On Nesting Location Choice Models Correctly: A Reply to Herger and McCorriston (*Economics Letters*, 2013)*

Kurt Schmidheiny‡
University of Basel

Marius Brülhart§
University of Lausanne

March 3, 2015

Abstract

As shown by Guimaraes, Figueiredo and Woodward (2003), a particular class of conditional logit models yield identical parameter estimates to a Poisson count data model. In Schmidheiny and Brülhart (2011), we have pointed out that the conditional logit model and the Poisson model can be seen as polar cases of a continuum of intermediate cases which emerge from a random utility nested logit model. Herger and McCorriston (2013) have proposed an alternative nested logit model. In this note, we show that their paper misrepresents our model, that their alternative specification does not in fact nest the Poisson model, and that their analysis contains a number of formal errors.

Highlights:

- The nested logit model in Schmidheiny and Brülhart (2011) is a random utility model that nests the conditional logit and Poisson models.
- The nested logit model in Herger and McCorriston (2013) is a restricted version of Schmidheiny and Brülhart (2011) that does not nest the Poisson model.
- The proposed estimation strategy in Herger and McCorriston (2013) is incorrect.

JEL Classification: C25, R3

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

*We thank Jörg Kalbfuss for carefully reading the manuscript. We gratefully acknowledge financial support from the Swiss National Science Foundation (Sinergia Grants 130648 and 147668; and NCCR "Trade Regulation").

‡Wirtschaftswissenschaftliche Fakultät, Universität Basel, Peter Merian-Weg 6 4125 Basel, Switzerland; e-mail: kurt.schmidheiny@unibas.ch. Also affiliated with the the Centre for Economic Policy Research (CEPR) and CESifo.

§Department of Economics, Faculty of Business and Economics, University of Lausanne, CH - 1015 Lausanne, Switzerland. E-mail: Marius.Brulhart@unil.ch. Web page: www.hec.unil.ch/mbrulhar. Also affiliated with the Centre for Economic Policy Research (CEPR).

1 Introduction

Location choice models explain the spatial distribution of agents such as firms or households as the outcome of individual location decisions. The most popular representation of such choices is McFadden's (1974) conditional logit model. Guimaraes, Figueiredo and Woodward (2003) have shown that a particular class of conditional choice models yield identical parameter estimates to a Poisson count data model. Schmidheiny and Brülhart (2011, henceforth SB) show that, while the two models are indeed observationally equivalent in cross-section data, they are distinctively different descriptions of the world and imply different marginal effects and elasticities. SB also show that the elasticities of the conditional logit model and the Poisson model can be seen as polar cases of a continuum of intermediate cases which emerge from a random utility nested logit model (RUNL). Herger and McCorriston (2013, henceforth HM) propose an alternative nested logit model. They claim that their model, unlike the SB RUNL, allows the "dissimilarity parameter" to be identified in cross-section data, thus making it possible to estimate how closely the data correspond to the conditional logit or Poisson models. In this note, we show that HM wrongly represent SB's RUNL and that their proposed alternative specification is a restricted version of SB which does not nest the Poisson model.

2 Conditional Logit and Poisson

Using the notation of HM, the basic framework in Guimaraes, Figueiredo and Woodward (2003) is the following

- Firms $i = 1, ..., N$ choose among $h = 1, ..., H$ locations. h are e.g. states within a country.
- There are $s = 1, ..., S$ different types of firms. s are e.g. industrial sectors or country of origin for foreign firms.
- The number of firms n_{sh} of each type s is observed in each location h. n_{sh} is used to denote both the observed number and the *random variable* in the data generating process.
- A vector of observable explanatory variables is denoted x_{sh}, characterizing locations h and potentially also differing across types s.

The location choice of an individual firm i is assumed to correspond to a conditional logit model. Thus, the profit function

$$\pi_{ish} = x'_{sh}\beta + \varepsilon_{ish} \tag{1}$$

consists of a type-specific deterministic part $x'_{sh}\beta$ and a stochastic part ε_{ish}, which has zero mean and is known to the firm but unknown to the econometrician. Firms choose the location which yields the highest profit.

Comment 1 *The expected profit function $E(\pi_{ish})$ in the conditional logit model does* not *contain an error term as stated in HM eq. (1) and (2). It is* not *useful to include a constant (δ_s in HM) or type-specific fixed effect in (1) as stated in HM eq. (2), because this type-specific constant is not identified in the conditional logit model.*

The conditional logit model contains the assumption that the stochastic term ε_{ish} is independent across i and h and follows an extreme value type 1 distribution. The conditional logit model implies the following choice probabilities:

$$P_{sh}^{CL} = \frac{e^{x'_{sh}\beta}}{\sum_{g=1}^{H} e^{x'_{sg}\beta}}. \tag{2}$$

Comment 2 *The choice probabilities in (2) are* conditional *on the type s. The denominator in eq. 2 does* not *sum over the firm types s as stated in HM eq. (3). The double summation in HM eq. (3) implies that the firm's type s is a* choice *of the firm rather than a* characteristic*, which is not what the basic framework assumes and which contradicts HM's own decision rule in HM eq. (1).*

The conditional logit model only predicts probabilities (relative numbers) and not absolute numbers of location choices, as it does not explicitly model the total number n_s of firms of each type. The total number n_s of firms of type s is often implicitly assumed as given in the conditional logit model. With exogenous n_s, the expected absolute number of choices is

$$E(n_{sh}^{CL}) = n_s P_{sh}. \tag{3}$$

SB derive the effect of a change in the k-th characteristic x_{shk} on the expected number of firms of type s in location h as the following semi-elasticity:

$$\eta_{sh}^{CL} = \frac{\partial log E(n_{sh}^{CL})}{\partial x_{shk}} = \frac{\partial E(n_{sh}^{CL})}{\partial x_{shk}} \frac{1}{E(n_{sh}^{CL})} = (1 - P_{sh})\beta_k, \tag{4}$$

assuming an exogenous number n_s of firms of each type.

Comment 3 $(1-P_{sh})\beta_k$ *equals* $[\partial E(n_{sh}^{CL})/\partial x_{shk}]\cdot[1/E(n_{sh}^{CL})]$ *and not* $[\partial E(n_{sh}^{CL})/\partial x_{shk}]\cdot[x_{shk}/E(n_{sh}^{CL})]$ *as stated in HM eq. (10). This expression is a semi-elasticity if x_{shk} is defined in levels and an elasticity if x_{shk} is defined in logs.*[1]

Guimaraes, Figueiredo and Woodward (2003) introduce a Poisson count data model which yields identical parameter estimates for the slope coefficients β. The Poisson approach directly models the expected absolute number of choices:

$$E(n_{sh}^{PC}) = e^{\alpha_s + x'_{sh}\beta}. \tag{5}$$

The Poisson model is a distinctively different data generating process from the conditional logit model and hence a different description of the world.

Comment 4 *The type-specific constants α_s in the Poisson model are unrelated to fixed effects in the profit function (1) of the conditional logit model, contrary to the implied assertion by HM eq. (2) and (5).*[2]

In the Poisson model, the effect of a change in the k-th local characteristic x_{shk} on the expected number of firms of type s in location h as semi-elasticity is

$$\eta_{sh}^{PC} = \frac{\partial log E(n_{sh}^{PC})}{\partial x_{shk}} = \frac{\partial E(n_{sh}^{PC})}{\partial x_{shk}} \frac{1}{E(n_{sh}^{CL})} = \beta_k. \tag{6}$$

Comment 5 β_k *equals* $[\partial E(n_{sh}^{PC})/\partial x_{shk}]\cdot[1/E(n_{sh}^{PC})]$ *and not* $[\partial E(n_{sh}^{PC})/\partial x_{shk}]\cdot[x_{shk}/E(n_{sh}^{PC})]$ *as stated in HM eq. (9). This expression is a semi-elasticity if x_{shk} is defined in levels and an elasticity if x_{shk} is defined in logs.*

[1] HM define x_{shk} as reported in logs. We see no value added in restricting x_{shk} to strictly positive values.

[2] We find the derivation of the Poisson model in HM to be formally incorrect and its interpretation to be misleading. While the definition of the Poisson model in HM eq. (5) is accurate, it is *not* the result of multiplying the CL probability in HM eq. (3) by its denominator. The constant term $\alpha_s = exp(\delta_s)$ in HM eq. (5) can therefore not be related to the non-identified constant δ_s in HM eg. (2). We also find the interpretation in the first line after HM eq. (7) misleading, because $E(\bar{n}_s^{PC}) = \sum_{h=1}^{H} e^{x'_{sh}\beta}$ is *not* the expected number of location choices in the Poisson model (which instead is $E(n_s^{PC}) = E(\sum_{h=1}^{H} n_{sh}^{PC}) = \sum_{h=1}^{H} e^{\alpha_s + x'_{sh}\beta}$). Moreover, while the ML estimator of α_s in HM eq. (7) is correct, it *cannot* be interpreted as the "discrepancy" between expected and observed numbers of observations.

3 The RUNL

In SB, we have proposed the following random utility nested logit (RUNL) model.[3] The profit of a firm i of type s in location $h = 1, ..., H$ is given by

$$\pi_{ish} = x'_{sh}\gamma_s + \varepsilon_{ish} = x'_{sh}\beta\lambda_s + \varepsilon_{ish}, \tag{7}$$

where the reparametrization $\gamma_s = \beta\lambda_s$ is described in SB, p. 216.[4] This reparametrization, in combination with the assumptions that β is constant across types s and γ_s is type-specific, formally allows us to relate the coefficient vector β to both the CL and the Poisson model. The RUNL model of SB moreover considers an outside option $h = 0$, which can be interpreted as locating in another country or not starting up a company. The outside option yields profit

$$\pi_{iso} = \delta_s + \varepsilon_{iso}. \tag{8}$$

It is assumed in the RUNL model that the stochastic term ε_{ish} follows a special type of generalized extreme value (GEV) distribution (see McFadden, 1978) such that all locations $h = 1, ..., H$ belong to one *nest* and the outside option $h = 0$ belongs to another (degenerate) nest. Following Amemiya (1985, p. 303, eq. 9.3.60) the choice probabilities are then expressed as:[5]

$$P_{sh}^{RUNL} = (1 - P_{so}^{RUNL}) \cdot P_{sh|\varnothing}^{RUNL} \quad \text{for all } h > 0, \tag{9}$$

$$P_{so}^{RUNL} = \frac{e^{\delta_s}}{e^{\delta_s} + \left(\sum_{g=1}^{H} e^{x'_{sg}\gamma_s/\lambda_s}\right)^{\lambda_s}} = \frac{e^{\delta_s}}{e^{\delta_s} + \left(\sum_{g=1}^{H} e^{x'_{sg}\beta}\right)^{\lambda_s}}, \text{and} \tag{10}$$

$$P_{sh|\varnothing}^{RUNL} = \frac{e^{x'_{sh}\gamma_s/\lambda_s}}{\sum_{g=1}^{H} e^{x'_{sg}\gamma_s/\lambda_s}} = \frac{e^{x'_{sh}\beta}}{\sum_{g=1}^{H} e^{x'_{sg}\beta}}, \tag{11}$$

where $P_{sh|\varnothing}^{RUNL}$ is the probability of choosing location $h > 0$ conditional on not choosing the alternative. The parameter λ_s can be interpreted as $\lambda_s = \sqrt{1 - \rho_s}$, where ρ_s is the correlation of the stochastic term ε_{ish} within the choices $h > 0$. Note that this version of the RUNL is fully general; other versions of the nested logit model which are consistent with a random utility model are just reparametrizations of this specification.

Comment 6 *Adding a constant Δ to the profit functions of all choice options $h = 0, 1, ..., H$ does not alter the choice probabilities in the SB RUNL. HM p. 290 wrongly state the opposite, implying that the SB nested logit is not a RUNL.*

The nested logit (NL) model in Hunt (2000) encompasses specifications which are consistent with a random utility model and others which are not:

$$P_{sh}^{NL} = (1 - P_{so}^{NL}) \cdot P_{sh|\varnothing}^{NL} \quad \text{for all } h > 0, \tag{12}$$

$$P_{so}^{NL} = \frac{\left(e^{\delta_s\varsigma_s^o}\right)^{\lambda_s^o}}{\left(e^{\delta_s\varsigma_s^o}\right)^{\lambda_s^o} + \left(\sum_{g=1}^{H} e^{x'_{sg}\beta_s\varsigma_s^{\varnothing}}\right)^{\lambda_s^{\varnothing}/\varsigma_s^{\varnothing}}}, \text{and} \tag{13}$$

[3]The term random utility nested logit model is taken from Heiss (2002).

[4] In SB Section 2.2, we assume that $\lambda_s = \lambda$ is not type-specific. The generalization used here is trivial and has no empirical content, as λ_s is not identified independently of type-specific δ_s. All results in SB hold unchanged and no additional results can be derived with type-specific λ_s.

[5]See also McFadden (1977, p. 6, eq. 20), Anderson, de Palma, Thisse (1992, p. 46, eq. 2.39), Koppelman and Wen (1998, p. 290, eq. 4), Heiss (2002, p. 234, eq. 10), or Train (2009, p. 80, eq. 4.2).

$$P^{NL}_{sh|\o} = \frac{e^{x'_{sh}\beta_s \varsigma^{\o}_s}}{\sum_{g=1}^{H} e^{x'_{sg}\beta_s \varsigma^{\o}_s}}. \tag{14}$$

The RUNL in SB is a special case of the Hunt (2000) NL model with $\beta_s = \beta$, $\lambda^{\o}_s = \lambda_s$, $\lambda^{o}_s = 1$, $\varsigma^{\o}_s = 1$ and $\varsigma^{o}_s = 1$.

In the SB RUNL, the expected number of firms of type s in location h is

$$E(n^{RUNL}_{sh}) = (n_{so+s\o})P^{RUNL}_{sh}, \tag{15}$$

where $n_{so+s\o}$ is the exogenous total number of firms of type s which choose either $h > 0$ or $h = 0$. The total, $n_{so+s\o}$, is typically unobservable. The semi-elasticity of a change in the k-th local characteristic x_{shk} on the expected number of firms of type s in location $h > 0$ is

$$\eta^{RUNL}_{sh} = \frac{\partial log E(n^{RUNL}_{sh})}{\partial x_{shk}} = [1 - P^{RUNL}_{sh|\o}(1 - \lambda_s P^{RUNL}_{so})]\beta_k. \tag{16}$$

The RUNL elasticity η^{RUNL}_{sh} equals the CL elasticity if $\lambda_s = 0$ and the Poisson elasticity if $\lambda_s = 1$ and $\delta_s \to \infty$, hence $P^{RUNL}_{so} = 1$. The SB RUNL therefore nests the conditional logit model, the Poisson model and the continuum of cases in-between.

4 The Herger and McCorriston RUNL

The NL model proposed in HM is a special case of the Hunt (2000) NL model with $\beta_s = \beta$, $\lambda^{\o}_s = \lambda_s$, $\lambda^{o}_s = \lambda_s$, $\varsigma^{\o}_s = 1$, $\varsigma^{o}_s = 0$ and/or $\delta^{o}_s = 0$, resulting in (see HM eq. 20 and 21)

$$P^{HM}_{sh} = (1 - P_{so}) \cdot P_{sh|\o} = \frac{e^{x'_{sh}\beta}\left(\sum_{g=1}^{H} e^{x'_{sg}\beta}\right)^{\lambda_s - 1}}{1 + \left(\sum_{g=1}^{H} e^{x'_{sg}\beta}\right)^{\lambda_s}} \quad \text{for all } h > 0, \tag{17}$$

$$P^{HM}_{so} = \frac{1^{\lambda_s}}{1^{\lambda_s} + \left(\sum_{g=1}^{H} e^{x'_{sg}\beta}\right)^{\lambda_s}} = \frac{1}{1 + \left(\sum_{g=1}^{H} e^{x'_{sg}\beta}\right)^{\lambda_s}}, \text{and} \tag{18}$$

$$P^{HM}_{sh|\o} = \frac{e^{x'_{sh}\beta}}{\sum_{g=1}^{H} e^{x'_{sg}\beta}}.[6] \tag{19}$$

Comparing (10) and (18) shows that the HM specification is a restricted version of the SB RUNL with the implicit restriction $\delta_s = 0$ for all s.

Comment 7 *The HM RUNL is a restricted version of the fully general SB RUNL. There is no parameter to control for the value of the outside option in general, and the value of the outside option is assumed to be identical across types s.*

The semi-elasticity in the HM RUNL is identical to the one in the SB RUNL in eq. (16) but with $P^{HM}_{sh|\o}$ and $P^{HM}_{sh|0}$ in place of $P^{RUNL}_{sh|\o}$ and $P^{RUNL}_{sh|0}$, respectively:

$$\eta^{HM}_{sh} = \frac{\partial log E(n^{HM}_{sh})}{\partial x_{shk}} = [1 - P^{HM}_{sh|\o}(1 - \lambda_s P^{HM}_{so})]\beta_k. \tag{20}$$

This semi-elasticity cannot be equal to the one of the Poisson model (eq. 6), because $P^{HM}_{so} < 1$ (strictly) in eq. (18) and hence $1 - \lambda_s P^{HM}_{so} \neq 0$.

[6]Setting $\varsigma^{o}_s = 0$ has the same effect as setting $\delta^{o}_s = 0$ in eq. (13).

Comment 8 *The HM RUNL model does not nest the Poisson model. It cannot in fact be used to study the main question at hand: the equivalence of the CL and Poisson models. The elasticity in HM eq. (24) and Appendix B is incorrect.*[7]

5 Estimation

Guimaraes, Figueiredo and Woodward (2003) show that the conditional logit model (2) and the Poisson model (5) have identical concentrated likelihood functions (up to a additive constant) for the parameter vector β when the estimation is based on one cross-section of observations of n_{sh}, $h = 1, ..., H$. SB show in Appendix A.3 that this is also true for the coefficient vector β in the RUNL model, assuming $\lambda_s = \lambda$ (see also footnote 4). It is straightforward to show that this result also holds for type-specific λ_s.[8] Based on data for $h > 0$, the three models are therefore observationally equivalent while the implied marginal effects and elasticities differ. Note that the nested logit parameters δ_s and λ_s in the SB RUNL are not identified without observing n_{so}. If n_{so} were observed, either λ_s or δ_s would be identified but not independently. In practice, n_{so} is typically unknown. For example, the number of potential firms that could start up but do not materialize cannot be observed. So, in practice, neither λ_s nor δ_s are identified.

Unlike in the SB RUNL, λ_s is identified in the HM RUNL for the (unlikely) case that n_{so} were observed. So, as usual, parameter restrictions (here $\delta_s = 0$) help identify structural parameters. But, as shown in Comment 8, the restriction $\delta_s = 0$ is not a useful one in this framework. The ML estimator of λ_s with the HM restriction $\delta_s = 0$ is

$$\widehat{\lambda}_s = \frac{\log\left(n_{s\o}/n_{so}\right)}{\log[\sum_{h=1}^{H} exp(x'_{sh}\widehat{\beta})]}, \tag{21}$$

where $n_{s\o} = \sum_{h=1}^{H} n_{sh}$ and $\widehat{\beta}$ is the ML estimator of β (see the derivation in the Appendix). HM propose a different estimator for λ_s in HM eq. (23). Substituting HM eq. (7) and eq. (21) into HM eq. (23) yields:[9]

$$\widehat{\lambda}_s = \frac{\log\left(n_{s\o}\right)}{\log[\sum_{h=1}^{H} exp(x'_{sh}\widehat{\beta})]}, \tag{22}$$

Comment 9 *The estimator of λ_s in HM eq. (23) implicitly assumes that the number of firms choosing the outside option equals one, $n_{so} = 1$. We cannot see a justification for this assumption in a typical application.*

Brülhart and Schmidheiny (2015) show how a panel of observations of n_{sht} over multiple time periods t allows the researcher to identify a single parameter ρ_s which is a function of both δ_s and λ_s without the need to observe n_{sot}.

[7] The correct elasticity in the HM model is $\partial log E(n_{sh}^{HM})/\partial log\, x_{shk} = [1 - P_{sh|\o}^{HM}(1 - \lambda P_{so}^{HM})]\beta_k \cdot x_{shk}$, which contains the term P_{so}^{HM}. The definitions $E(N^{\o}) \equiv \sum_{g=1}^{H} e^{x'_{sg}\beta}$ (HM, eq. 20), and $E(N) \equiv 1 + E(N^{\o})_s^{\lambda}$ (HM, eq. 22) ares not meaningful, and we see no formal basis for HM eq. (23).

[8] Sketch of proof: Solve the first order condition $\partial \log L(\beta, \delta, \lambda)/\partial \delta_s$ for δ_s. Plug the resulting δ_s into $\log L(\beta, \delta, \lambda)$ for all s. λ_s cancels out in the resulting concentrated likelihood function $\log L(\beta)$.

[9] HM equation (21) implicitly defines $E(n_{\o}) = \sum_{s=1}^{S}\sum_{h=1}^{H} exp(x'_{sh}\widehat{\beta})$. According to our comment 2 this should actually read $E(n_{s\o}) = \sum_{h=1}^{H} exp(x'_{sh}\widehat{\beta})$. HM eq. (7) estimates $\widehat{\alpha}_s = n_{s\o}/\sum_{h=1}^{H} exp(x'_{sh}\widehat{\beta})$. Plugging these two elements in HM eq. (23) yields our eq. (22).

6 Appendix

The log likelihood function of the HM RUNL is

$$\begin{aligned}\log L(\beta,\lambda) &= \sum_{s=0}^{S}\left[n_{so}\log(P_{so}^{HM}) + \sum_{h=1}^{H} n_{sh}\log(P_{sh}^{HM})\right]\\ &= \sum_{s=0}^{S}\left[n_{so}\log\left(\frac{1}{1+e^{IV_s\lambda_s}}\right) + \sum_{h=1}^{H} n_{sh}\log\left(\frac{e^{x'_{sh}\beta}e^{IV_s(\lambda_s-1)}}{1+e^{IV_s\lambda_s}}\right)\right]\end{aligned}$$

where $IV_s = \log\sum_{g=1}^{H} e^{x'_{sg}\beta}$ is often called the *inclusive value*.

The first order condition with respect to λ_s is for all s

$$\frac{\partial\log L(\beta,\lambda)}{\partial\lambda_s} = -n_{so}\frac{IV_s e^{IV_s\lambda_s}}{1+e^{IV_s\lambda_s}} + \sum_{h=1}^{H} n_{sh}\left(IV_s - \frac{IV_s e^{IV_s\lambda_s}}{1+e^{IV_s\lambda_s}}\right) = 0.$$

Solving the first order condition for λ_s yields

$$\lambda_s = \frac{\log\left(n_{s\emptyset}/n_{so}\right)}{IV_s}.$$

Plugging λ_s into the log likelihood function gives the concentrated log likelihood function

$$\begin{aligned}\log L(\beta) &= \sum_{s=1}^{S}\left[\sum_{h=1}^{H} n_{sh}x'_{sh}\beta - n_{s\emptyset}\log\left(\sum_{h=1}^{H} e^{x'_{sh}\beta}\right)\right.\\ &+ \left. n_{so}\log(n_{so}) + n_{s\emptyset}\log(n_{s\emptyset}) - (n_{so}+n_{s\emptyset})\log(n_{so}+n_{s\emptyset})\right]\end{aligned}$$

which is a function of β only and which is up to a constant identical to the concentrated likelihood functions of the conditional logit, the Poisson and the SB RUNL models. Hence all four models yield identical ML estimates for β.

References

[1] Amemiya, Takeshi (1985) *Advanced Econometrics*, Cambridge: Harvard University Press.

[2] Anderson, Simon, Andre de Palma and Jacques-Francois Thisse (1992) *Discrete Choice Theory of Product Differentiation*. Cambridge: MIT Press.

[3] Brülhart, Marius and Kurt Schmidheiny (2015) Estimating the Rivalness of State-Level Inward FDI. *Journal of Regional Science*, 55(1): 139–148.

[4] 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.

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

[6] Heiss, Florian (2002) Structural Choice Analysis with Nested Logit Models. *The Stata Journal*, 2(3): 227–252.

[7] Herger, Nils and Steve McCorriston (2013) On discrete Choice Models. *Economics Letters*, 120: 288–291.

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

[9] McFadden, Daniel (1977) Quantitative Methods for Analyzing Travel Behavior of Individuals: Some Recent Developments. *Cowles Foundation Discussion Paper*, 474.

[10] McFadden, Daniel (1978) Modelling the Choice of Residential Location. In: A. Karlqvist et al. (Eds.), Spatial Interaction Theory and Plann ing Models, Amsterdam: North-Holland.

[11] Koppelman, Frank S. and Chieh-Hua Wen (1998) Alternative Nested Logit Models: Structure, Properties and Estimation. *Transportation Research B*, 32 (5), 289–298.

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

[13] Schmidheiny, Kurt and Marius Brülhart (2011) On the Equivalence of Location Choice Models: Conditional Logit, Nested Logit and Poisson. *Journal of Urban Economics*, 69(2): 214–222.

[14] Train, Kenneth (2009) *Discrete Choice Methods with Simulation.* Cambridge: Cambridge University Press.