, the number of times the outside option $h = 0$ or location $h > 0$ has been chosen.

Adopting the following choice-consistent normalisation avoids these caveats.

Proposition 1. *Setting $\lambda_s^{o} = \lambda_s^{\o} = \lambda_s$, $\varsigma_s^{\o} = 1$, and $\varsigma_s^{o} = 0$ represents a choice-consistent normalisation in the sense that adding a constant Δ to the profits (2) and (11) does not affect the choice outcome.*

Proof. Appendix A. □

With Proposition 1, (13) and (15) become

$$P_{sh} = \frac{\exp(x'_{sh}\beta)\left[\sum_{s=1}^{S}\sum_{h=1}^{H}\exp(x'_{sh}\beta)\right]^{(\lambda_s-1)}}{1 + \left[\sum_{s=1}^{S}\sum_{h=1}^{H}\exp(x'_{sh}\beta)\right]^{\lambda_s}} = \frac{(E[N^{\o}])^{(\lambda_s-1)}}{1 + (E[N^{\o}])^{\lambda_s}}\exp(x'_{sh}\beta) \quad (20)$$

$$P_{s0} = \frac{1}{1 + \left[\sum_{s=1}^{S}\sum_{h=1}^{H}\exp(x'_{sh}\beta)\right]^{\lambda_s}} = \frac{1}{1 + (E[N^{\o}])^{\lambda_s}}, \quad (21)$$

Proposition 1 implies that $\exp(\delta_s\varsigma_s^{o})^{\lambda_s^{o}} = \exp(0)^{\lambda_s} = 1$ and, hence, normalises to 1 the contribution of the outside option. This is unproblematic since the parameter λ_s pertaining to $E[N^{\o}]^{\lambda_s}$ already weights the importance between the outside option $h = 0$ and the locations $h > 0$.

When $\varsigma_s^{\o} = 1$, (14) implies that $\lambda_s = \sqrt{1 - \rho_s}$, and the correlation ρ_s reflects the degree of dissimilarity between the locations.

3. Poisson regression with λ_s

Estimating λ_s from (19) involves n_{os}. However, how many times the outside option has been chosen is often unobservable. Counting location choices provides a possible remedy for this.

Similar to Section 1, transforming the nested logit model into a Poisson regression requires the multiplication of P_{sh} of (20) with the denominator $E[N] = 1 + (E[N^{\o}])^{\lambda_s}$ to obtain

$$E[\widetilde{n}_{sh}^{pcu}] = P_{sh}E[N] = \underbrace{(E[N^{\o}])^{(\lambda_s-1)}}_{=\alpha_s}\exp(x'_{sh}\beta) = \alpha_s E[n_{sh}^{pc}]. \quad (22)$$

This resembles (5) with group effects parameterised by $\alpha_s = E[N^{\o}]^{(\lambda_s-1)}$. When $\lambda_s = 1$, we have $E[N^{\o}]^{(\lambda_s-1)} = 1$ and the basic Poisson regression with completely dissimilar elemental options arises. The outside option is irrelevant implying that $E[\widetilde{n}_{sh}^{pcu}]$ equals $E[n_{sh}^{pc}]$. Conversely, when $\lambda_s < 1$, the outside option $h = 0$ is similar to the option of investing in locations $h > 0$. The greater this similarity, the more the number $E[\widetilde{n}_{sh}^{pcu}]$ of observed location choices differs from $E[n_{sh}^{pc}]$ of a basic count process.

Solving $\alpha_s = E[N^{\o}]^{(\lambda_s-1)}$ for the λ_s yields

$$\lambda_s = \frac{\ln(E[N^{\o}]\alpha_s)}{\ln(E[N^{\o}])}. \quad (23)$$

Recall that α_s can be estimated from (7) and reflects the degree of underreporting. The Poisson regression with $\lambda_s = 1$ requires that $\alpha_s = 1$. This means that there is no underreporting, which is maybe intuitive since the outside option does not affect the choice of a location that differs completely from the alternatives. The conditional logit model with $\lambda_s = 0$ requires $\alpha_s = 1/E[N^{ø}]$. This implies a degree of underreporting increasing with the number $E[N^{ø}]$ of location choices. Typically, α_s is close to 0, even when the sample contains a relatively modest number of location choices $N^{ø}$.

The identification of (23) requires that $E[N^{ø}] > 0$ (we do actually have location choices) and that we have panel data permitting us to calculate the group effects (7).

Appendix B shows that the elasticity of $E[\widetilde{n}^{pcu}_{sh}]$ with respect to $x_{sh,k}$ equals

$$\eta^{pcu}_{sh,k} = [1 - P_{sh|ø}(1 - \lambda_s)]\beta_k. \quad (24)$$

This is again consistent with SB. Specifically, $\lambda_s = 1$ returns the Poisson elasticity of (9) and $\lambda_s = 0$ the conditional logit elasticity of (10). Also, evaluating (24) yields $\eta^{pcu}_{sh,k} = \lambda_s\beta_k + (1-\lambda_s)(1-P_{sh|ø})\beta_k$, meaning that the elasticity of the Poisson regression with a fixed outside option is a linear average between (9) and (10).

4. Conclusion

This paper extends SB by a choice consistent framework, which permits us to identify the dissimilarity parameter λ. Furthermore, using a Poisson regression facilitates the estimation of the dissimilarity parameter.

Acknowledgement

Insightful comments provided by the referee helped us to clarify several aspects of this paper.

Appendix A. Proof Proposition 1

The choice consistency (as defined in the text) of the normalisation of Proposition 1 ($\lambda^o_s = \lambda^{ø}_s = \lambda_s$; $\varsigma^{ø}_s = 1$; $\varsigma^o_s = 0$) reflects the well known result that multinomial logit distributions are homogeneous of degree zero. This implies that adding a constant Δ to all arguments entering $P_{sh} = P_{øs}P_{sh|ø}$ of (20) leaves the probabilities P_{sh} – and hence the choice outcome – unchanged. The same holds for P_{s0} of (21).[2]

The choice consistency does not hold for the SB normalisation ($\varsigma^{ø}_s = 1$, $\varsigma^o_s = 1$; $\lambda^o_s = 1$). For example, adding Δ to P_{s0} of (17) yields

$$P^*_{s0} = \frac{\exp(\delta_s + \Delta)}{[\exp(\delta_s + \Delta)] + \left[\sum_{s=1}^{S}\sum_{h=1}^{H}\exp\left(x'_{sh}\beta + \Delta\right)\right]^{\lambda^{ø}_s}}$$

and

$$P^*_{s0} = \frac{\exp(\delta_s)\exp(\Delta)}{\exp(\delta_s)\exp(\Delta) + \left\{\left[\sum_{s=1}^{S}\sum_{h=1}^{H}\exp\left(x'_{sh}\beta\right)\right]\exp(\Delta)\right\}^{\lambda^{ø}_s}}$$
$$= \frac{\exp(\delta_s)\exp(\Delta)}{\exp(\delta_s)\exp(\Delta) + \exp(\Delta)^{\lambda^{ø}_s}\left[\left(\sum_{s=1}^{S}\sum_{h=1}^{H}\exp(x'_{sh}\beta)\right)\right]^{\lambda^{ø}_s}}$$
$$\neq P_{s0}.$$

Appendix B. Defining the elasticity $\eta^{pcu}_{sh,k}$

The elasticity $\eta^{pcu}_{sh,k}$ is defined as

$$\eta^{pcu}_{sh} = \frac{\partial E[\widetilde{n}^{pcu}_{sh}]}{\partial x_{sh,k}}\frac{x_{sh,k}}{E[\widetilde{n}^{pcu}_{sh}]}.$$

With $E[\widetilde{n}^{pcu}_{sh}] = (E[N^{ø}])^{(\lambda_s-1)}\exp(x'_{sh}\beta)$ of (22) and $E[N^{ø}] = \sum_{s=1}^{S}\sum_{h=1}^{H}\exp(x'_{sh}\beta)$, we have

$$\eta^{pcu}_{sh} = \left[(\lambda_s - 1)(E[N^{ø}])^{(\lambda_s-2)}\exp(x_{sh,k})\frac{\beta_k}{x_{sh,k}}\exp(x_{sh,k}) + E[N^{ø}]^{(\lambda_s-1)}\exp(x_{sh,k})\frac{\beta_k}{x_{sh,k}}\right]\frac{x_{sh,k}}{E[N^{ø}]^{(\lambda_s-1)}\exp(x_{sh,k})}.$$

Cancelling terms yields

$$\eta^{pcu}_{sh} = (\lambda_s - 1)(E[N^{ø}])^{-1}\exp(x'_{sh}\beta)\beta_k + \beta_k.$$

Since $(E[N^{ø}])^{-1}\exp(x'_{sh}\beta)$ equals $P_{sh|ø}$ according to (12), we have

$$\eta^{pcu}_{sh} = [1 + (\lambda_s - 1)P_{sh|ø}]\beta_k = [1 - (1 - \lambda_s)P_{sh|ø}]\beta_k.$$

References

Cameron, C.A., Trivedi, P.K., 1998. Regression Analysis of Count Data. Cambridge University Press, Cambridge.
Guimarães, P., Figueirdo, O., Woodward, D., 2003. A tractable approach to the firm location decision problem. The Review of Economics and Statistics 85, 201–204.
Hausman, J., Hall, B.H., Griliches, Z., 1984. Econometric models for count data with an application to the patents-R&D relationship. Econometrica 52, 909–938.
Hensher, D.A, Greene, W.H., 2002. Specification and estimation of the nested logit model: alternative normalisations. Transportation Research Part B 36, 1–17.
Hunt, G.L., 2000. Alternative nested logit model structures and the special case of partial degeneracy. Journal of Regional Science 40, 89–113.
Koppelman, F.S., Wen, C.H., 1998. Alternative nested logit models: structure properties and estimation. Transportation Research 32, 289–298.
Schmidheiny, K., Brülhart, M., 2011. On the equivalence of location choice models: conditional logit, nested logit and Poisson. Journal of Urban Economics 69, 214–222.

[2] A formal derivation of these results can be made available on request.