Choice of Residential Telephone Services Case

Estimation of a Nested Logit Model

Files to use with BIOGEME:
Model file: GEV_Tel_NL_unrestricted.mod
Data file: telephone.dat

The application of the IIA McFadden test in the case study on specification testing revealed that the IIA assumption does not hold between the SM and BM alternatives and does not hold among the EF, LF, and MF alternatives as well. We start by giving some examples of possible nesting structures for the Nested Logit (NL) model in Figure 1.

The sample model file describes the first nesting structure shown in Figure 1. The expressions of the utilities for this simple NL model are

$$
\begin{aligned}
V_{BM} &= ASC_{BM} + \beta_{cost} \ln(cost_{BM}) \\
V_{SM} &= \beta_{cost} \ln(cost_{SM}) \\
V_{LF} &= ASC_{LF} + \beta_{cost} \ln(cost_{LF}) \\
V_{EF} &= ASC_{EF} + \beta_{cost} \ln(cost_{EF}) \\
V_{MF} &= ASC_{MF} + \beta_{cost} \ln(cost_{MF}).
\end{aligned}
$$

We show a snapshot of the BIOGEME code in Figure 2. In the first column, we write the name of the nest and in the last column the alternatives that belong to it. Here the alternative numbers must correspond to those used in the utility function under the column ID. The estimation results of the NL model are shown in Table 1.

To be consistent with random utility theory, the inequality $\frac{\mu}{\mu_m} < 1$ with μ being normalized to 1 implies $\mu_m > 1$. To see if this is the case here, we can test the null hypothesis $H_0 : \mu_{meas} = \mu_{flat} = 1$. Since there are multiple restrictions here, we cannot do multiple t-tests. We should do a likelihood

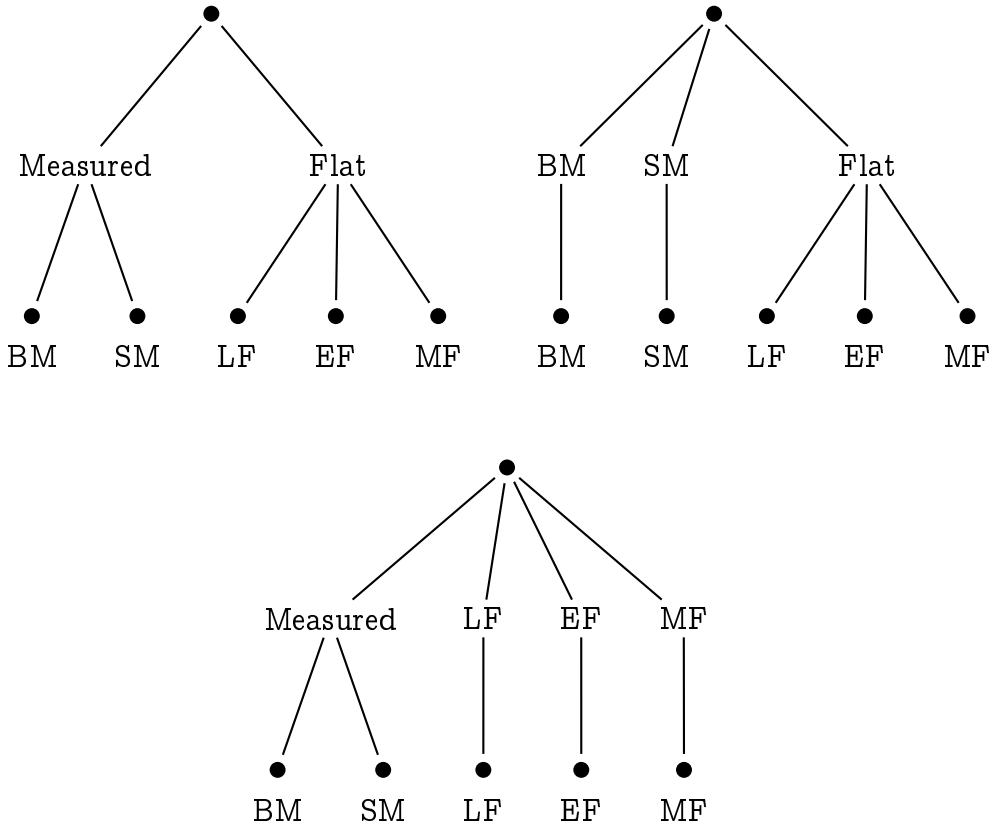

Figure 1: The possible nesting structures

```
[NLNests]
// Name paramvalue  LowerBound UpperBound  status list of alt
N_MEAS    1.0          1.0        10.0       0      1 2
N_FLAT    1.0          1.0        10.0       0      3 4 5
```

Figure 2: BIOGEME snapshot

NL with generic attributes					
Parameter number	Parameter name	Parameter estimate	Robust standard error	Robust *t stat. 0*	Robust *t stat. 1*
1	ASC_{BM}	-0.378	0.117	-3.22	
2	ASC_{LF}	0.893	0.158	5.64	
3	ASC_{EF}	0.847	0.391	2.17	
4	ASC_{MF}	1.41	0.238	5.90	
5	β_{cost}	-1.49	0.243	-6.12	
6	μ_{meas}	2.06	0.573	3.60	1.86
7	μ_{flat}	2.29	0.764	3.00	1.69

Summary statistics
Number of observations $= 434$
$\mathcal{L}(0) = -560.250$
$\mathcal{L}(\hat{\beta}) = -473.219$
$\bar{\rho}^2 = 0.143$

Table 1: NL with generic attributes

ratio test as follows. The test statistic for the null hypothesis is given by

$$-2(\mathcal{L}_R - \mathcal{L}_U) = -2(-477.557 + 473.219) = 8.676$$

where the restricted model is the MNL model (*MNL_Tel_generic.mod*) and the unrestricted model is the nested logit model. The test statistic is asymptotically χ^2 distributed with 2 degrees of freedom since there are 2 restrictions. Since $8.676 > 5.991$ (the critical value of the χ^2 distribution with 2 degrees of freedom at a 95 % level of confidence), we reject the null hypothesis (MNL model) and accept the nested logit model.

The μ_m's of the two nests can be set equal to each other too. This can be done in two ways. One way is to keep the μ_m's fixed to 1 and estimate μ (the related BIOGEME code is shown in Figure 3).

Alternatively, we can also constrain the two nest coefficients to be equal while keeping μ fixed to 1 (Figure 4).

```
[Mu]
// Value LowerBound UpperBound Status
+1.0000000e+00 +0.0000000e+00 +1.0000000e+00 0

[NLNests]
// Name paramvalue  LowerBound UpperBound  status list of alt
N_MEAS    1.0          1.0        10.0       1      1 2
N_FLAT    1.0          1.0        10.0       1      3 4 5
```

Figure 3: BIOGEME snapshot

```
[NLNests]
// Name paramvalue  LowerBound UpperBound  status list of alt
N_MEAS    1.0          1.0        10.0       0      1 2
N_FLAT    1.0          1.0        10.0       0      3 4 5

[ConstraintNestCoef]
// List of pairs of nests for which the associated
// coefficients  must be constrained to be equal
// Syntax: COEF_NEST_A = COEF_NEST_B
N_MEAS = N_FLAT
```

Figure 4: BIOGEME snapshot

The estimation results for this last specification are shown in Table 2.

NL with linear constraints					
Parameter number	Parameter name	Parameter estimate	standard error	*t stat.* 0	Robust *t stat.* 1
1	ASC_{BM}	-0.368	0.110	-3.35	
2	ASC_{LF}	0.882	0.167	5.29	
3	ASC_{EF}	0.833	0.398	2.09	
4	ASC_{MF}	1.39	0.251	5.51	
5	β_{cost}	-1.50	0.257	-5.83	
6	μ_{meas}	2.16	0.519	4.17	2.24
7	μ_{flat}	2.16	0.519	4.17	2.24

Summary statistics
Number of observations $= 434$
$\mathcal{L}(0) = -560.250$
$\mathcal{L}(\hat{\beta}) = -473.288$
$\bar{\rho}^2 = 0.145$

Table 2: NL with linear constraint on nest parameters

Estimation of a Cross-Nested Logit Model with Fixed Alphas

Files to use with BIOGEME:
Model file: GEV_Tel_CNL_fix.mod
Data file: telephone.dat

In this section and the next one, we specify two different Cross-Nested Logit (CNL) models using both fixed and variable degrees of membership. The major premise here is that such specifications are mainly for demonstration purposes. However, an assumption that might make sense is that the standard measured alternative (SM) is likely to be correlated with both

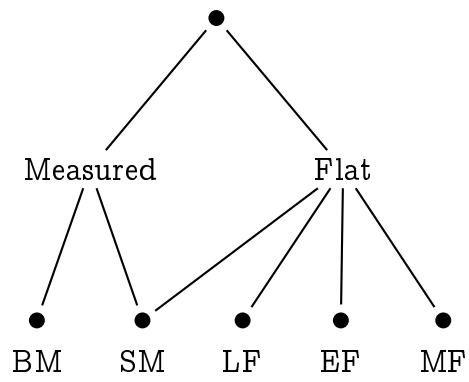

Figure 5: The cross-nested structure

measured and flat options. Indeed, if we look at its definition, it turns out that it may belong to both nests, having also a fixed monthly charge. Based on this hypothesis, the proposed cross-nested structure is shown in Figure 5.

We present the CNL model with the same deterministic utility functions as in the previous model. The corresponding snapshot from the BIOGEME code for this cross-nesting specification is shown in Figure 6.

Note that we define α_{CNL} so that the SM alternative belongs equally to both the flat and the measured nests. This assumption will be relaxed in the next section. Thus, CNL with fixed α's is a restricted model of CNL with variable α's. The estimation results are shown in Table 3.

Cross-Nested Logit Model with Variable Alphas

Files to use with BIOGEME:
Model file: GEV_Tel_CNL_var.mod
Data file: telephone.dat

In the previous CNL model, we assumed that the SM alternative belongs equally to the measured nest and the flat nest by fixing $\alpha_{\text{SM_meas}}$ and $\alpha_{\text{SM_flat}}$ to be equal to 0.5. This assumption can be relaxed, and we can

```
[CNLNests]
// Name        paramvalue LowerBound UpperBound  status
N_MEAS          1.0        1          10         0
N_FLAT          1.0        1          10         0

[CNLAlpha]
// Alt     Nest        value   LowerBound    UpperBound  status
BM      N_MEAS  1        0         1.0         1
SM      N_MEAS     0.5     0         1.0         1
SM      N_FLAT     0.5     0         1.0         1
LF      N_FLAT     1       0         1.0         1
EF      N_FLAT     1       0         1.0         1
MF      N_FLAT     1       0         1.0         1
```

Figure 6: BIOGEME snapshot

CNL estimation results					
Parameter number	Parameter name	Parameter estimate	Robust standard error	Robust *t stat. 0*	Robust *t stat. 1*
1	ASC_{BM}	-0.791	0.0769	-10.28	
2	ASC_{LF}	0.460	0.241	1.91	
3	ASC_{EF}	0.405	0.393	1.03	
4	ASC_{MF}	0.845	0.329	2.57	
5	β_{cost}	-1.21	0.311	-3.91	
6	μ_{meas}	3.14	1.18	2.66	1.81
7	μ_{flat}	2.36	1.14	2.08	1.19
Summary statistics					
Number of observations $= 434$					
$\mathcal{L}(0) = -560.250$					
$\mathcal{L}(\hat{\beta}) = -474.429$					
$\bar{\rho}^2 = 0.141$					

Table 3: CNL estimation results

estimate the share of SM in each nest during the estimation of the model parameters. The corresponding BIOGEME snapshot is shown in Figure 7. From the results presented in Table 4, we see that the alternative SM has a very small share in the flat nest.

We also want to underline the fact that in both CNL specifications the condition

$$\sum_m \alpha_{jm} = 1$$

has been imposed. Such a condition is not necessary for the validity of the model. It is imposed for identification purposes. We refer the interest reader to ? for more theoretical details.

To select between the nested logit and CNL model with variable α's, we can test the null hypothesis $H_0 : \alpha_{SM_flat} = 0$. Since there is a single restriction, we can use either a t-test or a likelihood ratio test which are asymptotically equivalent. The t-statistic with respect to 0 is 0.00, which indicates that α_{SM_flat} is not significantly different from 0, and hence we accept the null hypothesis (nested logit model) and reject the CNL model with variable α's.

We can also do a likelihood ratio test as follows. The test statistic for the null hypothesis is given by

$$-2(\mathcal{L}_R - \mathcal{L}_U) = -2(-473.219 + 473.219) = 0.000$$

where the restricted model is the nested logit model and the unrestricted model is the CNL model. The test statistic is asymptotically χ^2 distributed with 1 degree of freedom since there is 1 restriction. Since $0.000 < 3.841$ (the critical value of the χ^2 distribution with 1 degree of freedom at a 95 % level of confidence), we accept the null hypothesis (nested logit model) and reject the CNL model with variable α's. We can thus conclude that the SM alternative is correlated only with the measured nest but not with the flat nest.

To select between the CNL model with fixed α's and the CNL model with variable α's, we can test the null hypothesis $H_0 : \alpha_{SM_flat} = 0.5$. Since there is a single restriction, we can use either a t-test or a likelihood ratio test

which are asymptotically equivalent. The t-statistic with respect to 0.5 is -0.58, which indicates that α_{SM_flat} is not significantly different from 0.5, and hence we accept the null hypothesis (CNL model with fixed α's) and reject the CNL model with variable α's.

We can also do a likelihood ratio test as follows. The test statistic for the null hypothesis is given by

$$-2(\mathcal{L}_R - \mathcal{L}_U) = -2(-474.429 + 473.219) = 2.420$$

where the restricted model is the CNL model with fixed α's and the unrestricted model is the CNL model with variable α's. The test statistic is asymptotically χ^2 distributed with 1 degree of freedom since there is 1 restriction. Since $2.420 < 3.841$ (the critical value of the χ^2 distribution with 1 degree of freedom at a 95 % level of confidence), we accept the null hypothesis (CNL model with fixed α's) and reject the CNL model with variable α's.

Since both the nested logit model and the CNL model with fixed α's are preferred to the unrestricted model (CNL model with variable α's), we select the nested logit model because it has a higher $\bar{\rho}^2$ than the CNL model with fixed α's (0.143 vs. 0.141).

```
[CNLNests]
// Name       paramvalue LowerBound UpperBound  status
N_MEAS         1.0        1          10        0
N_FLAT         1.0        1          10        0

[CNLAlpha]
// Alt     Nest       value   LowerBound    UpperBound  status
BM      N_MEAS  1       0        1.0        1
SM      N_MEAS     0.5     0        1.0        0
SM      N_FLAT     0.5     0        1.0        0
LF      N_FLAT     1       0        1.0        1
EF      N_FLAT     1       0        1.0        1
MF      N_FLAT     1       0        1.0        1
```

Figure 7: BIOGEME snapshot

CNL with α_{CNL} variable					
Parameter number	Parameter name	Parameter estimate	standard error	*t stat. 0*	*t stat. 1*
1	ASC_{BM}	-0.379	0.863	-0.44	
2	ASC_{LF}	0.893	0.872	1.02	
3	ASC_{EF}	0.847	0.938	0.90	
4	ASC_{MF}	1.41	0.894	1.57	
5	β_{cost}	-1.49	0.257	-5.80	
6	μ_{meas}	2.06	0.575	3.59	1.85
7	μ_{flat}	2.29	0.640	3.58	2.02
8	$\alpha_{\text{SM_meas}}$	1.00	0.855	1.17	0.00
9	$\alpha_{\text{SM_flat}}$	0.000415	0.855	0.00	-1.17

Summary statistics
Number of observations $= 434$
$\mathcal{L}(0) = -560.250$
$\mathcal{L}(\hat{\beta}) = -473.219$
$\bar{\rho}^2 = 0.141$

Table 4: CNL α_{CNL} variable