

AN ABSTRACT OF THE THESIS OF

Dale Wesley Usner for the degree of Doctor of Philosophy in Statistics presented on May 16, 2000. Title: Persistence and Heterogeneity in Habitat Selection Studies.

Abstract approved: _____

Redacted for Privacy

Fred L. Ramsey

Recently the independent multinomial selections model (IMS) with the multinomial logit link has been suggested as an analysis tool for radio-telemetry habitat selection data. This model assumes independence between animals, independence between sightings within an animal, and identical multinomial habitat selection probabilities for all animals.

We propose two generalizations to the IMS model. The first generalization is to allow a Markov chain dependence between consecutive sightings of the same animal. This generalization allows for both positive correlation (individuals persisting in the same habitat class in which they were previously sighted) and negative correlation (individual vacating the habitat class in which they were previously sighted). The second generalization is to allow for heterogeneity. Here, a hierarchical Dirichlet-multinomial distribution is used to allow for variability in selection probabilities between animals. This generalization accounts for over-dispersion of selection probabilities and allows for inference to the population of animals, assuming that the animals studied constitute a random sample from that population..

Both generalizations are one parameter extensions to the multinomial logit model and allow for testing the assumptions of identical multinomial selection probabilities and independence. These tests are performed using the score, Wald, and asymptotic likelihood ratio statistics. Estimates of model parameters are obtained using maximum likelihood techniques, and habitat characteristics are tested using drop-in-deviance statistics.

Using example data, we show that persistence and heterogeneity exist in habitat selection data and illustrate the difference in analysis results between the IMS model and the persistence and heterogeneity models. Through simulation, we show that analyzing persistence data assuming independence between sightings within an animal gives liberal tests of significance for habitat characteristics when the data are generated with positive correlation and conservative tests of significance when the data are generated with negative correlation. Similarly, we show that analyzing heterogeneous data, assuming identical multinomial selection probabilities, gives liberal tests of significance for habitat characteristics.

Persistence and Heterogeneity in Habitat Selection Studies

by

Dale Wesley Usner

A THESIS

submitted to

Oregon State University

in partial fulfillment of
the requirements for the
degree of

Doctor of Philosophy

Presented May 16, 2000
Commencement June 2001

Doctor of Philosophy thesis of Dale Wesley Usner presented on May 16, 2000

APPROVED:

Redacted for Privacy

Major Professor, representing Statistics

Redacted for Privacy

Chair of Department of Statistics

Redacted for Privacy

Dean of Graduate School

I understand that my thesis will become part of the permanent collection of Oregon State University libraries. My signature below authorizes release of my thesis to any reader upon request.

Redacted for Privacy

Dale Wesley Usner, Author

ACKNOWLEDGMENTS

I am indebted to Dr. Fred L. Ramsey for his guidance and support. Dr. Ramsey has contributed substantially to my education through his patience, understanding, and wisdom. I thank Dr. N. Scott Urquhart for always making me think and Dr. Paul Murtaugh for allowing me to stray from assigned paths to satisfy my curiosity.

I appreciate the faculty and staff of the Department of Statistics for their encouragement and contribution to my education and for the funding which allowed my education to continue. A sincere thank you to all for having an open door policy, which fosters a comfortable learning environment.

I offer my love and respect to my wife, Tamara; son, Kord; and daughter, Kaitlyn for being there when I needed them, but most of all for being there when I thought I needed to be alone. I extend my love and appreciation to my brother, Dean, and my parents Peggy and Randy Usner, who did not understand, but accepted and supported, my decision to move across the country to go back to school.

I am also thankful for current and previous colleagues who endured my lunacy and contributed to my education. Finally, I am grateful to Dr. Alan Rossman, who first taught me probability and statistics, for his guidance and encouragement.

TABLE OF CONTENTS

	<u>Page</u>
1. INTRODUCTION	1
2. INDEPENDENT MULTINOMIAL SELECTIONS (IMS) MODEL	4
2.1 Multinomial Logit	4
2.2 Maximization of π Parameterization	5
2.3 Maximization of β Parameterization	5
3. PERSISTENCE MODEL	7
3.1 Markov Chain Transition Probability Matrix	8
3.2 Spectral Decomposition	9
3.3 Correlation	10
3.4 Likelihood	10
3.5 Maximization	11
3.5.1 η, π Parameterization	11
3.5.2 η, β Parameterization	12
3.6 Tests	13
3.6.1 Independent Multinomial Selections Model	13
3.6.2 Habitat Characteristics	13
3.7 Time Between Sightings to Give Approximate Independence	14
3.8 Example	16
3.8.1 Independent Multinomial Selection Analysis	18
3.8.2 Persistence Analysis	19

TABLE OF CONTENTS (Continued)

	<u>Page</u>
3.9 Simulation Results	21
3.9.1 Test Size for IMS vs Persistence	21
3.9.2 Test Power for IMS vs Persistence	22
3.9.3 Habitat Characteristic Test Size	23
3.9.4 Habitat Characteristic Test Power and Model Lack of Fit (LOF) ...	25
3.10 Discussion	27
4. HETEROGENEITY MODEL	29
4.1 Hierarchical Dirichlet-Multinomial Distribution	29
4.2 Likelihood	31
4.3 Maximization	32
4.3.1 γ, π Parameterization	32
4.3.2 γ, β Parameterization	33
4.4 Tests	33
4.4.1 Independent Multinomial Selections Model	33
4.4.2 Habitat Characteristics	34
4.5 Example	35
4.5.1 Independent Multinomial Selection Analysis	36
4.5.2 Heterogeneity Analysis	37
4.6 Simulation Results	39
4.6.1 Test Size for IMS vs Heterogeneity	39
4.6.2 Test Power for IMS vs Heterogeneity	40
4.6.3 Habitat Characteristic Test Size	41
4.6.4 Habitat Characteristic Test Power and Model Lack of Fit (LOF) ...	43
4.7 Discussion	44

TABLE OF CONTENTS (Continued)

	<u>Page</u>
BIBLIOGRAPHY	47
APPENDICES	48
Appendix A.1 Logistic Regression Parameter Bias	49
Appendix A.2 Persistence Parameter, η , Boundaries	51
Appendix A.3 Correlation Between Consecutive Sightings	52
Appendix A.4 Proof of Functional Equality between $\hat{\eta}$ and π	54
Appendix A.5 First and Second Derivatives of $\ell(\eta, \beta X)$	61
Appendix A.6 Delta Method Estimates of Variance for $\hat{\pi}_i$ and $\frac{\hat{\pi}_i}{\hat{\pi}_K}$	66
Appendix A.7 Asymptotic Tests of $H_0 : \eta = 1$ vs $H_1 : \eta \neq 1$	68
Appendix A.8 Sighting History Matrix	72
Appendix A.9 Tables of Size for $H_0 : \eta = 1$ vs $H_1 : \eta \neq 1$	73
Appendix A.10 Tables of Power for $H_0 : \eta = 1$ vs $H_1 : \eta \neq 1$	76
Appendix A.11 Tables of Size for $H_0 : \beta_1 = 0$ vs $H_1 : \beta_1 \neq 0$	79
Appendix A.12 Unconditional Distribution of Y_t	80
Appendix A.13 Moments of Y_t	82
Appendix A.14 Calculation and Reduction of the Loglikelihood, $\ell(\gamma, \pi \mathbf{y})$	83
Appendix A.15 Initial Estimate of $\hat{\gamma}$	85
Appendix A.16 Check for Decreasing Loglikelihood at $\gamma = 0$	86
Appendix A.17 Heterogeneity Multinomial Logit Model Derivatives	87
Appendix A.18 Asymptotic Tests of $H_0 : \gamma = 0$ vs $H_1 : \gamma > 0$	89
Appendix A.19 Tables of Size for $H_0 : \gamma = 0$ vs $H_1 : \gamma > 0$	92
Appendix A.20 Tables of Power for $H_0 : \gamma = 0$ vs $H_1 : \gamma > 0$	95
Appendix A.21 Tables of Power and LOF for $H_0 : \beta_1 = 0$ vs $H_1 : \beta_1 \neq 0$	98

LIST OF TABLES

<u>Table</u>	<u>Page</u>
3.1 Multipliers for sighting times to give approximate independence	16
3.2 Home-range: area, proportion of area available and availability ratio	17
3.3 Independent multinomial selections model estimates of habitat selection probabilities and selection probability ratios	18
3.4 Persistence model estimates of habitat selection probabilities and selection probability ratios	19
3.5 Size of $H_0 : \eta = 1$ vs $H_1 : \eta \neq 1$	21
3.6 Power of $H_0 : \eta = 1$ vs $H_1 : \eta \neq 1$	23
3.7 Size of $H_0 : \beta_1 = 0$ vs $H_1 : \beta_1 \neq 0$	24
3.8 Power and lack of fit of $H_0 : \beta_1 = 0$ vs $H_1 : \beta_1 \neq 0$	26
4.1 Study Area: proportion of area available and availability ratio	35
4.2 Independent multinomial selections model estimates of habitat selection probabilities and selection probability ratios	36
4.3 Heterogeneity model estimates of habitat selection probabilities and selection probability ratios	38
4.4 Reduced heterogeneity model estimates of habitat selection probabilities and selection probability ratios	38
4.5 Size of $H_0 : \gamma = 0$ vs $H_1 : \gamma > 0$	40
4.6 Power of $H_0 : \gamma = 0$ vs $H_1 : \gamma > 0$	41
4.7 Size of $H_0 : \beta_1 = 0$ vs $H_1 : \beta_1 \neq 0$	42
4.8 Power and lack of fit of $H_0 : \beta_1 = 0$ vs $H_1 : \beta_1 \neq 0$	44

This thesis is dedicated to the loving memory of Gail Usner, George Coon, Jane Hill,
and Robert Paul. Peace.

PERSISTENCE AND HETEROGENEITY IN HABITAT SELECTION STUDIES

CHAPTER 1

INTRODUCTION

Radio telemetry is increasingly used in studies to track animals through time. Studies incorporate this technology for many different purposes, such as estimating migratory paths (Mate et al., 1997 and Mate et al., 1998) and determining habitat selection (Vander Heyden and Meslow, 1999) of animals. In habitat selection studies, individual animals are located (sighted) using signals that are sent from radio transmitters (affixed to the animals) to receivers, which have three main forms: hand held, stationary tower, and satellite. After sighting the animals, the habitat type for each sighting location is determined either from maps or from visitations, and measurements of habitat characteristics are made. The objective of habitat selection studies is to determine habitat characteristics which influence the animal population's selection.

Data from these studies consist of two matrices. The first matrix, X , is a sighting history matrix with rows for individual animals, columns for sighting times, and entries identifying the habitat type the animal was sighted in at the given time. The second matrix, Z , is the design matrix, with rows corresponding to habitat types and columns to habitat characteristics. The entries in the design matrix are measurements of each characteristic within each habitat type. An example of the X and Z data matrices might look like:

$$X = \begin{array}{c} \text{Time} \\ \text{Animal 1} \\ \text{Animal 2} \\ \text{Animal 3} \end{array} \begin{array}{ccccc} 1 & 2 & 3 & 4 & 5 \\ \left[\begin{array}{ccccc} 1 & 1 & 1 & 3 & 2 \\ 2 & 2 & 1 & 1 & 1 \\ 1 & 1 & 3 & 3 & 2 \end{array} \right] \end{array} \quad Z = \begin{array}{c} \% \text{ Old} \\ \text{Growth} \\ \text{Clay} \\ \text{Soil} \\ \text{Habitat 1} \\ \text{Habitat 2} \\ \text{Habitat 3} \end{array} \begin{array}{cc} \left[\begin{array}{cc} 54 & 1 \\ 38 & 0 \\ 87 & 1 \end{array} \right] \end{array}$$

X_t will denote the sighting history for animal t , and Z_i will denote the habitat characteristic measurements for habitat i .

Logistic regression is regularly used to analyze habitat selection study data (Ramsey et al., 1994 and Manly et al., 1993). This method is most appropriate when animals leave signs of habitat usage such as nesting. Measurements of habitat characteristics are then made at used sites as well as randomly selected unused sites. Logistic regression is performed to determine if the odds of usage is associated with the characteristics measured. This analysis assumes independence between animals, independence between repeat sightings of the same animal, and identical selection structure between animals.

When animals do not leave signs of habitat usage or a census of the study area is not obtainable, randomly selecting unused sites is not possible. Instead, random sites are selected among those that are available, and an analysis of use versus availability is performed. In discussing the application of logistic regression to resource selection studies, Manly et al. (1993) suggest that this technique “has much to recommend it when census data are available...,” implying that there may be some limitations to the technique when census data are not available. Indeed, logistic regression applied to use versus availability data yields estimates of use versus nonuse selection parameters that are biased toward zero, where the bias increases with the extent of the habitat that is used (Appendix A.1).

The independent multinomial selections (IMS) model, which uses the multinomial logit, has recently been suggested as an analysis tool for habitat selection studies (McCracken et al., 1998). The method allows for estimation of habitat selection probabilities using habitat characteristics such as the proportion of the study area covered by each habitat type as explanatory variables. The independent multinomial selections model assumes identical multinomial habitat selection probabilities for all animals, independence between animals and independence of repeat sightings of the same animal. The IMS model is presented briefly in Chapter 2.

The assumption of independence between repeat sightings of the same animal is a concern when the time interval between sightings of the same animal is relatively short (Thomas and Taylor, 1990). The time interval giving approximate independence between sightings of the same animal could depend on the speed of the species and the sizes of the habitats. For example, a time interval that gives approximate independence for one species may be too short for another slower species, which could result in dependence between sightings of the same animal. Specifically, relatively short intervals can result in positive correlation between consecutive sightings of an animal. Correlation between consecutive sightings is addressed in Chapter 3 using a persistence model.

The assumption that every animal selects its habitats according to the same multinomial selection probabilities leads to a very specific structure for the variances of these probabilities, namely multinomial variance. Often there is more variability in nature than is allotted for by the multinomial variance. Allowing for heterogeneity of multinomial habitat selection probabilities is addressed in Chapter 4.

CHAPTER 2

INDEPENDENT MULTINOMIAL SELECTIONS (IMS) MODEL

Consider T animals the t^{th} of which is sighted n_t times, $t = 1, \dots, T$. Further, consider that each of these sightings is in one of K possible habitats. Let $\boldsymbol{\pi} = [\pi_1, \dots, \pi_K]$, be the population selection probabilities for habitats 1 through K respectively. Assume that animals act independently of each other and that repeat sightings of the same animal are independent. Let $\mathbf{Y}_t = [Y_{t1}, \dots, Y_{tK}]$ be the number of sightings in each of the K habitats for animal t . Then conditioning on n_t , $\mathbf{Y}_t \sim \text{Multinomial}(n_t, \boldsymbol{\pi})$. Similarly, conditioning on the sighting numbers for all T animals, $\mathbf{Y} = [\mathbf{Y}'_1, \dots, \mathbf{Y}'_T]'$, the likelihood of $\boldsymbol{\pi}$ is:

$$L(\boldsymbol{\pi} | [n_1, \dots, n_T], \mathbf{y}) = \prod_{t=1}^T \left(n_t! \prod_{i=1}^K \frac{\pi_i^{y_{ti}}}{y_{ti}!} \right) \quad (2.1)$$

Hence, conditional on $[n_1, \dots, n_T]$, the sighting history matrix, \mathbf{X} , can be reduced to the sufficient statistic: $T(\mathbf{y}) = \left[\sum_{t=1}^T y_{t1}, \dots, \sum_{t=1}^T y_{tK} \right]$, where $T(\mathbf{y})$ registers the total numbers of sightings in each of the habitat types.

2.1 Multinomial Logit

The multinomial logit function (2.2) expresses the log selection probability ratio (SPR) of habitat i to a reference habitat as a linear function of habitat characteristics. Without loss of generality the K^{th} habitat is used as the reference.

$$\ln(\pi_i/\pi_k) = \beta_0(A_i - A_k) + (Z_i - Z_k)\underline{\beta} \quad (2.2)$$

where $\underline{\beta} = [\beta_1, \dots, \beta_p]'$ is the vector of habitat characteristic coefficients, A_i is the log of the total study area covered by habitat i , and β_0 is the coefficient for A_i . Let $\underline{\beta} = [\beta_0, \underline{\beta}']$. Equation (2.2) can be rewritten to give π_i as a function of habitat characteristics:

$$\pi_i = \frac{e^{\beta_0 A_i + Z_i \underline{\beta}}}{\sum_{j=1}^K e^{\beta_0 A_j + Z_j \underline{\beta}}} \quad (2.3)$$

2.2 Maximization of π Parameterization

The maximum likelihood estimates (MLE) of π are obtained by the usual

multinomial estimators: $\hat{\pi}_0 = \left[\frac{\sum_{t=1}^T y_{t1}}{n}, \frac{\sum_{t=1}^T y_{t2}}{n}, \dots, \frac{\sum_{t=1}^T y_{tK}}{n} \right]$, where n is the total number

of sightings over all the animals $\left(n = \sum_{t=1}^T n_t \right)$. Estimates of the standard errors for $\hat{\pi}$

are obtained using the inverse of the observed information matrix.

2.3 Maximization Under β Parameterization

Substituting equation (2.3) for π_i into the likelihood equation (2.1) allows for estimation of habitat selection probabilities in terms of habitat characteristics. The

maximum number of habitat characteristic variables is $K-1$. This limitation results from the multinomial distribution with K classes having only $K-1$ independent parameters $\left(\sum_{i=1}^K \pi_i = 1 \right)$. Maximum likelihood estimates of $\boldsymbol{\beta}$ are obtained using a maximization routine, for example Newton-Raphson techniques. Initial estimates of $\hat{\boldsymbol{\beta}}$ are obtained by first updating $\hat{\boldsymbol{\pi}}_0$ through projection onto the Z -space giving: $\hat{\boldsymbol{\pi}}_{0_{\text{new}}} = Z'(ZZ')^{-1} Z\hat{\boldsymbol{\pi}}_0$. The initial estimates of $\hat{\boldsymbol{\beta}}$ are then found by solving the multinomial logit, equation (2.2), for $\hat{\boldsymbol{\beta}}$ using $\hat{\boldsymbol{\pi}}_{0_{\text{new}}}$. Approximate standard errors for $\hat{\boldsymbol{\beta}}$ are obtained using the inverse of the observed information matrix. Approximate standard errors of the habitat selection probabilities, $\hat{\boldsymbol{\pi}}$, and selection probability ratios, $\hat{\pi}_i/\hat{\pi}_k$, are obtained using the delta approximation method (Appendix A.6).

CHAPTER 3

PERSISTENCE MODEL

If positive correlation is present in data but the modeling technique assumes independence, the estimates of variance from the model will underestimate the true variance in the parameter estimates. Furthermore, the assumption of independence, when it does not exist, leads to an overestimation of degrees of freedom. Both of these estimation errors result in liberal tests of hypotheses for the parameters of interest. If negative correlation is present but not modeled, then estimates of variance overestimate the true variance, which results in conservative tests of hypotheses. Manly et al. (1993) recognize the possibility of correlation between sightings of the same animal and suggest analyzing each animal separately – after which tests of parameters can be made using the mean parameter value across animals and its sample standard error. Swihart and Slade (1985a) show that estimates of home-range derived from correlated observations have negative bias compared to estimates derived from independent observations. Swihart and Slade (1985b) also present tests of independence of animal locations, using sighting coordinates, and introduce a technique to estimate the sampling interval needed to attain approximate independence.

In this paper, we introduce a persistence model, which is a one-parameter generalization of the independent multinomial selections model, to account for correlation between sightings of the same animal. Its use yields more accurate tests of hypotheses than the IMS model when correlation exists. Furthermore, this model allows for estimation of a time interval between sightings to give approximate independence. Independence between animals and identical multinomial habitat selection probabilities for all animals are assumed.

3.1 Markov Chain Transition Probability Matrix

Consider that there is temporal dependence between sightings of the same animal and that this dependence follows a Markov chain. The probability of being sighted in habitat i at time $q+1$ depends only on the sighting at time q . These probabilities can be described with a transition probability matrix. The form of the transition probability matrix considered in this paper,

$$\Pi(\eta) = \begin{bmatrix} 1-\eta(1-\pi_1) & \eta\pi_2 & \cdots & \eta\pi_{K-1} & \eta\pi_K \\ \eta\pi_1 & 1-\eta(1-\pi_2) & \cdots & \eta\pi_{K-1} & \eta\pi_K \\ \eta\pi_1 & \eta\pi_2 & \ddots & \vdots & \vdots \\ \vdots & \vdots & \cdots & 1-\eta(1-\pi_{K-1}) & \eta\pi_K \\ \eta\pi_1 & \eta\pi_2 & \cdots & \eta\pi_{K-1} & 1-\eta(1-\pi_K) \end{bmatrix},$$

incorporates one additional parameter, η , to the IMS model. The persistence parameter, η , is bounded between 0 and the minimum of $\left\{ \frac{1}{\pi_i}, \frac{1}{1-\pi_i} : \forall i \right\}$ (Appendix A.2), where the π_i 's are the stationary probabilities of being sighted in each of the habitat types. The stationarity is straightforward using equations (3.1) or (3.6) below.

Note that at $\eta = 1$,

$$\Pi(1) = \begin{bmatrix} \pi_1 & \pi_2 & \cdots & \pi_{K-1} & \pi_K \\ \pi_1 & \pi_2 & \cdots & \pi_{K-1} & \pi_K \\ \pi_1 & \pi_2 & \ddots & \vdots & \vdots \\ \vdots & \vdots & \cdots & \pi_{K-1} & \pi_K \\ \pi_1 & \pi_2 & \cdots & \pi_{K-1} & \pi_K \end{bmatrix}$$

which is the probability transition matrix assuming independence between sightings (IMS model). Moreover, as η decreases from 1 to 0, the probabilities on the main diagonal increase over that obtained under the IMS model; thereby increasing the probability that the animals will persist in the habitat in which they were previously sighted. The off-diagonals decrease proportionately to their probabilities under the IMS model. As η increases from 1 to $\min\left\{\frac{1}{\pi_i}, \frac{1}{1-\pi_i} : \forall i\right\}$, the probabilities on the main diagonal decrease from that obtained under the IMS model; thereby increasing the probability that the animals will vacate the habitat in which they were previously sighted. The off-diagonals increase proportionately.

3.2 Spectral Decomposition

The form of the probability transition matrix, $\Pi(\eta)$, given above has a simple spectral decomposition. The eigenvalues of $\Pi(\eta)$ are $\varepsilon_1 = 1$ and $\varepsilon_2 = \varepsilon_3 = \dots = \varepsilon_k = 1 - \eta$. The vectors $\mathbf{1}$ and $\boldsymbol{\pi}$ are the right and left eigenvectors associated with ε_1 . This results in the following spectral decomposition of $\Pi(\eta)$: If $\Lambda(\eta) = \text{diag}(1, 1 - \eta, \dots, 1 - \eta)$, then

$$\begin{aligned}\Pi(\eta) &= \mathbf{R}\Lambda(\eta)\mathbf{L}' \\ &= \mathbf{R}\left[\Lambda(1) + (1 - \eta)(\mathbf{I}_k - \Lambda(1))\right]\mathbf{L}' \\ &= \Pi(1) + (1 - \eta)\left[\mathbf{I}_k - \Pi(1)\right]\end{aligned}\tag{3.1}$$

Here, $\mathbf{R}\mathbf{L}' = \mathbf{I}_k$ by orthogonality of eigenvectors. This representation of $\Pi(\eta)$ will be used later to determine sighting times to give approximate independence.

3.3 Correlation

The correlation between consecutive sightings of the same animal under the persistence model is $1 - \eta$ (Appendix A.3). Hence, $\eta \in (0, 1)$ leads to positive correlation between sightings, and $\eta \in \left(1, \min \left\{ \frac{1}{\pi_i}, \frac{1}{1 - \pi_i} : \forall i \right\} \right)$ leads to negative correlation between sightings.

3.4 Likelihood

Let f_{ti} be an indicator of the initial sighting in habitat i for animal t :

$$f_{ti} = \begin{cases} 1 & \text{animal } t \text{ initially sighted in habitat } i \\ 0 & \text{animal } t \text{ initially sighted in habitat } j \neq i \end{cases}$$

Let n_{tji} be the number of transitions from habitat j to habitat i for animal t and v_{ti} be the number of transitions to habitat i from habitats other than i for animal t , $v_{ti} = \sum_{j \neq i} n_{tji}$.

Then the probability of an observed sighting history for a given animal t under the persistence model is:

$$\begin{aligned} P(X_t | \boldsymbol{\pi}, \eta) &\propto \prod_{i=1}^K \pi_i^{f_{ti}} (\eta \pi_i)^{v_{ti}} (1 - \eta(1 - \pi_i))^{n_{ti}} \\ &= \eta^{V_t} \prod_{i=1}^K \pi_i^{f_{ti} + v_{ti}} (1 - \eta(1 - \pi_i))^{n_{ti}} \end{aligned} \tag{3.2}$$

where $V_t = \sum_{i=1}^K v_{ti}$. Assuming independence between animals, the likelihood for the persistence model parameters is:

$$\begin{aligned}
 L(\boldsymbol{\pi}, \eta | \mathbf{X}) &\propto \prod_{t=1}^T \left[\eta^{V_t} \prod_{i=1}^K \pi_i^{f_{ti} + v_{ti}} (1 - \eta(1 - \pi_i))^{n_{ti}} \right] \\
 &= \eta^{\sum_{t=1}^T V_t} \prod_{i=1}^K \pi_i^{\sum_{t=1}^T f_{ti} + \sum_{t=1}^T v_{ti}} (1 - \eta(1 - \pi_i))^{\sum_{t=1}^T n_{ti}} \\
 &= \eta^{V \cdot} \prod_{i=1}^K \pi_i^{f_{\cdot i} + v_{\cdot i}} (1 - \eta(1 - \pi_i))^{n_{\cdot i}}
 \end{aligned} \tag{3.3}$$

Hence the sighting history matrix can be reduced to the sufficient statistic:

$$T(\mathbf{X}) = [V_{\cdot}, f_{\cdot i} + v_{\cdot i}, n_{\cdot ij} : i = 1, \dots, K].$$

3.5 Maximization

3.5.1 η, π Parameterization

The Newton-Raphson algorithm provides a convenient method for obtaining maximum likelihood estimators (MLE) for the π_i 's and η . The convergence of this technique is improved using the equality:

$$\hat{\eta}_{\pi} = \frac{B}{(K-1)n_{\cdot 11} + B(1 - \hat{\pi}_1)} \tag{3.4}$$

where

$$B = \sum_{i=2}^K (f_{\cdot i} + v_{\cdot i}) \left(\frac{1 - \hat{\pi}_i}{\hat{\pi}_i} \right) - (f_{\cdot 1} + v_{\cdot 1}) \left(\frac{(K-2) + \hat{\pi}_1}{\hat{\pi}_1} \right) + V. \quad (3.5)$$

The notation $\hat{\eta}_\pi$ is used to indicate the MLE of η under the η, π parameterization. See Appendix A.4 for inductive proof. Initial estimates of $\hat{\pi}$ are $\hat{\pi}_0$, the maximum likelihood estimates under the IMS model.

3.5.2 η, β Parameterization

Substituting the multinomial logit equation (2.3) into equation (3.3) gives the likelihood in terms of $\beta = [\beta_0, \underline{\beta}']$, log area, and the habitat characteristics Z . A similar relationship to equation (3.4) between the parameter estimates, $\hat{\eta}$ and $\hat{\beta}$ was not found – unless the number of β 's is equal to $K-1$. Therefore, in these cases $\hat{\eta}$ had to be found through the Newton-Raphson algorithm as well as $\hat{\beta}$. However, using $\hat{\eta}_\pi$ as the initial value of η in the maximization routine under the multinomial logit representation accelerated convergence. Initial estimates of $\hat{\beta}$ are obtained using the same algorithm discussed under the IMS model.

The first and second derivatives of the loglikelihood, $\ell(\eta, \beta | X)$, are given in Appendix A.5. Approximate standard errors of $\hat{\beta}$ and $\hat{\eta}$ are obtained from the inverse of the observed information matrix. Approximate standard errors of the habitat selection probabilities, $\hat{\pi}$, and selection probability ratios, $\hat{\pi}_i / \hat{\pi}_K$, are obtained using the delta method (Appendix A.6).

3.6 Tests

3.6.1 *Independent Multinomial Selections Model*

Tests of the null hypothesis of independent multinomial selections between sightings within an animal versus the alternative hypothesis of persistence are conducted by testing $H_0 : \eta = 1$ vs $H_1 : \eta \neq 1$. The score, Wald, and asymptotic likelihood ratio chi-squared statistics were constructed to test for independence (Appendix A.7). These tests are calculated under the full (saturated) model, where there are $K-1$ parameters.

3.6.2 *Habitat Characteristics*

The test of no effect of habitat characteristic j , $H_0 : \beta_j = 0$ vs $H_1 : \beta_j \neq 0$ is completed through the usual Type III drop-in-deviance chi-squared test statistic. The test of the coefficient for log area equal to one, $H_0 : \beta_0 = 1$, vs $H_1 : \beta_0 \neq 1$ is also constructed. This test determines if log area can be offset – thereby accounting for random selection of habitat classes giving selection probabilities proportional to area. Rejecting $H_0 : \beta_0 = 1$ could be the result of missing a necessary habitat characteristic in the model, or it could be the result of some other random selection procedure by the animal, such as randomly picking an initial point and walking a straight line in some direction.

Consider the case where the habitat characteristics are indicator variables for the habitat classes, habitat class K is the reference class, and area is offset. Here, the test of $H_0 : \beta_j = 0$ tests if the log selection probability ratio (SPR) of habitat j to habitat K is equal to the ratio of the areas of habitat j to habitat K . Positive β_j gives a SPR of

habitat j to habitat K that is greater than expected given the respective areas. Negative β_j gives a SPR of habitat j to habitat K that is less than expected given the respective areas.

3.7 Time Between Sightings to Give Approximate Independence

While the persistence model incorporates a time dependence structure and therefore can be used to more accurately model data with time dependence, an investigator might be interested in the length of time between sightings that is necessary to give approximate independence. For example, in almost all studies there are time or cost constraints to the amount of sampling that can be done or the number of experimental units that can be treated. One way to reduce costs in habitat selection studies is to make fewer sightings of an individual animal. An efficient way of approaching this reduction in sightings is to determine the length of time between sightings that gives approximate independence. When the sightings of an individual animal are independent, the investigator obtains more information per unit cost than if there is dependence.

The persistence model allows for easy estimation of the necessary time between sightings to give approximate independence. Consider that data from a particular study, with Δ time units between sightings, show persistence model type dependence with estimate $\hat{\eta}$. Therefore, the Markov transition probability matrix has the form $\Pi(\eta)$, and the transition probability matrix for sightings Δq time units apart has the form $[\Pi(\eta)]^q$. Now by (3.1),

$$\begin{aligned}
[\Pi(\eta)]^q &= [\Pi(1) + (1-\eta)(I_k - \Pi(1))]^q \\
&= \sum_{j=0}^q \binom{q}{j} \Pi(1)^{q-j} (1-\eta)^j (I_k - \Pi(1))^j \\
&= [\Pi(1)]^q + (1-\eta)^q [(I_k - \Pi(1))]^q \\
&= \Pi(1) + (1-\eta)^q (I_k - \Pi(1)),
\end{aligned} \tag{3.6}$$

because $\Pi(1)$ and $I_k - \Pi(1)$ are idempotent projection operators on orthogonal spaces. Here, $1 - (1 - \eta)^q$ is the persistence parameter for sightings Δq time units apart. The multiplier used to calculate the length of time between sightings necessary to obtain the desired persistence, c , is found by solving $1 - (1 - \eta)^q = c$ for q , giving:

$$q = \frac{\ln(1-c)}{\ln(1-\eta)} \tag{3.7}$$

For approximate independence between sightings, we suggest that c be between 0.90 and 1.00. This stems from simulation results, which show that both the size of the test of habitat selection characteristics, $H_0 : \beta_1 = 0$ vs $H_1 : \beta_1 \neq 0$, and the percent of model lack of fit under the IMS model are close to those obtained under the persistence model when $\eta \in [.90, 1]$ (Tables 3.7 and 3.8). The entries in Table 3.1 are the multipliers necessary to obtain persistence parameter c given $\hat{\eta}$.

Table 3.1 Multipliers for sighting times to give approximate independence

$\hat{\eta}$	c		
	0.90	0.95	0.99
0.4	4.51	5.86	9.02
0.5	3.32	4.32	6.64
0.6	2.51	3.27	5.03
0.7	1.91	2.49	3.82
0.8	1.43	1.86	2.86

For example, if a given study sights animals every Δ time units, and the estimated persistence parameter for the study is $\hat{\eta} = 0.4$, then 4.51Δ time units between sightings are needed to obtain a persistence parameter of 0.90, and 9.02Δ time units between sightings are needed to obtain a persistence parameter of 0.99. Note that for a given $\hat{\eta}$ the higher the c, the longer the time interval necessary between sightings. Therefore, the investigator must strike a compromise between the suggested time interval and the lifespan of the transmitter.

3.8 Example

As part of a study to determine home range and habitat selection by female black bears (*Ursus americanus*) in central Oregon Cascades, female black bears were captured and radiocollared in the springs of 1993-1995 (Vander Heyden and Meslow, 1999). The bears were then radio-tracked, on average, every other day, with attempts to sight at three different times of the day (morning, afternoon, and evening) (Vander Heyden and Meslow, 1999). The study area was divided into six habitat classes: grass-forb, shrub, open-canopy sapling/pole, closed-canopy sapling/pole, closed-canopy mature timber, and open-canopy mature timber (exact definitions are given in Vander Heyden and Meslow, 1999).

Throughout 1993 to 1995 a total of twelve bears were collared and sighted. Home ranges were determined each year for each bear using adaptive kernel and minimum convex polygon algorithms (Vander Heyden and Meslow, 1999). Within these home ranges random samples of locations were made to determine availability of habitat. The size, in hectares, of the habitat patch surrounding each of the random locations was determined using a LANDSAT Thematic Mapper image (Vander Heyden and Meslow, 1999). We estimate habitat class availability by the sum of the areas of unique habitat patches of a given class.

The twelfth bear, analyzed herein, was collared in 1995 and sighted 64 times. The bear was never sighted in the Shrub habitat; therefore, this habitat was removed from the habitat selection analyses discussed below. For the five remaining habitat classes the estimated areas, proportions of home range coverage, and area availability ratios with respect to open-canopy mature timber habitat are given in Table 3.2.

Table 3.2 Home-range: area, proportion of area available and availability ratio

Habitat Class	Grass Forb	Open Sapling	Closed Sapling	Closed Mature	Open Mature
Area	1291.5	82.4	387.8	1693.1	620.2
Proportion Available	0.3169	0.0202	0.0952	0.4155	0.1522
Availability Ratio	2.0824	0.1329	0.6253	2.7299	1.0000

Habitat selections for this bear were analyzed employing log area and indicator variables for grass-forb, sapling/pole, and closed-canopy as explanatory variables. Both the persistence and independent multinomial selections models were fit to the data using open-canopy mature timber as the reference habitat to calculate the selection probability ratios (SPR).

3.8.1 Independent Multinomial Selection Analysis

The IMS model shows that the coefficient for log area, 1.3633 (1.1812), is not significantly different from one with drop-in-deviance statistic 0.0950 on one degree of freedom, giving p-value = 0.7579. The model offsetting area shows that the selection probability ratios of grass-forbs and closed-canopy are significantly smaller than their respective availability ratios with both p-values < 0.0001, and the SPR of sapling/pole is marginally significantly larger than its availability ratio with p-value = 0.0714. The estimated SPR of grass-forbs is $e^{-1.81546} = 0.1628$ times its availability ratio with 90% Wald CI = (0.0874, 0.3031), the estimated SPR of closed-canopy is $e^{-1.54131} = 0.2141$ times its availability ratio with 90% Wald CI = (0.1363, 0.3363), and the estimated SPR of sapling/pole is $e^{0.60677} = 1.8345$ times its availability ratio with 90% Wald CI = (1.0893, 3.0894). Table 3.3 gives the estimated habitat selection probabilities and SPR's with approximate standard errors for each habitat.

Table 3.3 Independent multinomial selections model estimates of habitat selection probabilities and selection probability ratios

IMS Habitat Class	Grass Forb	Open Sapling	Closed Sapling	Closed Mature	Open Mature
Prob. of Use	0.1405	0.1011	0.1018	0.2422	0.4145
SE	0.0431	0.0285	0.0286	0.0473	0.0560
SPR	0.3389	0.2438	0.2456	0.5845	1.0000
SE	0.1281	0.0773	0.0969	0.1605	

Comparing Table 3.3 with Table 3.2 shows the probabilities of use for grass-forb and closed-canopy mature timber are smaller than their respective proportions of area available. The probabilities of use for open-canopy sapling/pole and open-canopy mature timber are larger than their respective proportions of area available.

3.8.2 Persistence Analysis

The persistence model estimated η as 0.6749 with standard error: 0.0876. The asymptotic likelihood ratio statistic for testing $H_0: \eta = 1$ (IMS model) vs $H_a: \eta \neq 1$ (persistence model) was 18.1267 with 1 df, for a p-value < 0.0001 . Therefore, there is substantial evidence of persistence. Appendix A.8 contains the sighting history matrix for this bear, which shows the tendency of the bear to remain in the habitat in which it was previously sighted.

Furthermore, this model shows that the coefficient for log area, 0.8943 (1.4520), is not significantly different from one with drop-in-deviance statistic 0.0054 on one degree of freedom, giving p-value = 0.9415. The model offsetting area shows that the selection probability ratios of grass-forbs and closed-canopy are significantly smaller than available with p-values: 0.0002 and 0.0009 respectively. Here, sapling/pole does not have a statistically significant effect on SPR, p-value = 0.1098. After removing the indicator for sapling/pole, the estimated SPR of grass-forbs is $e^{-2.06350} = 0.1270$ times its availability ratio with 90% Wald CI = (0.0489, 0.3296), and the estimated SPR of closed-canopy is $e^{-1.26044} = 0.2835$ times its availability ratio with 90% Wald CI = (0.1531, 0.5251). Table 3.4 gives the estimated selection probabilities and SPR's with approximated standard errors for each habitat.

Table 3.4 Persistence model estimates of habitat selection probabilities and selection probability ratios

Persistence Habitat Class	Grass Forb	Open Sapling	Closed Sapling	Closed Mature	Open Mature
Prob. of Use	0.1126	0.0566	0.0755	0.3296	0.4258
SE	0.0547	0.0104	0.0159	0.0694	0.0783
SPR	0.2645	0.1329	0.1773	0.7740	1.0000
SE	0.1533	-	0.0664	0.2900	

Comparing Table 3.4 with Table 3.2 shows the probabilities of use for grass-forb, closed-canopy sapling/pole, and closed-canopy mature timber are smaller than their respective proportions of area available, and the probabilities of use for open-canopy sapling/pole and open-canopy mature timber are larger than their respective proportions of area available for each. Note that there is no standard error for the selection probability ratio of open-canopy sapling/pole because the specified model forces this SPR to be identical to its availability ratio.

This analysis, using the persistence model, showed the effect of sapling/pole to be non-significant. However, analysis of the data assuming independence between sightings results in a significant effect of sapling/pole, where the selection probability ratio is higher than the available ratio for habitats categorized as sapling/pole. Therefore, not only does this bear's data show substantial evidence of persistence, but it also shows that analysis assuming independence gives different results from analysis incorporating correlation.

As further illustration of persistence, we randomly ordered the sighting history (Appendix A.8), which should produce a history that has approximate independence between sightings. Analysis of the randomized history under the persistence model gives persistence parameter estimate of 1.0320 with standard error: 0.0639. The asymptotic likelihood ratio statistic was 0.2285 with p-value = 0.6326. Therefore, the persistence parameter of this random ordering does not test different from 1, as expected.

The sightings for these data were made approximately every 48 hours, which would seem long enough to give independence between sightings. However, the final model shows persistence between sightings with $\hat{\eta} = 0.6756$. Using (3.7) to calculate the amount of time necessary between sightings to obtain a persistence

parameter of 0.90 gives: $48 \frac{\ln(1-0.90)}{\ln(1-0.6756)} \approx 98$ hours or four days instead of two.

The amount of time necessary to obtain a persistence parameter of 0.95 is:

$48 \frac{\ln(1-0.95)}{\ln(1-0.6756)} \approx 128$ hours or five and one-third days.

3.9 Simulation Results

3.9.1 Test Size for IMS vs Persistence

The sizes of the score, Wald, and asymptotic LR tests for independence between sightings of an animal, $H_0 : \eta = 1$ vs $H_1 : \eta \neq 1$, were calculated in a saturated model using 5000 simulations with level $\alpha = 0.10, 0.05$, and 0.01 tests. Table 3.5 shows results for level $\alpha = 0.05$ tests with $K = 4$ habitats and $\boldsymbol{\pi} = [.25, .25, .25, .25]$.

Appendix A.9 gives further results for $K = 4$ habitats. The size of the test is generally close to the expected level, ranging between 8.5 to 11.0, 4.0 to 6.0, and 0.8 to 1.2 for $\alpha = 0.10, 0.05$, and 0.01 respectively. Size does not appear to be affected by the number of animals, T ; the number of sightings per animal, n_t ; or the habitat selection probabilities, $\boldsymbol{\pi}$. The results for $K = 3$ and 5 habitats are similar. With few sightings, ten per animal, the size ranges increase, and in certain simulations, convergence difficulties arose; this happened when at least one habitat had a small probability of selection, 0.10 , and the number of animals and sightings were small.

Table 3.5 Size of $H_0 : \eta = 1$ vs $H_1 : \eta \neq 1$

T	n_t	$\boldsymbol{\pi}$	SS	Size x 100		
				Score	Wald	LRT
50	50	.25, .25, .25, .25	5000	4.94	4.82	4.92
50	30	.25, .25, .25, .25	5000	4.96	4.86	4.92
50	10	.25, .25, .25, .25	5000	4.70	4.10	4.54
30	50	.25, .25, .25, .25	5000	5.26	5.26	5.34
30	30	.25, .25, .25, .25	5000	5.14	5.16	5.06
30	10	.25, .25, .25, .25	5000	5.20	4.38	5.04
10	50	.25, .25, .25, .25	5000	5.14	5.18	5.20
10	30	.25, .25, .25, .25	5000	5.04	5.10	5.02
10	10	.25, .25, .25, .25	4942	6.07	4.69	5.83

3.9.2 Test Power for IMS vs Persistence

The power of the score, Wald, and asymptotic LR tests for $H_0 : \eta = 1$ vs $H_1 : \eta \neq 1$ were calculated in a saturated model using 1000 simulations with level $\alpha = 0.05$ tests. Power was calculated for various η . Table 3.6 shows results for $K = 4$ habitats, $\boldsymbol{\pi} = [.25, .25, .25, .25]$, and $\eta = 0.8$ & 0.9 . Further results are given in Appendix A.10. These tables show that power to detect a difference from $\eta = 1$ increases rapidly as η moves away from one in either direction. At $\eta = 0.8$, the power is close to 1 in all cases except $T = n_t = 10$. Power also increases with increasing T , n_t , and K . At $\eta = 0.90$, $T = 50$, and $n_t = 50$ the power is 1.00, dropping to around 0.32 with $T = 10$ and $n_t = 10$. When $\eta < 1$ the Wald test gives the highest power, with the LRT test giving the next highest power. When $\eta > 1$ the Score test gives the highest power, with the LRT test giving the next highest power. The power of these tests do not appear to depend on habitat selection probabilities used to generate the data; hence, results for only one $\boldsymbol{\pi}$ per K are shown.

Table 3.6 Power of $H_0 : \eta = 1$ vs $H_1 : \eta \neq 1$

T	n_t	π	η	Power x 100		
				Score	Wald	LRT
50	50	.25, .25, .25, .25	0.8	100.00	100.00	100.00
50	30	.25, .25, .25, .25	0.8	100.00	100.00	100.00
50	10	.25, .25, .25, .25	0.8	100.00	100.00	100.00
30	50	.25, .25, .25, .25	0.8	100.00	100.00	100.00
30	30	.25, .25, .25, .25	0.8	100.00	100.00	100.00
30	10	.25, .25, .25, .25	0.8	99.90	100.00	99.90
10	50	.25, .25, .25, .25	0.8	100.00	100.00	100.00
10	30	.25, .25, .25, .25	0.8	100.00	100.00	100.00
10	10	.25, .25, .25, .25	0.8	76.65	85.97	81.26
50	50	.25, .25, .25, .25	0.9	100.00	100.00	100.00
50	30	.25, .25, .25, .25	0.9	100.00	100.00	100.00
50	10	.25, .25, .25, .25	0.9	91.90	94.70	93.10
30	50	.25, .25, .25, .25	0.9	100.00	100.00	100.00
30	30	.25, .25, .25, .25	0.9	99.70	99.90	99.70
30	10	.25, .25, .25, .25	0.9	72.50	79.70	75.40
10	50	.25, .25, .25, .25	0.9	95.10	96.90	95.90
10	30	.25, .25, .25, .25	0.9	74.40	82.60	77.50
10	10	.25, .25, .25, .25	0.9	24.47	39.58	30.72

3.9.3 Habitat Characteristic Test Size

The size of the asymptotic likelihood ratio (drop-in-deviance) test for habitat characteristics, $H_0 : \beta_1 = 0$ vs $H_1 : \beta_1 \neq 0$, was calculated using 2000 simulations with level $\alpha = 0.05$ tests. Results are given in Table 3.7 for $K = 3$, $T = 50$, and various n_t and η . The entry under Z is the indicator variable tested (all indicator variables gave similar results, the results for $[1, 0, 0]$ is presented); “Pers.” entries are 100 times the size of the test under the persistence model; and “IMS” entries are 100 times the size of the test under the IMS model. At $\eta = 0.4$, the persistence model appropriately shows a size of about 0.05. However, the IMS model shows liberal tests of significance with sizes between 0.28 and 0.33. The size obtained under the persistence model does not appear to

depend on η . However, the size decreases under the IMS model as $\eta \rightarrow 1.0$, until it reaches about 0.05 at $\eta = 1.0$. Moreover, as η increases past 1.0 the size continues to decrease.

Table 3.7 Size of $H_0 : \beta_1 = 0$ vs $H_1 : \beta_1 \neq 0$

T	n_t	π	η	Z	Size x 100	
					Pers.	IMS
50	50	1/3, 1/3, 1/3	.4	1, 0, 0	5.55	33.35
50	30	1/3, 1/3, 1/3	.4	1, 0, 0	4.85	31.55
50	10	1/3, 1/3, 1/3	.4	1, 0, 0	5.10	27.90
50	50	1/3, 1/3, 1/3	.6	1, 0, 0	4.80	19.40
50	30	1/3, 1/3, 1/3	.6	1, 0, 0	5.25	20.05
50	10	1/3, 1/3, 1/3	.6	1, 0, 0	4.65	17.75
50	50	1/3, 1/3, 1/3	.8	1, 0, 0	4.90	10.85
50	30	1/3, 1/3, 1/3	.8	1, 0, 0	5.45	11.20
50	10	1/3, 1/3, 1/3	.8	1, 0, 0	4.25	9.75
50	50	1/3, 1/3, 1/3	.9	1, 0, 0	4.65	7.40
50	30	1/3, 1/3, 1/3	.9	1, 0, 0	4.55	7.40
50	10	1/3, 1/3, 1/3	.9	1, 0, 0	5.60	8.00
50	50	1/3, 1/3, 1/3	.95	1, 0, 0	4.65	6.10
50	30	1/3, 1/3, 1/3	.95	1, 0, 0	4.50	6.05
50	10	1/3, 1/3, 1/3	.95	1, 0, 0	5.85	7.05
50	50	1/3, 1/3, 1/3	1.0	1, 0, 0	5.55	5.65
50	30	1/3, 1/3, 1/3	1.0	1, 0, 0	4.75	5.10
50	10	1/3, 1/3, 1/3	1.0	1, 0, 0	5.50	5.70
50	50	1/3, 1/3, 1/3	1.2	1, 0, 0	4.90	1.65
50	30	1/3, 1/3, 1/3	1.2	1, 0, 0	5.20	1.95
50	10	1/3, 1/3, 1/3	1.2	1, 0, 0	4.55	1.80

The reduction in size under the IMS model as η increases is to be expected. As $\eta \rightarrow 1.0$, the positive correlation between consecutive sightings of the same animal (Appendix A.3) decreases, until the independence model is obtained at $\eta = 1.0$. As η increases past 1.0, the consecutive sightings within an animal become negatively

correlated. Therefore, failing to adjust for persistence type dependence leads to liberal tests when $\eta < 1$ and conservative tests when $\eta > 1$. Appendix A.11 gives similar tables for $T = 30$ and 10 . Size does not appear to depend on T or n_t under either model. Similar results are seen using $K = 4$ and 5 .

3.9.4 *Habitat Characteristic Test Power and Model Lack of Fit (LOF)*

The power of the asymptotic likelihood ratio (drop-in-deviance) test for habitat characteristics, $H_0 : \beta_1 = 0$ vs $H_1 : \beta_1 \neq 0$, was calculated using 2000 simulations with a level $\alpha = 0.05$ test. Sighting history matrices were generated using various non-uniform selection probabilities. The logs of the selection probabilities were used as habitat characteristic variables in the models. Model lack of fit was tested between the fitted model with one explanatory variable and the saturated model with $K-1$ explanatory variables. The percent of model lack of fit using a level $\alpha = 0.05$ test was recorded.

Table 3.8 gives results for selection probabilities, $\pi = [.20, .23, .27, .30]$. This table shows that the power to detect a difference from uniform habitat selection probabilities decreases as η , T , and n_t decrease. While these results hold for both models, there is a greater effect seen under the persistence model. The percent of LOF is between 1.5 and 2.8 under the persistence model and does not appear to depend on η , T , or n_t . Under the IMS model, the percent of LOF is between 0.5 to 37.4, increasing as η decreases. There does not appear to be a clear effect of T or n_t on LOF under the IMS model. Similar effects on power and LOF are seen for $K = 3$ and 5 habitats.

Table 3.8. Power and lack of fit of $H_0 : \beta_1 = 0$ vs $H_1 : \beta_1 \neq 0$

T	n_t	π	η	Power x 100		LOF	
				Pers.	IMS	Pers.	IMS
50	50	.20, .23, .27, .30	.4	98.90	99.90	1.95	37.35
50	30	.20, .23, .27, .30	.4	92.10	98.05	2.20	35.10
50	10	.20, .23, .27, .30	.4	55.75	78.90	2.15	30.65
50	50	.20, .23, .27, .30	.6	99.95	100.00	1.85	18.60
50	30	.20, .23, .27, .30	.6	98.80	99.70	2.15	15.85
50	10	.20, .23, .27, .30	.6	71.95	84.20	1.60	15.90
50	50	.20, .23, .27, .30	.8	100.00	100.00	1.85	7.55
50	30	.20, .23, .27, .30	.8	99.90	100.00	2.45	8.20
50	10	.20, .23, .27, .30	.8	84.25	89.65	2.60	7.05
50	50	.20, .23, .27, .30	.9	100.00	100.00	1.30	3.15
50	30	.20, .23, .27, .30	.9	100.00	100.00	2.00	3.65
50	10	.20, .23, .27, .30	.9	88.00	90.30	2.20	3.80
50	50	.20, .23, .27, .30	1.0	100.00	100.00	2.05	2.10
50	30	.20, .23, .27, .30	1.0	100.00	100.00	1.65	1.65
50	10	.20, .23, .27, .30	1.0	92.65	92.10	2.10	2.00
50	50	.20, .23, .27, .30	1.1	100.00	100.00	1.65	0.55
50	30	.20, .23, .27, .30	1.1	100.00	100.00	1.75	0.80
50	10	.20, .23, .27, .30	1.1	96.55	94.85	2.25	1.00
30	50	.20, .23, .27, .30	.4	92.50	98.35	2.75	36.20
30	30	.20, .23, .27, .30	.4	74.90	91.25	1.65	34.70
30	10	.20, .23, .27, .30	.4	36.85	65.65	2.40	30.50
30	50	.20, .23, .27, .30	.6	98.65	99.55	2.30	18.00
30	30	.20, .23, .27, .30	.6	89.90	94.90	2.25	19.05
30	10	.20, .23, .27, .30	.6	48.00	66.15	2.40	16.45
30	50	.20, .23, .27, .30	.8	100.00	99.95	2.05	7.25
30	30	.20, .23, .27, .30	.8	97.35	98.75	2.10	7.85
30	10	.20, .23, .27, .30	.8	60.30	71.10	2.35	7.25
30	50	.20, .23, .27, .30	.9	99.99	99.90	2.30	4.25
30	30	.20, .23, .27, .30	.9	98.70	99.10	1.55	2.95
30	10	.20, .23, .27, .30	.9	68.25	73.20	2.25	4.40
30	50	.20, .23, .27, .30	1.0	100.00	100.00	2.65	2.55
30	30	.20, .23, .27, .30	1.0	99.45	99.60	2.60	2.45
30	10	.20, .23, .27, .30	1.0	75.50	75.45	1.80	1.75
30	50	.20, .23, .27, .30	1.1	100.00	100.00	2.05	0.95
30	30	.20, .23, .27, .30	1.1	99.90	99.75	2.05	0.80
30	10	.20, .23, .27, .30	1.1	83.55	76.75	1.40	0.60

3.10 Discussion

The assumption of independence underlies many statistical analyses. Analyzing correlated data under the assumption of independence is known to result in biased estimates of variance and incorrect test size. The assumption of independence between sightings of the same animal in habitat association studies has been questioned by Thomas and Taylor (1990) and Manly et al. (1993). Swihart and Slade (1985) discuss the effects of correlated data on estimates of home-range. The persistence model accounts for correlation between sightings within an animal and gives more accurate tests of habitat characteristic parameters than the IMS model when correlation exists.

Analyzing positively correlated persistence data under the independent multinomial selections model shows high percentages of lack of fit. The IMS model compensates for this lack of fit through the inclusion of erroneously significant habitat characteristic coefficients, as shown by the large sizes of the habitat characteristic tests. An opposite phenomenon is realized when analyzing negatively correlated persistence data under the IMS model. The bear data from Vander Heyden and Meslow (1999) demonstrate that persistence data does exist in habitat selection data and that different conclusions could be drawn if the data is analyzed assuming independence.

Testing the assumption of independence between sightings, $H_0 : \eta = 1$ vs $H_1 : \eta \neq 1$, shows that the score, Wald, and asymptotic likelihood ratio test perform equally well with respect to size. However, these tests do not perform similarly with respect to power. The Wald test has greater power for $\eta < 1$ and the score has greater power for $\eta > 1$. We recommend performing the score, Wald, and asymptotic likelihood ratio tests. If the results do not agree, use the Wald test for $\eta < 1$ and the score test for $\eta > 1$.

The power of the test of independence between sightings increases quickly as η moves away from one for all T and n_t . However, the power of the test of habitat

characteristics under the persistence model decreases as η decreases, especially with $n_t = 10$. Hence, though there might be enough power to detect persistence with ten sightings of each animal for low η , the power to detect habitat selection characteristics is low. The power to detect habitat selection characteristics is substantially greater for thirty sightings.

The persistence model was designed specifically to model the correlation that materializes through short time intervals between sightings. However, a similar correlation structure could emerge by sighting the animal at the same time of the day for a few consecutive days and then switching the time of the sightings to a different time. Correlation could also emerge through differential seasonal preferences by an animal.

We are presently working on estimates of the effective sample size for persistence data. The ability to estimate effective sample size would allow for a more straightforward characterization of power and for investigators to choose between different combinations of T and n_t to obtain the desired power.

CHAPTER 4

HETEROGENEITY MODEL

Individual animals having different multinomial selection probabilities results in population selection probabilities with larger variances than are accounted for by the multinomial model. Failure to account for this over-dispersion leads to estimates of variance that underestimate the true variance in the data. Manly et al. (1993) recognize the possibility of heterogeneity of habitat selection probabilities between animals, saying: “An interesting question would then be whether a model that allows each animal to have a different resource selection probability function gives a significantly better fit to the data than a model that assumes all animals have the same function.”

In this paper, we introduce a heterogeneity model, which is a one-parameter generalization of the independent multinomial selections (IMS) model to account for varying multinomial selection probabilities between animals. Its use yields more accurate tests of hypotheses than the IMS model when heterogeneity exists.

4.1 Hierarchical Dirichlet-Multinomial Distribution

Assume that each animal in the population chooses its multinomial habitat selection probabilities independently from a Dirichlet($\gamma, \boldsymbol{\pi}; \gamma > 0$) distribution, where $\boldsymbol{\pi} = [\pi_1, \dots, \pi_k]$ is the vector of population average multinomial habitat selection probabilities and γ is an over-dispersion parameter. Let the multinomial selection probabilities for animal t be given by: $\boldsymbol{\omega}_t = [\omega_{t1}, \dots, \omega_{tk}]$. This parameterization of the Dirichlet distribution has density function:

$$f(\boldsymbol{\omega}_t) = D(\boldsymbol{\pi}\boldsymbol{\gamma}^{-1})^{-1} \prod_{i=1}^K \omega_{ti}^{\pi_i \boldsymbol{\gamma}^{-1} - 1}$$

where

$$D(\boldsymbol{\pi}\boldsymbol{\gamma}^{-1}) = \int \cdots \int \prod_{i=1}^K \omega_{ti}^{\pi_i \boldsymbol{\gamma}^{-1} - 1} d\omega_{t1} \cdots d\omega_{tK} = \frac{\prod_{i=1}^K \Gamma(\pi_i \boldsymbol{\gamma}^{-1})}{\Gamma(\boldsymbol{\gamma}^{-1})} \quad (4.1)$$

Subsequently, the population habitat selection probabilities can be modeled using a hierarchical Dirichlet-multinomial distribution, giving the heterogeneity model.

Assuming independence between sightings within an animal reduces the sighting history, X_t , to: $Y_t = [Y_{t1}, \dots, Y_{tK}]$, where Y_{ti} is the number of times animal t was sighted in habitat i . The conditional distribution of Y_t given $\boldsymbol{\omega}_t$ is multinomial, $Y_t | \boldsymbol{\omega}_t \sim \text{multinomial}(n_t, \boldsymbol{\omega}_t)$, where n_t is the number of times animal t was sighted.

The unconditional distribution of Y_t (Appendix A.12) is:

$$f(y_t) = \binom{\boldsymbol{\gamma}^{-1} + n_t - 1}{n_t}_{\mathbb{R}} \prod_{i=1}^K \binom{\pi_i \boldsymbol{\gamma}^{-1} + y_{ti} - 1}{\pi_i \boldsymbol{\gamma}^{-1} - 1}_{\mathbb{R}}$$

where,

$$\binom{\boldsymbol{\gamma}^{-1} + n_t - 1}{n_t}_{\mathbb{R}} = \frac{(\boldsymbol{\gamma}^{-1} + n_t - 1) \cdots (\boldsymbol{\gamma}^{-1} + n_t - \lfloor \boldsymbol{\gamma}^{-1} \rfloor - n_t)}{n_t! (\boldsymbol{\gamma}^{-1} - 1) \cdots (\boldsymbol{\gamma}^{-1} - \lfloor \boldsymbol{\gamma}^{-1} \rfloor)}$$

and $\lfloor \gamma^{-1} \rfloor$ is defined as the integer part of γ^{-1} . The expected value and covariances of Y_{ti} are:

$$E[Y_{ti}] = n_t \pi_i \quad \text{Var}(Y_{ti}) = n_t \pi_i (1 - \pi_i) \frac{1 + n_t \gamma}{1 + \gamma}$$

$$\text{Cov}(Y_{ti}, Y_{tj}) = -n_t \pi_i \pi_j \frac{1 + n_t \gamma}{1 + \gamma}$$

(Appendix A.13). Hence, the expected number of sightings in habitat i for animal t is equal to that obtained under the IMS model. The variance and covariance, however, are a multiple of those obtained under the IMS model. As $\gamma \rightarrow \infty$ the factor, $\frac{1 + n_t \gamma}{1 + \gamma}$, goes to n_t , and as $\gamma \rightarrow 0$ the factor goes to one, which reduces the variance and covariance to that obtained under the IMS model. Hence, over-dispersion increases as γ increases. Moreover, testing $H_0 : \gamma = 0$ versus $H_1 : \gamma > 0$ tests the adequacy of the IMS model to the heterogeneity model.

4.2 Likelihood

Consider that there are T animals; let $\mathbf{Y} = [Y_1', \dots, Y_T']'$ be the matrix of sighting numbers for all T animals. Then assuming that the t^{th} animal is sighted on n_t occasions and that different animals contribute independent sightings, the likelihood for the heterogeneity model parameters is:

$$L(\boldsymbol{\pi}, \gamma | \mathbf{y}) = \prod_{t=1}^T \left[\binom{\gamma^{-1} + n_t - 1}{n_t}_{\mathbb{R}} \prod_{i=1}^K \binom{\pi_i \gamma^{-1} + y_{ti} - 1}{\pi_i \gamma^{-1} - 1}_{\mathbb{R}} \right] \quad (4.2)$$

and the loglikelihood, calculated in Appendix A.14, is:

$$\ell(\gamma, \boldsymbol{\pi} | \mathbf{y}) \propto - \sum_{t=1}^T \sum_{j=1}^{n_t} \ln(1 + [j-1]\gamma) + \sum_{t=1}^T \sum_{i=1}^K \sum_{j=1}^{y_{ti}} \ln(\pi_i + [j-1]\gamma) \quad (4.3)$$

4.3 Maximization

4.3.1 γ, π Parameterization

Maximum likelihood estimates (MLE's) of $\boldsymbol{\pi}$ and γ are obtained using the Newton-Raphson (N-R) algorithm. Initial estimates of $\hat{\boldsymbol{\pi}}$ are $\hat{\boldsymbol{\pi}}_0$, the maximum likelihood estimates under the IMS model, and the initial estimate of $\hat{\gamma}$ is (Appendix A.15):

$$\hat{\gamma} = \frac{Q-1}{n_* - Q}$$

where

$$Q = \frac{\sum_{t=1}^T (\hat{\pi}_{ti} - \hat{\pi}_{0i})^2 n_*}{(T-1) \hat{\pi}_{0i} (1 - \hat{\pi}_{0i})}$$

Here, $\hat{\pi}_{ti}$ is the MLE of π_i for animal t under IMS model assumptions and $\hat{\pi}_{0i}$ is MLE of π_i overall animals under the IMS model; $i = 1, \dots, K$. n_* is a measure of central n_t , $t = 1, \dots, T$; we use $n_* = \text{mean}(n_t)$.

4.3.2 γ, β Parameterization

Substituting the multinomial logit equation (2.3) into equation (4.3) gives the loglikelihood in terms of $\beta = [\beta_0, \underline{\beta}']'$, log area, and the habitat characteristics Z . Again, the N-R algorithm is used to obtain the MLE's for β and γ . Initial estimates of $\hat{\beta}$ are obtained using the same algorithm discussed under the IMS model. The initial estimate of $\hat{\gamma}$ is calculated, as described above, after projecting $\hat{\pi}_0$ onto Z -space.

The Newton-Raphson algorithm is based on the maximum being in the interior of the parameter space. However, it is possible that $\hat{\gamma} = 0$. Therefore, to circumvent convergence difficulties in using the N-R algorithm, the first derivative of the loglikelihood with respect to γ is computed at $\gamma = 0$ and $\hat{\pi}_0$. If the first derivative is less than 0, showing that the likelihood is decreasing as γ increases, then $\hat{\gamma}$ is set to zero and $\hat{\pi}$ is set to $\hat{\pi}_0$ (Appendix A.16).

The first and second derivatives of the loglikelihood, $\ell(\gamma, \beta | \mathbf{y})$ are given in Appendix A.17. Approximate standard errors of $\hat{\beta}$ and $\hat{\gamma}$ are obtained from the inverse of the observed information matrix, and approximate standard errors of the habitat selection probabilities, $\hat{\pi}$, and selection probability ratios, $\hat{\pi}_i / \hat{\pi}_k$, are obtained using the delta method (Appendix A.6).

4.4 Tests

4.4.1 Independent Multinomial Selections Model

Tests of the null hypothesis of homogeneity versus the alternative hypothesis of heterogeneity are conducted by testing $H_0 : \gamma = 0$ vs $H_1 : \gamma > 0$. The score, Wald, and asymptotic likelihood ratio chi-squared statistics were constructed to test for

independence (Appendix A.18). These tests are calculated under the full (saturated) model, where there are $K-1$ parameters.

The score test is valid when the MLE, $\hat{\gamma}$, is in the interior of the parameter space, $\gamma > 0$. When the MLE is on the boundary, $\hat{\gamma} = 0$, the score statistic could be extremely negative; thereby, rejecting $H_0 : \gamma = 0$, even though $\hat{\gamma} = 0$. $H_0 : \gamma = 0$ should not be rejected when $\hat{\gamma} = 0$, regardless of the test outcome.

4.4.2 Habitat Characteristics

The test of no effect of habitat characteristic j , $H_0 : \beta_j = 0$ vs $H_1 : \beta_j \neq 0$ is completed through the usual Type III drop-in-deviance chi-squared test statistic. The test of the coefficient for log area equal to one, $H_0 : \beta_0 = 1$, vs $H_1 : \beta_0 \neq 1$ is also constructed. This test determines if log area can be offset – thereby accounting for random selection of habitat classes giving selection probabilities proportional to area. Rejecting $H_0 : \beta_0 = 1$ could be the result of missing a necessary habitat characteristic in the model, or it could be the result of some other random selection procedure by the animal, such as randomly picking an initial point and walking a straight line in some direction.

Consider the case where the habitat characteristics are indicator variables for the habitat classes, habitat class K is the reference class, and area is offset. Here, the test of $H_0 : \beta_j = 0$ tests if the log selection probability ratio (SPR) of habitat j to habitat K is equal to the ratio of the areas of habitat j to habitat K . Positive β_j gives a SPR of habitat j to habitat K that is greater than expected given the respective areas. Negative β_j gives a SPR of habitat j to habitat K that is less than expected given the respective areas.

4.5 Example

As part of a study to determine home range and habitat selection by female black bears (*Ursus americanus*) in central Oregon Cascades, female black bears were captured and radiocollared in the springs of 1993-1995 (Vander Heyden and Meslow, 1999). The bears were then radio-tracked, on average, every other day, with attempts to sight at three different times of the day (morning, afternoon, and evening) (Vander Heyden and Meslow, 1999). The study area was divided into six habitat classes: grass-forb, shrub, open-canopy sapling/pole, closed-canopy sapling/pole, closed-canopy mature timber, and open-canopy mature timber (exact definitions are given in Vander Heyden and Meslow, 1999).

Throughout 1993 to 1995 a total of twelve bears were collared and sighted. Home ranges were determined each year for each bear using adaptive kernel and minimum convex polygon algorithms (Vander Heyden and Meslow, 1999). Within these home ranges random samples of locations were made to determine availability of habitat. The size, in hectares, of the habitat patch surrounding each of the random locations was determined using a LANDSAT Thematic Mapper image (Vander Heyden and Meslow, 1999).

The 1995 data, analyzed herein, contained between 38 and 69 sightings on eight different female black bears. The analyses use the proportion of each habitat in the entire study area to adjust for availability. The proportion of study area (Vander Heyden and Meslow, 1999) and area availability ratio with respect to open-canopy mature timber habitat are given for each of the six habitat classes in Table 4.1.

Table 4.1 Study Area: proportion of area available and availability ratio

Habitat Class	Grass Forb	Shrub	Open Sapling	Closed Sapling	Closed Mature	Open Mature
Proportion Available	0.17	0.10	0.10	0.13	0.32	0.18
Availability Ratio	0.9444	0.5556	0.5556	0.7222	1.7778	1.0000

Habitat selections for this year were analyzed employing log area and indicator variables for grass-forb, shrub, sapling/pole, and closed-canopy as explanatory variables. Both the heterogeneity and independent multinomial selections models were fit to the data using open-canopy mature timber as the reference habitat to calculate the selection probability ratios (SPR).

4.5.1 *Independent Multinomial Selection Analysis*

The IMS model shows that the coefficient for log area, 1.4256 (0.6988), is not significantly different from one with drop-in-deviance 0.3723 on 1 df, giving p-value = 0.5418. Therefore, area is offset. The indicator for sapling/pole is then removed from the model with drop-in-deviance statistic 0.0044 on 1 df, giving p-value = 0.9471. The resulting model shows that the SPR's of grass-forbs, shrubs, and closed-canopy are significantly smaller than their respective available area ratios with 1 df drop-in-deviance p-values: <0.0001, <0.0001, and 0.0087 respectively. Table 4.2 gives the estimated selection probabilities and SPR's with standard errors for each habitat.

Table 4.2 Independent multinomial selections model estimates of habitat selection probabilities and selection probability ratios

IMS Habitat Class	Grass Forb	Shrubs	Open Sapling	Closed Sapling	Closed Mature	Open Mature
Prob. of Use	0.0969	0.0463	0.1372	0.1366	0.3362	0.2470
SE	0.0140	0.0100	0.0082	0.0068	0.0168	0.0148
SPR	0.3924	0.1873	0.5556	0.5530	1.3612	1.0000
SE	0.0668	0.0437	-	0.0569	0.1401	

Comparing Table 4.2 with Table 4.1 shows the SPR for grass-forb, shrubs, closed-canopy sapling/pole, and closed-canopy mature timber are smaller than their respective availability ratios. The specified model forces the open-canopy sapling/pole SPR to be identical to its availability ratio and hence has no standard error.

4.5.2 Heterogeneity Analysis

The heterogeneity model estimated γ as 0.0746 with standard error: 0.0223. The asymptotic likelihood ratio statistic for testing $H_0: \gamma = 0$ (IMS model) vs $H_0: \gamma \neq 0$ (heterogeneity model) was 73.5209 with 1 df, for a p-value < 0.0001. Therefore, there is substantial evidence of heterogeneity. This model shows that the coefficient for log area, 1.5937 (1.5076), is not significantly different from one with drop-in-deviance statistic 0.1559 on 1 df, giving p-value = 0.6930. Therefore, area is offset. The indicator for sapling/pole is removed from the model with drop-in-deviance statistic 0.0086 on 1 df, giving p-value = 0.9261. The resulting model shows that the SPR's of grass-forbs, and shrubs are significantly smaller than their respective availability ratios with 1 df drop-in-deviance p-values: 0.0483 and 0.0378 respectively. The SPR's for habitats classified as closed canopy are not significantly different from their availability ratios, 1 df chi-squared p-value = 0.5976. Table 4.3 gives the estimated selection probabilities and SPR's with standard errors for each habitat using the model including the closed canopy term in order to make comparisons with the IMS model results.

Table 4.3 Heterogeneity model estimates of habitat selection probabilities and selection probability ratios

Heterogeneity Habitat Class	Grass Forb	Shrubs	Open Sapling	Closed Sapling	Closed Mature	Open Mature
Prob. of Use	0.1102	0.0545	0.1229	0.1419	0.3492	0.2213
SE	0.0302	0.0206	0.0174	0.0147	0.0363	0.0314
SPR	0.4983	0.2463	0.5556	0.6413	1.5785	1.0000
SE	0.1686	0.1050	-	0.1473	0.3626	

Comparing Table 4.3 with Table 4.2 shows the probabilities of use and SPR's are similar between the IMS and heterogeneity models. However, the standard errors for the probabilities of use and SPR's in the heterogeneity model are more than twice as large as those obtained from the IMS model. This increase in standard error is due to heterogeneity model's ability to account for extra multinomial variation.

Proceeding with the heterogeneity model, the non-significant closed-canopy term was removed from the model, leaving significant indicators for grass-forb and shrub, 1 df chi-squared p-values: 0.0482 and 0.0405 respectively. The estimated SPR of grass-forbs is $e^{-0.5661} = 0.5678$ times the area ratio with 90% Wald CI = (0.3461, 0.9314) and the estimated SPR of shrub is $e^{-0.7394} = 0.4774$ times the area ratio with 90% Wald CI = (0.2500, 0.9117). Table 4.4 gives the estimated selection probabilities and SPR's with approximated standard errors for each habitat.

Table 4.4 Reduced heterogeneity model estimates of habitat selection probabilities and selection probability ratios

Heterogeneity Habitat Class	Grass Forb	Shrubs	Open Sapling	Closed Sapling	Closed Mature	Open Mature
Prob. of Use	0.1104	0.0546	0.1144	0.1487	0.3660	0.2059
SE	0.0293	0.0202	0.0048	0.0062	0.0153	0.0086
SPR	0.5362	0.2652	0.5556	0.7222	1.7778	1.0000
SE	0.1614	0.1043	-	-	-	

The estimates of selection probability ratios for open-canopy sapling/pole, closed-canopy sapling/pole, and closed-canopy mature timber are forced to be identical to their respective availability ratios due to the explanatory variables remaining in the model. Therefore, there are no model based estimates of standard errors for these estimates. However, there are model based estimates of standard errors for the estimates of probability of use for all habitats. Note that the estimates of probability of use in this reduced model are very similar to those in the previous model, Table 4.3. The estimated standard errors for grass-forbs and shrubs are also similar to the previous model, but the estimated SE's for the remaining habitat types are substantially smaller.

Analyzing this habitat selection data assuming identical selection probabilities for all bears (IMS model) results in estimates of selection probability ratios for habitat types with closed-canopy that are significantly smaller than their availability ratios. Allowing for heterogeneous selection probabilities results in estimates of selection probability ratios for habitat types with closed-canopy that are not significantly different from their availability ratios. Hence, analysis results depend on the assumption of identical selection probabilities between animals; therefore, this assumption should be checked.

4.6 Simulation Results

4.6.1 Test Size for IMS vs Heterogeneity

The sizes of the score, Wald, and asymptotic LR tests for homogeneity of selection probabilities, $H_0 : \gamma = 0$ vs $H_0 : \gamma > 0$, were calculated within a saturated model using 2000 simulations with level $\alpha = 0.10, 0.05$, and 0.01 tests. Table 4.5 shows results for level $\alpha = 0.05$ tests with $K = 4$ habitats and $\pi = [.25, .25, .25, .25]$. Appendix A.19 gives further results for $K = 4$ habitats.

Table 4.5 Size of $H_0 : \gamma = 0$ vs $H_1 : \gamma > 0$

T	n_t	π	SS	Size x 100		
				score	Wald	LRT
50	50	.25, .25, .25, .25	2000	0.55	4.05	1.45
50	30	.25, .25, .25, .25	2000	0.80	3.60	1.65
50	10	.25, .25, .25, .25	2000	0.50	4.20	1.60
30	50	.25, .25, .25, .25	2000	0.40	4.05	1.30
30	30	.25, .25, .25, .25	2000	0.55	3.75	1.60
30	10	.25, .25, .25, .25	2000	0.15	4.10	1.35
10	50	.25, .25, .25, .25	2000	0.15	4.50	0.80
10	30	.25, .25, .25, .25	2000	0.15	4.40	1.00
10	10	.25, .25, .25, .25	2000	0.30	4.25	0.70

The score and LRT test are both very conservative, rejecting H_0 about 0.5 and 1.5 percent of the time respectively using level $\alpha = 0.05$ tests. This conservative testing using score and LRT is also seen with the level $\alpha = 0.10$ and $\alpha = 0.01$ tests. The Wald test works fairly well for level $\alpha = 0.05$ tests, rejecting H_0 about 4.25 percent of the time. However, this test is more conservative for level $\alpha = 0.10$ tests and liberal for level $\alpha = 0.01$ rejecting approximately 6.5 and 2.0 percent of the time respectively.

4.6.2 Test Power for IMS vs Heterogeneity

The power of the score, Wald, and asymptotic LR tests for $H_0 : \gamma = 0$ vs $H_0 : \gamma > 0$ were calculated within a saturated model using 1000 simulations with level $\alpha = 0.05$ tests. Power was calculated for various γ . Table 4.6 shows results for $K = 4$ habitats, $\pi = [.25, .25, .25, .25]$, and $\gamma = 0.25$ & 0.05 . Further results are given in Appendix A.20. These tables show that power to detect a difference from $\gamma = 0$ increases as γ increases. At $\gamma = 0.25$, the power is 1.00 for all cases except $T = n_t = 10$. Power also increases with increasing T and n_t . The Wald test gives the highest

power for all combinations of T , n_t , and γ . For example, at $\gamma = 0.05$, $T = 10$, and $n_t = 30$ the power of the score, Wald, and asymptotic LR tests are: 67.90, 93.30, and 83.70 respectively. The power of these tests do not appear to depend on habitat selection probabilities used to generate the data; hence, results for only one π per K are shown.

Table 4.6 Power of $H_0 : \gamma = 0$ vs $H_1 : \gamma > 0$

T	n_t	π	γ	Power x 100		
				Score	Wald	LRT
50	50	.25, .25, .25, .25	1/4	100.00	100.00	100.00
50	30	.25, .25, .25, .25	1/4	100.00	100.00	100.00
50	10	.25, .25, .25, .25	1/4	100.00	100.00	100.00
30	50	.25, .25, .25, .25	1/4	100.00	100.00	100.00
30	30	.25, .25, .25, .25	1/4	100.00	100.00	100.00
30	10	.25, .25, .25, .25	1/4	100.00	100.00	100.00
10	50	.25, .25, .25, .25	1/4	100.00	100.00	100.00
10	30	.25, .25, .25, .25	1/4	100.00	100.00	100.00
10	10	.25, .25, .25, .25	1/4	91.50	99.20	96.50
50	50	.25, .25, .25, .25	1/20	100.00	100.00	100.00
50	30	.25, .25, .25, .25	1/20	100.00	100.00	100.00
50	10	.25, .25, .25, .25	1/20	78.30	92.60	86.20
30	50	.25, .25, .25, .25	1/20	100.00	100.00	100.00
30	30	.25, .25, .25, .25	1/20	100.00	100.00	100.00
30	10	.25, .25, .25, .25	1/20	50.50	81.50	66.30
10	50	.25, .25, .25, .25	1/20	95.80	99.50	98.20
10	30	.25, .25, .25, .25	1/20	67.90	93.30	83.70
10	10	.25, .25, .25, .25	1/20	6.30	41.60	20.70

4.6.3 Habitat Characteristic Test Size

The size of the asymptotic likelihood ratio (drop-in-deviance) test for habitat characteristics, $H_0 : \beta_1 = 0$ vs $H_1 : \beta_1 \neq 0$, was calculated using 2000 simulations with level $\alpha = 0.05$ tests. Results are given in Table 4.7 for $K = 4$, $T = 30$, and various n_t and γ . The entry under Z is the indicator variable tested (all indicator variables gave

similar results, only one is presented); “Het.” entries are 100 times the size of the test under the heterogeneity model; and “IMS” entries are 100 times the size of the test under the IMS model. At $\gamma = 1/8$, the heterogeneity model shows a size of about 0.05. However, the IMS model shows liberal tests of significance with sizes between 0.16 and 0.45, increasing with n_t .

Table 4.7 Size of $H_0 : \beta_1 = 0$ vs $H_1 : \beta_1 \neq 0$

T	n_t	π	γ	Z	Size x 100	
					Het.	IMS
30	50	.25, .25, .25, .25	1/8	1, 0, 0, 0	5.54	44.54
30	30	.25, .25, .25, .25	1/8	1, 0, 0, 0	4.96	35.90
30	10	.25, .25, .25, .25	1/8	1, 0, 0, 0	5.53	16.33
30	50	.25, .25, .25, .25	1/20	1, 0, 0, 0	4.40	28.05
30	30	.25, .25, .25, .25	1/20	1, 0, 0, 0	5.21	20.77
30	10	.25, .25, .25, .25	1/20	1, 0, 0, 0	4.26	8.47
30	50	.25, .25, .25, .25	1/40	1, 0, 0, 0	4.75	18.75
30	30	.25, .25, .25, .25	1/40	1, 0, 0, 0	4.45	12.55
30	10	.25, .25, .25, .25	1/40	1, 0, 0, 0	5.42	7.60
30	50	.25, .25, .25, .25	1/100	1, 0, 0, 0	5.55	10.95
30	30	.25, .25, .25, .25	1/100	1, 0, 0, 0	5.25	8.65
30	10	.25, .25, .25, .25	1/100	1, 0, 0, 0	4.90	6.36

As γ decreases toward zero, the size under the IMS model decreases to the level of the test. This reduction in the size is expected since at $\gamma = 0$, the heterogeneity model is equivalent to the IMS model. The size obtained under the heterogeneity model does not appear to depend on K , γ , T , or n_t .

4.6.4 *Habitat Characteristic Test Power and Model Lack of Fit (LOF)*

The power of the asymptotic likelihood ratio (drop-in-deviance) test for habitat characteristics, $H_0 : \beta_1 = 0$ vs $H_1 : \beta_1 \neq 0$, was calculated using 2000 simulations with a level $\alpha = 0.05$ test. Sighting history matrices were generated using various non-uniform selection probabilities. The logs of the selection probabilities were used as habitat characteristic variables in the models. Model lack of fit was tested between the fitted model with one explanatory variable and the saturated model with $K-1$ explanatory variables. The percent of model lack of fit using a level $\alpha = 0.05$ test was recorded.

Table 4.8 gives results for selection probabilities, $\pi = [.278 \ .333 \ .389]$. This table shows that the power to detect a difference from uniform habitat selection probabilities decreases as γ increases and T and n_t decrease. While these results hold for both models, there is a greater effect seen under the heterogeneity model. The percent of LOF is between 0.0 and 0.5 under the heterogeneity model and does not appear to depend on γ , T , or n_t . Under the IMS model, the percent of LOF is between 0.5 to 37.7, increasing with γ and n_t . There does not appear to be effect of T on LOF under the IMS model.

Table 4.8 Power and lack of fit of $H_0 : \beta_1 = 0$ vs $H_1 : \beta_1 \neq 0$

T	n_t	π	γ	Power x 100		LOF x 100	
				Het.	IMS	Het.	IMS
50	50	.278, .333, .389	10/36	56.75	92.55	0.20	37.65
50	30	.278, .333, .389	10/36	54.75	89.85	0.25	25.75
50	10	.278, .333, .389	10/36	46.00	75.75	0.15	7.95
50	50	.278, .333, .389	1/18	95.10	99.70	0.00	10.80
50	30	.278, .333, .389	1/18	91.55	97.80	0.15	5.15
50	10	.278, .333, .389	1/18	70.30	82.25	0.15	1.40
50	50	.278, .333, .389	1/36	99.35	99.90	0.40	5.90
50	30	.278, .333, .389	1/36	97.15	99.25	0.30	2.35
50	10	.278, .333, .389	1/36	77.35	83.95	0.30	0.50
30	50	.278, .333, .389	10/36	38.35	86.60	0.10	35.95
30	30	.278, .333, .389	10/36	35.95	79.05	0.25	26.10
30	10	.278, .333, .389	10/36	31.20	61.20	0.15	8.00
30	50	.278, .333, .389	1/18	81.10	96.65	0.30	10.90
30	30	.278, .333, .389	1/18	76.10	91.80	0.45	4.80
30	10	.278, .333, .389	1/18	50.55	62.55	0.20	1.20
30	50	.278, .333, .389	1/36	93.85	98.80	0.35	4.05
30	30	.278, .333, .389	1/36	86.05	94.05	0.15	1.20
30	10	.278, .333, .389	1/36	55.60	63.60	0.35	0.65

Similar effects on power and LOF are seen for $K = 4$ and 5 habitats, given in Appendix A.21. These tables show similar results as Table 4.8 above. Moreover, these tables show that power and LOF, under both the heterogeneity and IMS models, increase as the number of habitats, K , increases.

4.7 Discussion

The question of heterogeneity or over-dispersion arises often in statistical analyses, specifically in generalized linear models. Analyzing heterogeneous data under the assumption of homogeneity is known to result in biased estimates of variance and incorrect test size. The assumption of identical multinomial habitat

selection probabilities between animals has been questioned by Manly et al. (1993). The heterogeneity model accounts for varying habitat selection probabilities between animals and gives more accurate tests of habitat characteristic parameters than the IMS model when heterogeneity.

Analyzing heterogeneous data under the independent multinomial selections model shows high percentages of lack of fit. The IMS model compensates for this lack of fit through the inclusion of erroneously significant habitat characteristic coefficients, as shown by the large sizes of the habitat characteristic tests. The bear data from Vander Heyden and Meslow (1999) demonstrate that heterogeneous data does exist in habitat selection data and that different conclusions could be drawn if the data is analyzed assuming independence.

Testing the assumption of homogeneity, $H_0 : \gamma = 0$ vs $H_1 : \gamma > 0$, shows that the score and asymptotic likelihood ratio test are extremely conservative. The Wald test, while being somewhat conservative attains a size that is closest to the expected level, with a size of about 4.25 under a level $\alpha = 0.05$ test. Moreover, the Wald test consistently attains the highest power of the three tests followed by the asymptotic LR and score test. We recommend performing the score, Wald, and asymptotic likelihood ratio tests, noting that the Wald test performs best for both size and power.

The power of the test of independence between sightings increases quickly as γ increases from zero for all T and n_t . However, the power of the test of habitat characteristics under the persistence model decreases as γ increases. Hence, though there might be enough power to detect persistence with ten sightings of each animal for low γ , the power to detect habitat selection characteristics is low. The power to detect habitat selection characteristics increases substantially with increased sightings when γ is small (little heterogeneity). However, when γ is not small (greater heterogeneity), there is more of a benefit to increase the number of animals in the study than to increase the number of sightings of a given animal.

The heterogeneity model was designed specifically to model heterogeneity habitat selection probabilities between animals. Therefore, while this model has the ability to account for over-dispersion, it does not have the ability to handle under-

dispersion. Were the habitat selection probabilities under-dispersed, the heterogeneity model would estimate $\gamma = 0$, and therefore, revert back to the independent multinomial selections model.

BIBLIOGRAPHY

- Johnson, D.H. (1980). The comparison of usage and availability measurements for evaluating resource preference. *Ecology* **61**: 65-71.
- Manly, B.F.J., McDonald, L.L., and Thomas, D.L. (1993). *Resource selection by animals: statistical design and analysis for field studies*. Chapman and Hall, London.
- Mate, B.R., Nieuwirth, S.L., and Kraus, S.D. (1997). Satellite-monitored movements of the northern right whale. *Journal of Wildlife Management* **61**(4):1393-1405.
- Mate, B.R., Gisiner, R., and Mobley, J. (1998). Local and migratory movements of Hawaiian humpback whales tracked by satellite telemetry. *Canadian Journal of Zoology* **76**:863-868.
- McCracken, M.L., Manly, B.F., and Vander Heyden, M. (1998). The use of discrete-choice models for evaluating resource selection. *Journal of Agricultural, Biological and Environmental Statistics* **3**(3):268-279.
- Ramsey, F.L., McCracken, M., Crawford, J.A., Drut, M.S., and Ripple, W.J. (1994). Habitat Association Studies of the Northern Spotted Owl, Sage Grouse, and Flammulated Owl. *Case Studies in Biometry*. Eds: N. Lange, L. Ryan, L. Billard, D. Brillinger, L. Conquest, and J. Greenhouse. John Wiley and Sons, Inc., New York, pages 189-209.
- Swihart, R.K. and Slade, N.A. (1985a). Influence of sampling interval on estimates of home-range size. *Journal of Wildlife Management* **49**(4):1019-1025
- Swihart, R.K. and Slade, N.A. (1985b). Testing for independence of observations animal movements. *Ecology* **66**(4):1176-1184
- Thomas, D.L. and Taylor, E.J. (1990). Study Designs and tests for comparing resource use and availability. *Journal of Wildlife Management* **54**(2):322-330.
- Vander Heyden, M. and Meslow, C.E. (1999). Habitat selection by female black bears in the central Cascades of Oregon. *Northwest Science* **73**(4): 283-294.

APPENDICES

APPENDIX A.1

LOGISTIC REGRESSION PARAMETER BIAS

Consider a study area with two habitat classes and the use versus habitat contingency table given below.

		Population Habitat Class		Total
		A	B	
Used	Y	$N_U - Y$	N_U	
Unused	N_{NA}	N_{NB}	N_N	
Available	N_A	N_B	N	

N_U sites are known to be used, with Y being the number in habitat A. Let λ be the probability of a site being in habitat A given the site was used. Further, consider a simple random sample of size M from sites known to be unused, with X being the number of sampled sites in habitat A. Let μ be the probability of a site being in habitat A given the site is unused. Then the joint distribution of Y and X is:

$$P(Y, X) = \binom{N_U}{y} \lambda^y (1-\lambda)^{N_U-y} \binom{M}{x} \mu^x (1-\mu)^{M-x}, \text{ with maximum likelihood estimates:}$$

$$\hat{\lambda} = \frac{y}{N_U} \text{ and } \hat{\mu} = \frac{x}{M}.$$

Let β denote the log-odds ratio of use for habitat A versus B. Then

$$\beta = \frac{P(\text{Used} | A)/P(\text{Unused} | A)}{P(\text{Used} | B)/P(\text{Unused} | B)} \text{ and, via Bayes theorem, } \hat{\beta} = \ln \left(\frac{\hat{\lambda}(1-\hat{\mu})}{(1-\hat{\lambda})\hat{\mu}} \right).$$

Now consider that the randomly sampled sites are from those available, not just those unused. Let H be the number of sites sampled, Z be the number in habitat A, and θ be the overall probability of a site being in habitat A. Then, the joint distribution of

$$Y \text{ and } Z \text{ is: } P(Y, Z) = \binom{N_U}{y} \lambda^y (1-\lambda)^{N_U-y} \binom{H}{z} \theta^z (1-\theta)^{H-z}, \text{ with:}$$

$$\hat{\lambda} = \frac{y}{N_U} \text{ and } \hat{\theta} = \frac{z}{H} \text{ and estimated log-odds ratio } \hat{\beta}_b = \ln \left(\frac{\hat{\lambda}/\hat{\theta}}{(1-\hat{\lambda})/(1-\hat{\theta})} \right).$$

However, θ does not measure the probability of being sighted in habitat A for unused sites. Instead, θ measures the overall probability of being in habitat A. Therefore,

$$\hat{\beta}_b = \hat{\beta} + \ln \left(\frac{\hat{\mu}/(1-\hat{\mu})}{\hat{\theta}/(1-\hat{\theta})} \right)$$

does not estimate the log-odds ratio of usage for habitat A versus habitat B.

Consider that θ can be rewritten in terms of λ and μ defined above:

$$\theta = \lambda \left(\frac{N_U}{N} \right) + \mu \left(1 - \frac{N_U}{N} \right), \text{ where } N \text{ is the total number of habitats. Then}$$

$$\hat{\beta}_b = \hat{\beta} + \ln \left(\frac{\hat{\mu}}{1-\hat{\mu}} \frac{1 - \left[\hat{\lambda} \left(\frac{N_U}{N} \right) + \hat{\mu} \left(1 - \frac{N_U}{N} \right) \right]}{\hat{\lambda} \left(\frac{N_U}{N} \right) + \hat{\mu} \left(1 - \frac{N_U}{N} \right)} \right).$$

From this equation it can be seen that as the proportion of used sites, $\frac{N_U}{N}$, goes to 0,

the estimate of $\hat{\beta}_b \rightarrow \hat{\beta}$. However, as $\frac{N_U}{N}$ goes to 1,

$$\hat{\beta}_b \rightarrow \hat{\beta} + \ln \left(\frac{\hat{\mu}}{1-\hat{\mu}} \frac{1-\hat{\lambda}}{\hat{\lambda}} \right) = \hat{\beta} - \ln \left(\frac{\hat{\lambda}}{1-\hat{\lambda}} \frac{1-\hat{\mu}}{\hat{\mu}} \right) = 0.$$

APPENDIX A.2**PERSISTENCE PARAMETER, η , BOUNDARIES**

Consider that each entry in $\Pi(\eta)$ must be between 0 and 1 since $\Pi(\eta)$ is a probability transition matrix. Hence,

$$1) \quad 1 - \eta(1 - \pi_i) \geq 0 \quad \forall i \Leftrightarrow \eta \leq \frac{1}{1 - \pi_i} \quad \forall i$$

$$2) \quad \eta\pi_i \leq 1 \quad \forall i \Leftrightarrow \eta \leq \frac{1}{\pi_i} \quad \forall i$$

and
$$3) \quad \eta\pi_i \geq 0 \quad \forall i \Leftrightarrow \eta \geq 0 \quad \forall i$$

APPENDIX A.3

CORRELATION BETWEEN CONSECUTIVE SIGHTINGS

Let W be the habitat the animal was sighted in at time t .

Let U be the habitat the animal was sighted in at time $t+1$.

Calculate variances of W and U .

$$E[W] = E[U] = \sum_{i=1}^K i\pi_i$$

$$E[W^2] = E[U^2] = \sum_{i=1}^K i^2\pi_i$$

Hence,

$$\text{Var}(W) = \text{Var}(U) = \sum_{i=1}^K i^2\pi_i(1-\pi_i) - 2\sum_{i<j} ij\pi_i\pi_j,$$

and $\sqrt{\text{Var}(W)\text{Var}(U)} = \text{Var}(W) = \text{Var}(U)$.

Calculate covariance of W and U .

$$\begin{aligned} E[WU] &= \sum_{i=1}^K i^2\pi_i(1-\eta(1-\pi_i)) + 2\sum_{i<j} ij\eta\pi_i\pi_j \\ &= \sum_{i=1}^K [i^2\pi_i - i^2\eta\pi_i + i^2\eta\pi_i^2] + 2\sum_{i<j} ij\eta\pi_i\pi_j \end{aligned}$$

where the first term results from the animal being sighted in the same habitat at time t and $t+1$, and the second term results from the animal being sighted in different habitats. Hence,

$$\begin{aligned}
 \text{Cov}(W, U) &= \sum_{i=1}^K \left[i^2 \pi_i - i^2 \eta \pi_i + i^2 \eta \pi_i^2 \right] + 2 \sum_{i < j}^K ij \eta \pi_i \pi_j - \sum_{i=1}^K i^2 \pi_i^2 - 2 \sum_{i < j}^K ij \pi_i \pi_j \\
 &= \sum_{i=1}^K \left[(1-\eta) i^2 \pi_i - (1-\eta) i^2 \pi_i^2 \right] - 2(1-\eta) \sum_{i < j}^K ij \pi_i \pi_j \\
 &= (1-\eta) \left\{ \sum_{i=1}^K i^2 \pi_i (1-\pi_i) - 2 \sum_{i < j}^K ij \pi_i \pi_j \right\}
 \end{aligned}$$

$$\text{Corr}(W, U) = 1 - \eta$$

APPENDIX A.4

PROOF OF FUNCTIONAL EQUALITY BETWEEN $\hat{\eta}$ AND π

Consider the loglikelihood for the persistence model parameters given the sighting history matrix:

$$\begin{aligned} \ell(\boldsymbol{\pi}, \eta | \mathbf{X}) = & \sum_{i=2}^K (f_{\cdot i} + v_{\cdot i}) \ln \pi_i + (f_{\cdot 1} + v_{\cdot 1}) \ln \pi_1 + V \cdot \ln \eta + \\ & \sum_{i=2}^K n_{\cdot ii} \ln(1 - \eta(1 - \pi_i)) + n_{\cdot 11} \ln(1 - \eta(1 - \pi_1)) \end{aligned}$$

Show

$$\hat{\eta} = \frac{B}{(K-1)n_{\cdot 11} + B(1 - \hat{\pi}_1)} \quad (\text{I.1})$$

where

$$B = \sum_{i=2}^K (f_{\cdot i} + v_{\cdot i}) \left(\frac{1 - \hat{\pi}_i}{\hat{\pi}_i} \right) - (f_{\cdot 1} + v_{\cdot 1}) \left(\frac{(K-2) + \hat{\pi}_1}{\hat{\pi}_1} \right) + V. \quad (\text{I.2})$$

First derivatives of $\ell(\mathbf{X} | \boldsymbol{\pi}, \eta)$ for π_j and η .

$$\frac{\partial \ell}{\partial \pi_j} = \frac{f_{\cdot j} + v_{\cdot j}}{\pi_j} - \frac{f_{\cdot 1} + v_{\cdot 1}}{\pi_1} + \frac{\eta n_{\cdot jj}}{1 - \eta(1 - \pi_j)} - \frac{\eta n_{\cdot 11}}{1 - \eta(1 - \pi_1)} \quad \forall j \neq 1$$

$$\frac{\partial \ell}{\partial \eta} = \frac{V}{\eta} - \sum_{i=2}^K \frac{n_{\cdot ii}(1 - \pi_i)}{1 - \eta(1 - \pi_i)} - \frac{n_{\cdot 11}(1 - \pi_1)}{1 - \eta(1 - \pi_1)}$$

Solve $\frac{\partial \ell}{\partial \eta} \equiv 0$ for $\frac{\eta n_{.22}}{1 - \eta(1 - \pi_2)}$

$$\begin{aligned} \frac{\partial \ell}{\partial \eta} = 0 &\Leftrightarrow \frac{V_{.}}{\eta} - \sum_{i=2}^K \frac{n_{.ii}(1 - \pi_i)}{1 - \eta(1 - \pi_i)} - \frac{n_{.11}(1 - \pi_1)}{1 - \eta(1 - \pi_1)} = 0 \\ &\Leftrightarrow \frac{n_{.22}(1 - \pi_2)}{1 - \eta(1 - \pi_2)} = \frac{V_{.}}{\eta} - \sum_{i=3}^K \frac{n_{.ii}(1 - \pi_i)}{1 - \eta(1 - \pi_i)} - \frac{n_{.11}(1 - \pi_1)}{1 - \eta(1 - \pi_1)} \\ &\Leftrightarrow \frac{\eta n_{.22}}{1 - \eta(1 - \pi_2)} = \frac{V_{.}}{(1 - \pi_2)} - \sum_{i=3}^K \frac{\eta n_{.ii}}{1 - \eta(1 - \pi_i)} \left(\frac{(1 - \pi_i)}{(1 - \pi_2)} \right) - \frac{\eta n_{.11}}{1 - \eta(1 - \pi_1)} \left(\frac{(1 - \pi_1)}{(1 - \pi_2)} \right) \end{aligned}$$

Now use mathematical induction to show that

$$\frac{\eta n_{.qq}}{1 - \eta(1 - \pi_q)} = \frac{V_{.}}{(1 - \pi_q)} + \sum_{i=2}^{q-1} \frac{f_{.i} + v_{.i}}{1 - \pi_q} \left(\frac{1 - \pi_i}{\pi_i} \right) - \frac{f_{.1} + v_{.1}}{1 - \pi_q} \left(\frac{(q-2) - \sum_{i=2}^{q-1} \pi_i}{\pi_1} \right) - \sum_{i=q+1}^K \frac{\eta n_{.ii}}{1 - \eta(1 - \pi_i)} \left(\frac{(1 - \pi_i)}{(1 - \pi_q)} \right) - \frac{\eta n_{.11}}{1 - \eta(1 - \pi_1)} \left(\frac{(q-1) - \sum_{i=1}^{q-1} \pi_i}{1 - \pi_q} \right) \quad (I.3)$$

1) $\frac{\eta n_{.22}}{1 - \eta(1 - \pi_2)}$ above is of this form.

2) Show that given $\frac{\eta n_{.q-1q-1}}{1 - \eta(1 - \pi_{q-1})}$ is of the above form, then $\frac{\eta n_{.qq}}{1 - \eta(1 - \pi_q)}$ is also.

Substitute $\frac{\eta n_{.q-1q-1}}{1 - \eta(1 - \pi_{q-1})}$ of the above form into $\frac{\partial \ell}{\partial \pi_{q-1}} = \frac{f_{.q-1} + v_{.q-1}}{\pi_{q-1}} - \frac{f_{.1} + v_{.1}}{\pi_1} + \frac{\eta n_{.q-1q-1}}{1 - \eta(1 - \pi_{q-1})} - \frac{\eta n_{.11}}{1 - \eta(1 - \pi_1)} \equiv 0$.

Giving:

$$\begin{aligned} & \frac{f_{.q-1} + v_{.q-1}}{\pi_{q-1}} - \frac{f_{.1} + v_{.1}}{\pi_1} + \frac{V.}{(1 - \pi_{q-1})} + \sum_{i=2}^{q-2} \frac{f_{.i} + v_{.i}}{1 - \pi_{q-1}} \left(\frac{1 - \pi_i}{\pi_i} \right) - \frac{f_{.1} + v_{.1}}{1 - \pi_{q-1}} \left(\frac{(q-3) - \sum_{i=2}^{q-2} \pi_i}{\pi_1} \right) - \sum_{i=q}^K \frac{\eta n_{.ii}}{1 - \eta(1 - \pi_i)} \left(\frac{(1 - \pi_i)}{(1 - \pi_{q-1})} \right) - \\ & \frac{\eta n_{.11}}{1 - \eta(1 - \pi_1)} \left(\frac{(q-2) - \sum_{i=1}^{q-2} \pi_i}{1 - \pi_{q-1}} \right) - \frac{\eta n_{.11}}{1 - \eta(1 - \pi_1)} = 0 \end{aligned} \quad (I.4)$$

$$\Leftrightarrow \frac{V.}{(1-\pi_{q-1})} + \sum_{i=2}^{q-1} \frac{f_{.i} + v_{.i}}{1-\pi_{q-1}} \left(\frac{1-\pi_i}{\pi_i} \right) - \frac{f_{.1} + v_{.1}}{\pi_1} \left(\frac{(q-2) - \sum_{i=2}^{q-1} \pi_i}{1-\pi_{q-1}} \right) - \sum_{i=q}^K \frac{\eta n_{.ii}}{1-\eta(1-\pi_i)} \left(\frac{(1-\pi_i)}{(1-\pi_{q-1})} \right) - \frac{\eta n_{.11}}{1-\eta(1-\pi_1)} \left(\frac{(q-1) - \sum_{i=1}^{q-1} \pi_i}{1-\pi_{q-1}} \right) = 0. \quad (I.5)$$

Solve (I.5) for $\frac{\eta n_{.qq}}{1-\eta(1-\pi_q)}$:

$$\frac{\eta n_{.qq}}{1-\eta(1-\pi_q)} = \left(\frac{1-\pi_{q-1}}{1-\pi_q} \right) \left[\frac{V.}{(1-\pi_{q-1})} + \sum_{i=2}^{q-1} \frac{f_{.i} + v_{.i}}{1-\pi_{q-1}} \left(\frac{1-\pi_i}{\pi_i} \right) - \frac{f_{.1} + v_{.1}}{\pi_1} \left(\frac{(q-2) - \sum_{i=2}^{q-1} \pi_i}{1-\pi_{q-1}} \right) - \sum_{i=q+1}^K \frac{\eta n_{.ii}}{1-\eta(1-\pi_i)} \left(\frac{(1-\pi_i)}{(1-\pi_{q-1})} \right) - \frac{\eta n_{.11}}{1-\eta(1-\pi_1)} \left(\frac{(q-1) - \sum_{i=1}^{q-1} \pi_i}{1-\pi_{q-1}} \right) \right]$$

$$\Leftrightarrow \frac{\eta n_{.qq}}{1-\eta(1-\pi_q)} = \frac{V.}{1-\pi_q} + \sum_{i=2}^{q-1} \frac{f_{.i} + v_{.i}}{1-\pi_q} \left(\frac{1-\pi_i}{\pi_i} \right) - \frac{f_{.1} + v_{.1}}{\pi_1} \left(\frac{(q-2) - \sum_{i=2}^{q-1} \pi_i}{1-\pi_q} \right) - \sum_{i=q+1}^K \frac{\eta n_{.ii}}{1-\eta(1-\pi_i)} \left(\frac{1-\pi_i}{1-\pi_q} \right) - \frac{\eta n_{.11}}{1-\eta(1-\pi_1)} \left(\frac{(q-1) - \sum_{i=1}^{q-1} \pi_i}{1-\pi_q} \right)$$

(I.6)

Equation (I.6) is of the form of equation (I.3). Hence, by mathematical induction equation (I.3) is true for $2 \leq q \leq K$. Setting $q = K$ gives:

$$\frac{\eta n_{.KK}}{1-\eta(1-\pi_K)} = \frac{V.}{1-\pi_K} + \sum_{i=2}^{K-1} \frac{f_{.i} + v_{.i}}{1-\pi_K} \left(\frac{1-\pi_i}{\pi_i} \right) - \frac{f_{.1} + v_{.1}}{\pi_1} \left(\frac{(K-2) - \sum_{i=2}^{K-1} \pi_i}{1-\pi_K} \right) - \frac{\eta n_{.11}}{1-\eta(1-\pi_1)} \left(\frac{(K-1) - \sum_{i=1}^{K-1} \pi_i}{1-\pi_K} \right)$$

Substitute $\frac{\eta n_{.KK}}{1-\eta(1-\pi_K)}$ equality into $\frac{\partial \ell}{\partial \pi_K} = \frac{f_{.K} + v_{.K}}{\pi_K} - \frac{f_{.1} + v_{.1}}{\pi_1} + \frac{\eta n_{.KK}}{1-\eta(1-\pi_K)} - \frac{\eta n_{.11}}{1-\eta(1-\pi_1)} \equiv 0$.

$$\frac{\partial \ell}{\partial \pi_K} = \frac{f_{\cdot K} + v_{\cdot K}}{\pi_K} - \frac{f_{\cdot 1} + v_{\cdot 1}}{\pi_1} + \frac{V_{\cdot}}{1 - \pi_K} + \sum_{i=2}^{K-1} \frac{f_{\cdot i} + v_{\cdot i}}{1 - \pi_K} \left(\frac{1 - \pi_i}{\pi_i} \right) - \frac{f_{\cdot 1} + v_{\cdot 1}}{\pi_1} \left(\frac{(K-2) - \sum_{i=2}^{K-1} \pi_i}{1 - \pi_K} \right) - \frac{\eta n_{\cdot 11}}{1 - \eta(1 - \pi_1)} \left(\frac{(K-1) - \sum_{i=1}^{K-1} \pi_i}{1 - \pi_K} \right) - \frac{\eta n_{11}}{1 - \eta(1 - \pi_1)} = 0$$

$$\Leftrightarrow \frac{V_{\cdot}}{1 - \pi_K} + \sum_{i=2}^K \frac{f_{\cdot i} + v_{\cdot i}}{1 - \pi_K} \left(\frac{1 - \pi_i}{\pi_i} \right) - \frac{f_{\cdot 1} + v_{\cdot 1}}{\pi_1} \left(\frac{(K-1) - \sum_{i=2}^K \pi_i}{1 - \pi_K} \right) - \frac{\eta n_{\cdot 11}}{1 - \eta(1 - \pi_1)} \left(\frac{K - \sum_{i=1}^K \pi_i}{1 - \pi_K} \right) = 0$$

Now, $\sum_{i=2}^K \pi_i = 1 - \pi_1$ and $\sum_{i=1}^K \pi_i = 1$ giving:

$$\frac{V_{\cdot}}{1 - \pi_K} + \sum_{i=2}^K \frac{f_{\cdot i} + v_{\cdot i}}{1 - \pi_K} \left(\frac{1 - \pi_i}{\pi_i} \right) - \frac{f_{\cdot 1} + v_{\cdot 1}}{1 - \pi_K} \left(\frac{K - 2 + \pi_1}{\pi_1} \right) - \frac{\eta n_{\cdot 11}}{1 - \eta(1 - \pi_1)} \left(\frac{K - 1}{1 - \pi_K} \right) = 0$$

Solving for $\frac{\eta n_{.11}(K-1)}{1-\eta(1-\pi_1)}$ gives $\frac{\eta n_{.11}(K-1)}{1-\eta(1-\pi_1)} = \underbrace{V. + \sum_{i=2}^K (f_{.i} + v_{.i}) \left(\frac{1-\pi_i}{\pi_i} \right) - (f_{.1} + v_{.1}) \left(\frac{K-2+\pi_1}{\pi_1} \right)}_B$

$\frac{\eta n_{.11}(K-1)}{1-\eta(1-\pi_1)} = B \Leftrightarrow \eta = \frac{B}{(K-1)n_{.11} + B(1-\pi_1)}$ as desired.

APPENDIX A.5

FIRST AND SECOND DERIVATIVES OF $\ell(\eta, \boldsymbol{\beta} | \mathbf{X})$

$$\ell(\eta, \boldsymbol{\beta} | \mathbf{X}) \propto \sum_{i=1}^{K-1} (f_{\cdot i} + v_{\cdot i}) \ln \pi_i + (f_{\cdot K} + v_{\cdot K}) \ln \left(1 - \sum_{i=1}^{K-1} \pi_i \right) + V \cdot \ln \eta + \sum_{i=1}^{K-1} n_{\cdot ii} \ln (1 - \eta (1 - \pi_i)) + n_{\cdot KK} \ln \left(1 - \eta \sum_{i=1}^{K-1} \pi_i \right)$$

Where $1 - \sum_{i=1}^{K-1} \pi_i = \pi_K$ and $\sum_{i=1}^{K-1} \pi_i = 1 - \pi_K$.

Consider $\pi_i = \frac{e^{\beta_0 A_i + Z_i \boldsymbol{\beta}}}{\sum_{j=1}^K e^{\beta_0 A_j + Z_j \boldsymbol{\beta}}} = \frac{e^{Z_i \boldsymbol{\beta}}}{\sum_{j=1}^K e^{Z_j \boldsymbol{\beta}}}$ where $\boldsymbol{\beta} = [\beta_0, \beta_1, \dots, \beta_p]'$ are the unknown parameters for the habitat characteristics

$$Z_i = [A_i, z_{i1}, \dots, z_{ip}]$$

Then the first derivative of $\ell(\eta, \boldsymbol{\beta} | \mathbf{X})$ with respect to β_m is:

$$\frac{\partial \ell}{\partial \beta_m} = \sum_{i=1}^{K-1} (f_{\cdot i} + v_{\cdot i}) \pi_i^{-1} \frac{\partial \pi_i}{\partial \beta_m} - (f_{\cdot K} + v_{\cdot K}) \frac{\sum_{i=1}^{K-1} \frac{\partial \pi_i}{\partial \beta_m}}{1 - \sum_{i=1}^{K-1} \pi_i} + \sum_{i=1}^{K-1} n_{\cdot ii} [1 - \eta(1 - \pi_i)]^{-1} \left(\eta \frac{\partial \pi_i}{\partial \beta_m} \right) - n_{\cdot KK} \frac{\eta \sum_{i=1}^{K-1} \frac{\partial \pi_i}{\partial \beta_m}}{1 - \eta \sum_{i=1}^{K-1} \pi_i}$$

where

$$\frac{\partial \pi_i}{\partial \beta_m} = \frac{z_{im} e^{z_i \boldsymbol{\beta}} \left(\sum_{j=1}^K e^{z_j \boldsymbol{\beta}} \right) - e^{z_i \boldsymbol{\beta}} \left(\sum_{j=1}^K x_{jm} e^{z_j \boldsymbol{\beta}} \right)}{\left(\sum_{j=1}^K e^{z_j \boldsymbol{\beta}} \right)^2} = \frac{e^{z_i \boldsymbol{\beta}}}{\sum_{j=1}^K e^{z_j \boldsymbol{\beta}}} \left[\frac{\sum_{j=1}^K (z_{im} - z_{jm}) e^{z_j \boldsymbol{\beta}}}{\sum_{j=1}^K e^{z_j \boldsymbol{\beta}}} \right] \quad (I.7)$$

The second derivative of $\ell(\eta, \boldsymbol{\beta} | \mathbf{X})$ with respect to β_m and β_r is:

$$\begin{aligned} \frac{\partial^2 \ell}{\partial \beta_r \partial \beta_m} &= \left\{ \sum_{i=1}^{K-1} (f_{\cdot i} + v_{\cdot i}) \frac{\partial}{\partial \beta_r} \left(\pi_i^{-1} \frac{\partial \pi_i}{\partial \beta_m} \right) \right\} - (f_{\cdot K} + v_{\cdot K}) \left\{ \pi_K^{-2} \left(\sum_{i=1}^{K-1} \frac{\partial \pi_i}{\partial \beta_r} \right) \left(\sum_{i=1}^{K-1} \frac{\partial \pi_i}{\partial \beta_m} \right) + \pi_K^{-1} \sum_{i=1}^{K-1} \frac{\partial^2 \pi_i}{\partial \beta_r \partial \beta_m} \right\} - \\ &\sum_{i=1}^{K-1} n_{\cdot ii} \left[(1 - \eta(1 - \pi_i))^{-2} \eta^2 \left(\frac{\partial \pi_i}{\partial \beta_r} \right) \left(\frac{\partial \pi_i}{\partial \beta_m} \right) - (1 - \eta(1 - \pi_i))^{-1} \eta \left(\frac{\partial^2 \pi_i}{\partial \beta_r \partial \beta_m} \right) \right] - \\ &n_{\cdot KK} \left[(1 - \eta(1 - \pi_K))^{-2} \eta^2 \left(\sum_{i=1}^{K-1} \frac{\partial \pi_i}{\partial \beta_r} \right) \left(\sum_{i=1}^{K-1} \frac{\partial \pi_i}{\partial \beta_m} \right) + (1 - \eta(1 - \pi_K))^{-1} \eta \sum_{i=1}^{K-1} \left(\frac{\partial^2 \pi_i}{\partial \beta_r \partial \beta_m} \right) \right] \end{aligned}$$

where

$$\pi_i^{-1} \frac{\partial \pi_i}{\partial \beta_m} = \frac{\sum_{j=1}^K (z_{im} - z_{jm}) e^{z_j \boldsymbol{\beta}}}{\sum_{j=1}^K e^{z_j \boldsymbol{\beta}}}$$

$$\frac{\partial}{\partial \beta_r} \left(\pi_i^{-1} \frac{\partial \pi_i}{\partial \beta_m} \right) = \frac{\sum_{h=1}^K \sum_{j=1}^K (z_{im} - z_{jm}) (z_{jr} - z_{hr}) e^{(z_j + z_h)\beta}}{\left(\sum_{j=1}^K e^{z_j \beta} \right)^2}$$

$$\frac{\partial^2 \pi_i}{\partial \beta_r \partial \beta_m} = \frac{\pi_i \sum_{h=1}^K \sum_{j=1}^K (z_{im} - z_{jm}) (z_{ir} - 2z_{hr} + z_{jr}) e^{(z_j + z_h)\beta}}{\left(\sum_{j=1}^K e^{z_j \beta} \right)^2} \quad (I.8)$$

The first and second derivatives of $\ell(\eta, \boldsymbol{\beta} | X)$ with respect to η are:

$$\frac{\partial \ell}{\partial \eta} = \frac{V.}{\eta} - \sum_{i=1}^K \frac{n_{.ii} (1 - \pi_i)}{[1 - \eta(1 - \pi_i)]}$$

$$\frac{\partial^2 \ell}{\partial \eta^2} = -\frac{V.}{\eta^2} - \sum_{i=1}^K \frac{n_{.ii} (1 - \pi_i)^2}{[1 - \eta(1 - \pi_i)]^2}$$

The second derivative of $\ell(\eta, \boldsymbol{\beta} | X)$ with respect to β_m and η is:

$$\frac{\partial^2 \ell}{\partial \eta \partial \beta_m} = \sum_{i=1}^{K-1} n_{\cdot ii} \frac{\frac{\partial \pi_i}{\partial \beta_m}}{[1 - \eta(1 - \pi_i)]^2} - n_{\cdot KK} \frac{\sum_{i=1}^{K-1} \frac{\partial \pi_i}{\partial \beta_m}}{[1 - \eta(1 - \pi_K)]^2}$$

The observed information matrix is constructed using the negative of the second partial derivatives as per usual.

APPENDIX A.6

DELTA METHOD ESTIMATES OF VARIANCE FOR $\hat{\pi}_i$ AND $\frac{\hat{\pi}_i}{\hat{\pi}_k}$

If area is not offset then log area is included in the Z_j 's and β_0 is included in $\underline{\beta}$ below.

$$\text{Var}(\hat{\pi}_i) = \text{Var}\left(\frac{e^{Z_i\hat{\beta}}}{\sum_{j=1}^K e^{Z_j\hat{\beta}}}\right) \approx \left[\frac{\partial}{\partial \underline{\beta}} \frac{e^{Z_i\beta}}{\sum_{j=1}^K e^{Z_j\beta}}\right]_{\beta=\hat{\beta}} \text{Cov}(\hat{\underline{\beta}}) \left[\frac{\partial}{\partial \underline{\beta}} \frac{e^{Z_i\beta}}{\sum_{j=1}^K e^{Z_j\beta}}\right]'_{\beta=\hat{\beta}}$$

where $\frac{\partial}{\partial \underline{\beta}} \frac{e^{Z_i\beta}}{\sum_{j=1}^K e^{Z_j\beta}} = \frac{\partial}{\partial \underline{\beta}} \pi_i$ is given in appendix A.5 equation (I.7), and $\text{Cov}(\hat{\underline{\beta}})$ is the

inverse of the observed information matrix.

$$\text{Var}\left(\ln\left(\frac{\hat{\pi}_i}{\hat{\pi}_k}\right)\right) = \text{Var}\left((Z_i - Z_k)\hat{\underline{\beta}}\right) = (Z_i - Z_k)\text{Cov}(\hat{\underline{\beta}})(Z_i - Z_k)'$$

$$\text{Var}\left(\frac{\hat{\pi}_i}{\hat{\pi}_k}\right) = \text{Var}\left(\frac{e^{Z_i\hat{\beta}}}{e^{Z_k\hat{\beta}}}\right) \approx \left[\frac{\partial}{\partial \underline{\beta}} \frac{e^{Z_i\beta}}{e^{Z_k\beta}}\right]_{\beta=\hat{\beta}} \text{Cov}(\hat{\underline{\beta}}) \left[\frac{\partial}{\partial \underline{\beta}} \frac{e^{Z_i\beta}}{e^{Z_k\beta}}\right]'_{\beta=\hat{\beta}}$$

Here,

$$\frac{\partial}{\partial \underline{\beta}} \frac{e^{Z_i \underline{\beta}}}{e^{Z_K \underline{\beta}}} = \frac{Z_i e^{Z_i \underline{\beta}} e^{Z_K \underline{\beta}} - Z_K e^{Z_i \underline{\beta}} e^{Z_K \underline{\beta}}}{e^{2Z_K \underline{\beta}}} = \frac{(Z_i - Z_K) e^{Z_i \underline{\beta}}}{e^{Z_K \underline{\beta}}} = \frac{\pi_i}{\pi_K} (Z_i - Z_K)$$

giving,

$$\text{Var} \left(\frac{\hat{\pi}_i}{\hat{\pi}_K} \right) \approx \left(\frac{\hat{\pi}_i}{\hat{\pi}_K} \right)^2 (Z_i - Z_K) \text{Cov}(\hat{\underline{\beta}}) (Z_i - Z_K)' = \left(\frac{\hat{\pi}_i}{\hat{\pi}_K} \right)^2 \text{Var} \left(\ln \left(\frac{\hat{\pi}_i}{\hat{\pi}_K} \right) \right)$$

APPENDIX A.7

ASYMPTOTIC TESTS OF $H_0 : \eta = 1$ vs $H_1 : \eta \neq 1$

A.7.1 Asymptotic Likelihood Ratio Statistic and Test (LRT)

$$H_0 : \eta = 1 \text{ vs } H_1 : \eta \neq 1$$

The loglikelihood under the saturated persistence model is:

$$\ell(\eta, \boldsymbol{\pi} | \mathbf{X}) \propto \sum_{i=1}^K (f_{.i} + v_{.i}) \ln \pi_i + V \ln \eta + \sum_{i=1}^K n_{.ii} \ln(1 - \eta(1 - \pi_i)),$$

where $\pi_K = 1 - \sum_{i=1}^{K-1} \pi_i$. Under H_0 , this loglikelihood becomes:

$$\ell_0(\eta = 1, \boldsymbol{\pi} | \mathbf{X}) \propto \sum_{i=1}^K (f_{.i} + v_{.i} + n_{.ii}) \ln \pi_i,$$

which is identical to loglikelihood under the IMS model. The MLE's are obtained under both H_0 and H_1 , and the usual asymptotic LRT chi-squared statistic:

$$\text{LRT}_{\text{stat}} = -2(\ell_0(\eta = 1, \hat{\boldsymbol{\pi}}_0 | \mathbf{X}) - \ell(\hat{\eta}, \hat{\boldsymbol{\pi}} | \mathbf{X}))$$

is calculated. Under H_0 , LRT_{stat} is distributed as a chi-squared statistic on one degree of freedom, $\text{LRT}_{\text{stat}} \sim \chi_1^2$.

A.7.2 Score Statistic and Test

$$H_0 : \eta = 1 \text{ vs } H_1 : \eta \neq 1$$

The loglikelihood of the persistence model is:

$$\ell(\eta, \boldsymbol{\pi} | X) \propto \sum_{i=1}^K (f_{\cdot i} + v_{\cdot i}) \ln \pi_i + V \cdot \ln \eta + \sum_{i=1}^K n_{\cdot ii} \ln(1 - \eta(1 - \pi_i))$$

The first partial derivatives of $\ell(\eta, \boldsymbol{\pi} | X)$ at $\eta = 1$ are:

$$\left. \frac{\partial \ell(\eta, \boldsymbol{\pi} | X)}{\partial \pi_i} \right|_{\eta=1} = \frac{f_{\cdot i} + v_{\cdot i} + n_{\cdot ii}}{\pi_i} - \frac{f_{\cdot K} + v_{\cdot K} + n_{\cdot KK}}{\pi_K} \quad i = 1, \dots, K-1$$

$$\left. \frac{\partial \ell(\eta, \boldsymbol{\pi} | X)}{\partial \eta} \right|_{\eta=1} = V \cdot - \sum_{i=1}^K \frac{n_{\cdot ii} (1 - \pi_i)}{\pi_i}$$

The negative second partial derivatives of $\ell(\eta, \boldsymbol{\pi} | \mathbf{X})$ at $\eta = 1$ are:

$$-\frac{\partial^2 \ell(\eta, \boldsymbol{\pi} | \mathbf{X})}{\partial \pi_i^2} \Big|_{\eta=1} = \frac{f_{\cdot i} + v_{\cdot i} + n_{\cdot ii}}{\pi_i^2} + \frac{f_{\cdot K} + v_{\cdot K} + n_{\cdot KK}}{\pi_K^2}, \quad i = 1, \dots, K-1$$

$$-\frac{\partial^2 \ell(\eta, \boldsymbol{\pi} | \mathbf{X})}{\partial \pi_i \partial \pi_j} \Big|_{\eta=1} = \frac{f_{\cdot K} + v_{\cdot K} + n_{\cdot KK}}{\pi_K^2}, \quad \forall i \neq j$$

$$-\frac{\partial^2 \ell(\eta, \boldsymbol{\pi} | \mathbf{X})}{\partial \pi_i \partial \eta} \Big|_{\eta=1} = -\frac{n_{\cdot ii}}{\pi_i^2} + \frac{n_{\cdot KK}}{\pi_K^2}, \quad i = 1, \dots, K-1$$

$$-\frac{\partial^2 \ell(\eta, \boldsymbol{\pi} | \mathbf{X})}{\partial \eta^2} \Big|_{\eta=1} = V_{\cdot} + \sum_{i=1}^K \frac{n_{\cdot ii} (1 - \pi_i)^2}{\pi_i^2}$$

The score, $U(\eta = 1, \hat{\boldsymbol{\pi}}_0) = \frac{\partial \ell(\eta, \boldsymbol{\pi} | \mathbf{X})}{\partial \eta} \Big|_{\eta=1, \boldsymbol{\pi}=\hat{\boldsymbol{\pi}}_0}$ and the observed information matrix,

$$J(\eta = 1, \hat{\boldsymbol{\pi}}_0) =$$

$$\begin{bmatrix} -\frac{\partial^2 \ell}{\partial \pi_1^2} \Big|_{\eta=1, \hat{\boldsymbol{\pi}}_0} & -\frac{\partial^2 \ell}{\partial \pi_1 \partial \pi_2} \Big|_{\eta=1, \hat{\boldsymbol{\pi}}_0} & \dots & -\frac{\partial^2 \ell}{\partial \pi_1 \partial \pi_{K-1}} \Big|_{\eta=1, \hat{\boldsymbol{\pi}}_0} & -\frac{\partial^2 \ell}{\partial \pi_1 \partial \eta} \Big|_{\eta=1, \hat{\boldsymbol{\pi}}_0} \\ -\frac{\partial^2 \ell}{\partial \pi_1 \partial \pi_2} \Big|_{\eta=1, \hat{\boldsymbol{\pi}}_0} & -\frac{\partial^2 \ell}{\partial \pi_2^2} \Big|_{\eta=1, \hat{\boldsymbol{\pi}}_0} & & \vdots & -\frac{\partial^2 \ell}{\partial \pi_2 \partial \eta} \Big|_{\eta=1, \hat{\boldsymbol{\pi}}_0} \\ \vdots & & \ddots & & \vdots \\ -\frac{\partial^2 \ell}{\partial \pi_1 \partial \pi_{K-1}} \Big|_{\eta=1, \hat{\boldsymbol{\pi}}_0} & \dots & & -\frac{\partial^2 \ell}{\partial \pi_{K-1}^2} \Big|_{\eta=1, \hat{\boldsymbol{\pi}}_0} & -\frac{\partial^2 \ell}{\partial \pi_{K-1} \partial \eta} \Big|_{\eta=1, \hat{\boldsymbol{\pi}}_0} \\ -\frac{\partial^2 \ell}{\partial \pi_1 \partial \eta} \Big|_{\eta=1, \hat{\boldsymbol{\pi}}_0} & -\frac{\partial^2 \ell}{\partial \pi_2 \partial \eta} \Big|_{\eta=1, \hat{\boldsymbol{\pi}}_0} & & -\frac{\partial^2 \ell}{\partial \pi_{K-1} \partial \eta} \Big|_{\eta=1, \hat{\boldsymbol{\pi}}_0} & -\frac{\partial^2 \ell}{\partial \eta^2} \Big|_{\eta=1, \hat{\boldsymbol{\pi}}_0} \end{bmatrix}$$

where $\hat{\boldsymbol{\pi}}_0$ is the MLE for $\boldsymbol{\pi}$ under H_0 .

The score statistic, $\text{Score}_{\text{stat}} = U(\eta = 1, \hat{\boldsymbol{\pi}}_0)^2 * J^{-1}_{K,K}$ has a chi-squared distribution on one degree of freedom under H_0 , $\text{Score}_{\text{stat}} \sim \chi^2_1$. $J^{-1}_{K,K}$ is the K^{th} diagonal entry of the inverse of $J(\eta = 1, \hat{\boldsymbol{\pi}}_0)$.

A.7.3 Wald Statistic and Test

$\text{Wald}_{\text{stat}} = \frac{(\hat{\eta} - 1)^2}{J^{-1}_{K,K}}$ has a chi-squared distribution on one degree of freedom under H_0 ,

$\text{Wald}_{\text{stat}} \sim \chi^2_1$. Here $\hat{\eta}$ is the MLE for η under the persistence model.

APPENDIX A.8

SIGHTING HISTORY MATRIX

$$X = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 4 & 5 & 4 & 6 & 6 & 6 & 6 & 6 & 6 & 3 & 6 & 6 \\ 6 & 6 & 1 & 6 & 6 & 6 & 6 & 6 & 6 & 6 & 6 & 5 & 6 & 6 & 6 & 6 & 4 & 4 & 4 \\ 5 & 5 & 1 & 6 & 4 & 6 & 5 & 5 & 4 & 5 & 5 & 6 & 5 & 3 & 5 & 3 & 5 & 3 & 3 \\ 3 & 5 & 6 & 5 & 5 & 1 & 5 \end{bmatrix}$$

Random ordering gives:

$$X = \begin{bmatrix} 5 & 4 & 4 & 5 & 5 & 5 & 4 & 5 & 5 & 4 & 5 & 5 & 4 & 5 & 1 & 5 & 4 & 4 & 1 \\ 2 & 5 & 5 & 5 & 3 & 4 & 5 & 5 & 3 & 5 & 1 & 4 & 4 & 3 & 2 & 5 & 5 & 1 & 5 \\ 4 & 4 & 3 & 2 & 2 & 3 & 5 & 5 & 1 & 5 & 5 & 3 & 1 & 4 & 1 & 3 & 4 & 2 & 1 \\ 5 & 4 & 5 & 5 & 5 & 2 & 1 \end{bmatrix}$$

APPENDIX A.9

TABLES OF SIZE FOR $H_0 : \eta = 1$ vs $H_1 : \eta \neq 1$

Level $\alpha = 0.10$ test.				Size x 100		
T	n_r	π	SS	score	Wald	LRT
50	50	.25, .25, .25, .25	5000	9.60	9.64	9.54
50	30	.25, .25, .25, .25	5000	9.92	10.14	9.96
50	10	.25, .25, .25, .25	5000	9.60	9.30	9.68
30	50	.25, .25, .25, .25	5000	10.38	10.20	10.30
30	30	.25, .25, .25, .25	5000	10.42	10.02	10.46
30	10	.25, .25, .25, .25	5000	10.72	9.18	10.38
10	50	.25, .25, .25, .25	5000	10.78	9.90	10.70
10	30	.25, .25, .25, .25	5000	10.62	9.72	9.98
10	10	.25, .25, .25, .25	4942	11.19	9.57	10.56
50	50	.1, .2, .3, .4	5000	10.96	10.42	10.82
50	30	.1, .2, .3, .4	5000	10.60	10.10	10.32
50	10	.1, .2, .3, .4	4833	10.49	9.37	10.47
30	50	.1, .2, .3, .4	5000	9.66	8.86	9.40
30	30	.1, .2, .3, .4	4997	10.13	9.39	9.99
30	10	.1, .2, .3, .4	4434	9.22	7.89	8.77
10	50	.1, .2, .3, .4	4944	10.28	8.84	9.79
10	30	.1, .2, .3, .4	4621	9.35	8.35	8.59
10	10	.1, .2, .3, .4	2731	6.59	8.90	6.81
50	50	.15, .15, .35, .35	5000	10.18	10.10	9.98
50	30	.15, .15, .35, .35	5000	10.12	9.70	10.00
50	10	.15, .15, .35, .35	4998	10.64	9.58	10.32
30	50	.15, .15, .35, .35	5000	9.72	9.80	9.96
30	30	.15, .15, .35, .35	5000	10.28	10.02	10.18
30	10	.15, .15, .35, .35	4961	10.91	9.55	10.40
10	50	.15, .15, .35, .35	4998	10.12	9.32	10.00
10	30	.15, .15, .35, .35	4987	11.55	9.36	10.87
10	10	.15, .15, .35, .35	4122	9.75	8.61	9.15
50	50	.2, .2, .2, .4	5000	10.70	10.44	10.54
50	30	.2, .2, .2, .4	5000	9.92	10.12	9.88
50	10	.2, .2, .2, .4	5000	10.38	9.76	10.58
30	50	.2, .2, .2, .4	5000	9.42	9.68	9.58
30	30	.2, .2, .2, .4	5000	10.06	9.36	10.04
30	10	.2, .2, .2, .4	5000	10.48	8.96	10.08
10	50	.2, .2, .2, .4	5000	9.64	9.38	9.48
10	30	.2, .2, .2, .4	5000	10.92	9.96	10.54
10	10	.2, .2, .2, .4	4748	10.09	7.65	9.14

Level $\alpha = 0.05$ test.				Size x 100		
T	n_t	π	SS	Score	Wald	LRT
50	50	.25, .25, .25, .25	5000	4.94	4.82	4.92
50	30	.25, .25, .25, .25	5000	4.96	4.86	4.92
50	10	.25, .25, .25, .25	5000	4.70	4.10	4.54
30	50	.25, .25, .25, .25	5000	5.26	5.26	5.34
30	30	.25, .25, .25, .25	5000	5.14	5.16	5.06
30	10	.25, .25, .25, .25	5000	5.20	4.38	5.04
10	50	.25, .25, .25, .25	5000	5.14	5.18	5.20
10	30	.25, .25, .25, .25	5000	5.04	5.10	5.02
10	10	.25, .25, .25, .25	4942	6.07	4.69	5.83
50	50	.1, .2, .3, .4	5000	5.44	5.18	5.36
50	30	.1, .2, .3, .4	5000	5.50	5.22	5.34
50	10	.1, .2, .3, .4	4833	5.63	4.82	5.40
30	50	.1, .2, .3, .4	5000	4.64	4.36	4.58
30	30	.1, .2, .3, .4	4997	5.26	4.36	4.98
30	10	.1, .2, .3, .4	4434	4.71	3.81	4.31
10	50	.1, .2, .3, .4	4944	5.34	4.59	4.92
10	30	.1, .2, .3, .4	4621	4.37	4.13	4.18
10	10	.1, .2, .3, .4	2731	2.71	5.71	3.15
50	50	.15, .15, .35, .35	5000	5.04	4.76	4.88
50	30	.15, .15, .35, .35	5000	5.12	5.16	5.12
50	10	.15, .15, .35, .35	4998	5.32	4.84	5.10
30	50	.15, .15, .35, .35	5000	4.76	4.44	4.70
30	30	.15, .15, .35, .35	5000	5.30	5.02	5.00
30	10	.15, .15, .35, .35	4961	5.70	4.92	5.36
10	50	.15, .15, .35, .35	4998	5.26	4.44	4.94
10	30	.15, .15, .35, .35	4987	6.30	4.91	5.59
10	10	.15, .15, .35, .35	4122	4.76	4.88	4.54
50	50	.2, .2, .2, .4	5000	5.02	5.06	5.06
50	30	.2, .2, .2, .4	5000	5.24	4.96	5.14
50	10	.2, .2, .2, .4	5000	5.34	5.00	5.10
30	50	.2, .2, .2, .4	5000	4.88	4.82	4.62
30	30	.2, .2, .2, .4	5000	4.72	4.34	4.52
30	10	.2, .2, .2, .4	5000	5.16	3.98	4.64
10	50	.2, .2, .2, .4	5000	5.20	4.66	5.14
10	30	.2, .2, .2, .4	5000	5.76	4.80	5.28
10	10	.2, .2, .2, .4	4748	5.33	3.92	4.25

Level $\alpha = 0.01$ test.				Size x 100		
T	n_i	π	SS	score	Wald	LRT
50	50	.25, .25, .25, .25	5000	1.04	1.02	0.96
50	30	.25, .25, .25, .25	5000	1.20	1.16	1.18
50	10	.25, .25, .25, .25	5000	1.04	0.74	0.82
30	50	.25, .25, .25, .25	5000	1.08	1.00	1.04
30	30	.25, .25, .25, .25	5000	1.00	1.00	0.98
30	10	.25, .25, .25, .25	5000	1.26	1.04	1.00
10	50	.25, .25, .25, .25	5000	1.30	1.02	1.28
10	30	.25, .25, .25, .25	5000	1.22	1.02	1.08
10	10	.25, .25, .25, .25	4942	1.76	1.03	1.40
50	50	.1, .2, .3, .4	5000	0.88	0.94	0.86
50	30	.1, .2, .3, .4	5000	1.14	1.08	1.06
50	10	.1, .2, .3, .4	4833	1.28	1.08	0.97
30	50	.1, .2, .3, .4	5000	0.98	0.82	0.80
30	30	.1, .2, .3, .4	4997	1.02	0.90	0.84
30	10	.1, .2, .3, .4	4434	1.04	1.17	0.70
10	50	.1, .2, .3, .4	4944	1.42	1.15	1.15
10	30	.1, .2, .3, .4	4621	0.91	1.28	0.78
10	10	.1, .2, .3, .4	2731	0.51	2.31	0.55
50	50	.15, .15, .35, .35	5000	0.92	1.12	0.90
50	30	.15, .15, .35, .35	5000	1.20	1.10	1.16
50	10	.15, .15, .35, .35	4998	1.32	1.12	1.20
30	50	.15, .15, .35, .35	5000	0.86	0.92	0.72
30	30	.15, .15, .35, .35	5000	0.88	1.04	0.84
30	10	.15, .15, .35, .35	4961	1.59	1.27	1.25
10	50	.15, .15, .35, .35	4998	1.12	0.88	0.92
10	30	.15, .15, .35, .35	4987	1.60	1.18	1.14
10	10	.15, .15, .35, .35	4122	1.12	2.01	0.75
50	50	.2, .2, .2, .4	5000	0.98	0.92	1.04
50	30	.2, .2, .2, .4	5000	1.18	1.08	1.10
50	10	.2, .2, .2, .4	5000	1.04	0.96	1.06
30	50	.2, .2, .2, .4	5000	0.92	1.02	0.94
30	30	.2, .2, .2, .4	5000	1.08	0.98	0.94
30	10	.2, .2, .2, .4	5000	1.18	0.68	0.92
10	50	.2, .2, .2, .4	5000	1.14	1.22	1.12
10	30	.2, .2, .2, .4	5000	1.42	0.80	1.26
10	10	.2, .2, .2, .4	4748	1.58	1.10	1.16

APPENDIX A.10

TABLES OF POWER FOR $H_0 : \eta = 1$ VS $H_1 : \eta \neq 1$

K = 3				Power x 100		
T	n_i	π	η	Score	Wald	LRT
50	50	1/3, 1/3, 1/3	0.8	100.00	100.00	100.00
50	30	1/3, 1/3, 1/3	0.8	100.00	100.00	100.00
50	10	1/3, 1/3, 1/3	0.8	100.00	100.00	100.00
30	50	1/3, 1/3, 1/3	0.8	100.00	100.00	100.00
30	30	1/3, 1/3, 1/3	0.8	100.00	100.00	100.00
30	10	1/3, 1/3, 1/3	0.8	98.80	99.30	99.20
10	50	1/3, 1/3, 1/3	0.8	100.00	100.00	100.00
10	30	1/3, 1/3, 1/3	0.8	99.00	99.30	99.10
10	10	1/3, 1/3, 1/3	0.8	65.10	73.80	68.60
50	50	1/3, 1/3, 1/3	0.9	100.00	100.00	100.00
50	30	1/3, 1/3, 1/3	0.9	100.00	100.00	100.00
50	10	1/3, 1/3, 1/3	0.9	79.50	81.70	80.50
30	50	1/3, 1/3, 1/3	0.9	100.00	100.00	100.00
30	30	1/3, 1/3, 1/3	0.9	98.00	98.60	98.20
30	10	1/3, 1/3, 1/3	0.9	58.30	65.40	60.80
10	50	1/3, 1/3, 1/3	0.9	84.70	86.50	85.50
10	30	1/3, 1/3, 1/3	0.9	60.30	65.50	62.70
10	10	1/3, 1/3, 1/3	0.9	17.80	24.40	19.70
50	50	1/3, 1/3, 1/3	0.95	92.00	93.40	92.40
50	30	1/3, 1/3, 1/3	0.95	75.00	77.60	75.80
50	10	1/3, 1/3, 1/3	0.95	28.80	33.50	30.80
30	50	1/3, 1/3, 1/3	0.95	72.50	75.20	73.90
30	30	1/3, 1/3, 1/3	0.95	53.70	56.60	54.60
30	10	1/3, 1/3, 1/3	0.95	17.70	22.70	19.80
10	50	1/3, 1/3, 1/3	0.95	32.50	35.80	34.10
10	30	1/3, 1/3, 1/3	0.95	18.50	24.40	20.90
10	10	1/3, 1/3, 1/3	0.95	6.80	9.90	7.90
50	50	1/3, 1/3, 1/3	1.20	100.00	100.00	100.00
50	30	1/3, 1/3, 1/3	1.20	100.00	100.00	100.00
50	10	1/3, 1/3, 1/3	1.20	100.00	100.00	100.00
30	50	1/3, 1/3, 1/3	1.20	100.00	100.00	100.00
30	30	1/3, 1/3, 1/3	1.20	100.00	100.00	100.00
30	10	1/3, 1/3, 1/3	1.20	99.90	99.80	99.90
10	50	1/3, 1/3, 1/3	1.20	100.00	100.00	100.00
10	30	1/3, 1/3, 1/3	1.20	99.90	99.90	99.90
10	10	1/3, 1/3, 1/3	1.20	86.99	77.64	85.57

K = 4				Power x 100		
T	n_t	π	η	Score	Wald	LRT
50	50	.25, .25, .25, .25	0.8	100.00	100.00	100.00
50	30	.25, .25, .25, .25	0.8	100.00	100.00	100.00
50	10	.25, .25, .25, .25	0.8	100.00	100.00	100.00
30	50	.25, .25, .25, .25	0.8	100.00	100.00	100.00
30	30	.25, .25, .25, .25	0.8	100.00	100.00	100.00
30	10	.25, .25, .25, .25	0.8	99.90	100.00	99.90
10	50	.25, .25, .25, .25	0.8	100.00	100.00	100.00
10	30	.25, .25, .25, .25	0.8	100.00	100.00	100.00
10	10	.25, .25, .25, .25	0.8	76.65	85.97	81.26
50	50	.25, .25, .25, .25	0.9	100.00	100.00	100.00
50	30	.25, .25, .25, .25	0.9	100.00	100.00	100.00
50	10	.25, .25, .25, .25	0.9	91.90	94.70	93.10
30	50	.25, .25, .25, .25	0.9	100.00	100.00	100.00
30	30	.25, .25, .25, .25	0.9	99.70	99.90	99.70
30	10	.25, .25, .25, .25	0.9	72.50	79.70	75.40
10	50	.25, .25, .25, .25	0.9	95.10	96.90	95.90
10	30	.25, .25, .25, .25	0.9	74.40	82.60	77.50
10	10	.25, .25, .25, .25	0.9	24.47	39.58	30.72
50	50	.25, .25, .25, .25	0.95	98.30	98.60	98.50
50	30	.25, .25, .25, .25	0.95	87.60	89.60	88.30
50	10	.25, .25, .25, .25	0.95	38.80	46.40	41.30
30	50	.25, .25, .25, .25	0.95	87.20	89.20	87.50
30	30	.25, .25, .25, .25	0.95	67.60	72.30	69.80
30	10	.25, .25, .25, .25	0.95	20.10	28.40	23.00
10	50	.25, .25, .25, .25	0.95	41.70	48.90	44.30
10	30	.25, .25, .25, .25	0.95	24.00	31.70	26.90
10	10	.25, .25, .25, .25	0.95	7.83	13.86	9.54
50	50	.25, .25, .25, .25	1.20	100.00	100.00	100.00
50	30	.25, .25, .25, .25	1.20	100.00	100.00	100.00
50	10	.25, .25, .25, .25	1.20	100.00	100.00	100.00
30	50	.25, .25, .25, .25	1.20	100.00	100.00	100.00
30	30	.25, .25, .25, .25	1.20	100.00	100.00	100.00
30	10	.25, .25, .25, .25	1.20	100.00	100.00	100.00
10	50	.25, .25, .25, .25	1.20	100.00	100.00	100.00
10	30	.25, .25, .25, .25	1.20	100.00	100.00	100.00
10	10	.25, .25, .25, .25	1.20	97.84	89.88	96.90

K = 5				Power x 100		
T	n_i	π	η	Score	Wald	LRT
50	50	.2, .2, .2, .2, .2	0.8	100.00	100.00	100.00
50	30	.2, .2, .2, .2, .2	0.8	100.00	100.00	100.00
50	10	.2, .2, .2, .2, .2	0.8	100.00	100.00	100.00
30	50	.2, .2, .2, .2, .2	0.8	100.00	100.00	100.00
30	30	.2, .2, .2, .2, .2	0.8	100.00	100.00	100.00
30	10	.2, .2, .2, .2, .2	0.8	100.00	100.00	100.00
10	50	.2, .2, .2, .2, .2	0.8	100.00	100.00	100.00
10	30	.2, .2, .2, .2, .2	0.8	100.00	100.00	100.00
10	10	.2, .2, .2, .2, .2	0.8	85.25	93.08	88.81
50	50	.2, .2, .2, .2, .2	0.9	100.00	100.00	100.00
50	30	.2, .2, .2, .2, .2	0.9	100.00	100.00	100.00
50	10	.2, .2, .2, .2, .2	0.9	96.00	97.80	96.60
30	50	.2, .2, .2, .2, .2	0.9	100.00	100.00	100.00
30	30	.2, .2, .2, .2, .2	0.9	100.00	100.00	100.00
30	10	.2, .2, .2, .2, .2	0.9	81.80	89.10	85.70
10	50	.2, .2, .2, .2, .2	0.9	97.60	98.80	98.40
10	30	.2, .2, .2, .2, .2	0.9	84.80	89.70	86.70
10	10	.2, .2, .2, .2, .2	0.9	25.79	44.01	33.27
50	50	.2, .2, .2, .2, .2	0.95	99.60	99.80	99.60
50	30	.2, .2, .2, .2, .2	0.95	95.20	95.80	95.50
50	10	.2, .2, .2, .2, .2	0.95	45.80	55.30	49.10
30	50	.2, .2, .2, .2, .2	0.95	95.30	96.10	95.70
30	30	.2, .2, .2, .2, .2	0.95	78.10	83.40	79.90
30	10	.2, .2, .2, .2, .2	0.95	26.10	38.70	30.40
10	50	.2, .2, .2, .2, .2	0.95	50.00	58.70	53.40
10	30	.2, .2, .2, .2, .2	0.95	27.10	38.70	31.30
10	10	.2, .2, .2, .2, .2	0.95	7.33	18.39	12.19
50	50	.2, .2, .2, .2, .2	1.10	100.00	100.00	100.00
50	30	.2, .2, .2, .2, .2	1.10	100.00	100.00	100.00
50	10	.2, .2, .2, .2, .2	1.10	99.80	99.20	99.60
30	50	.2, .2, .2, .2, .2	1.10	100.00	100.00	100.00
30	30	.2, .2, .2, .2, .2	1.10	100.00	100.00	100.00
30	10	.2, .2, .2, .2, .2	1.10	96.58	91.04	95.77
10	50	.2, .2, .2, .2, .2	1.10	100.00	99.70	100.00
10	30	.2, .2, .2, .2, .2	1.10	98.19	94.26	97.78
10	10	.2, .2, .2, .2, .2	1.10	63.18	26.10	55.99

APPENDIX A.11

TABLES OF SIZE FOR $H_0 : \beta_1 = 0$ VS $H_1 : \beta_1 \neq 0$

T = 30

					Size x 100	
T	n_t	π	η	Z	Pers.	IMS
30	50	1/3, 1/3, 1/3	.4	1, 0, 0	5.45	32.25
30	30	1/3, 1/3, 1/3	.4	1, 0, 0	5.80	32.70
30	10	1/3, 1/3, 1/3	.4	1, 0, 0	5.15	28.25
30	50	1/3, 1/3, 1/3	.9	1, 0, 0	5.55	7.85
30	30	1/3, 1/3, 1/3	.9	1, 0, 0	4.80	7.80
30	10	1/3, 1/3, 1/3	.9	1, 0, 0	6.15	7.65
30	50	1/3, 1/3, 1/3	1.2	1, 0, 0	5.35	1.85
30	30	1/3, 1/3, 1/3	1.2	1, 0, 0	4.70	1.85
30	10	1/3, 1/3, 1/3	1.2	1, 0, 0	4.60	1.65

T = 10

					Size x 100	
T	n_t	π	η	Z	Pers.	IMS
10	50	1/3, 1/3, 1/3	.4	1, 0, 0	5.25	32.50
10	30	1/3, 1/3, 1/3	.4	1, 0, 0	4.90	32.25
10	10	1/3, 1/3, 1/3	.4	1, 0, 0	4.80	28.65
10	50	1/3, 1/3, 1/3	.9	1, 0, 0	4.50	7.00
10	30	1/3, 1/3, 1/3	.9	1, 0, 0	4.50	6.75
10	10	1/3, 1/3, 1/3	.9	1, 0, 0	5.25	7.80
10	50	1/3, 1/3, 1/3	1.2	1, 0, 0	5.35	1.65
10	30	1/3, 1/3, 1/3	1.2	1, 0, 0	4.30	1.50
10	10	1/3, 1/3, 1/3	1.2	1, 0, 0	5.30	1.80

APPENDIX A.12

UNCONDITIONAL DISTRIBUTION OF Y_t

$$\begin{aligned}
f(y_t) &= \int \cdots \int \mathbb{P}(Y_t = y_t \mid \boldsymbol{\omega}_t) f(\boldsymbol{\omega}_t) \partial\omega_{t1} \cdots \partial\omega_{tK} \\
&= \int \cdots \int \frac{n_t!}{\prod_{i=1}^K y_{ti}!} \prod_{i=1}^K \omega_{ti}^{y_{ti}} \frac{1}{D(\boldsymbol{\pi}\gamma^{-1})} \prod_{i=1}^K \omega_{ti}^{\pi_i\gamma^{-1}-1} \partial\omega_{t1} \cdots \partial\omega_{tK} \\
&= \frac{n_t!}{\prod_{i=1}^K y_{ti}!} \frac{1}{D(\boldsymbol{\pi}\gamma^{-1})} \int \cdots \int \prod_{i=1}^K \omega_{ti}^{y_{ti} + \pi_i\gamma^{-1} - 1} \partial\omega_{t1} \cdots \partial\omega_{tK} \\
&= \frac{n_t!}{\prod_{i=1}^K y_{ti}!} \frac{1}{D(\boldsymbol{\pi}\gamma^{-1})} D(\boldsymbol{\pi}\gamma^{-1} + y_t) \quad \text{by (2.4)} \\
&= \frac{n_t! \Gamma(\gamma^{-1})}{\Gamma(\gamma^{-1} + n_t)} \prod_{i=1}^K \frac{\Gamma(\pi_i\gamma^{-1} + y_{ti})}{\Gamma(\pi_i\gamma^{-1}) y_{ti}!}
\end{aligned}$$

Now,

$$\begin{aligned}
\frac{n_t! \Gamma(\gamma^{-1})}{\Gamma(\gamma^{-1} + n_t)} &= \frac{n_t! (\gamma^{-1} - 1) \cdots (\gamma^{-1} - \lfloor \gamma^{-1} \rfloor) \Gamma(\gamma^{-1} - \lfloor \gamma^{-1} \rfloor)}{(\gamma^{-1} + n_t - 1) \cdots (\gamma^{-1} + n_t - \lfloor \gamma^{-1} \rfloor - n_t) \Gamma(\gamma^{-1} + n_t - \lfloor \gamma^{-1} \rfloor - n_t)} \\
&= \binom{\gamma^{-1} + n_t - 1}{n_t}_{\mathbb{R}}
\end{aligned}$$

where $\lfloor \gamma^{-1} \rfloor$ is defined as the integer part of γ^{-1} . Further,

$$\frac{\Gamma(\pi_i \gamma^{-1} + y_{ii})}{\Gamma(\pi_i \gamma^{-1}) y_{ii}!} = \frac{(\pi_i \gamma^{-1} + y_{ii} - 1) \cdots (\pi_i \gamma^{-1} + y_{ii} - \lfloor \pi_i \gamma^{-1} \rfloor - y_{ii}) \Gamma(\pi_i \gamma^{-1} - \lfloor \pi_i \gamma^{-1} \rfloor)}{y_{ii}! (\pi_i \gamma^{-1} - 1) \cdots (\pi_i \gamma^{-1} - \lfloor \pi_i \gamma^{-1} \rfloor) \Gamma(\pi_i \gamma^{-1} - \lfloor \pi_i \gamma^{-1} \rfloor)}$$

$$= \binom{\pi_i \gamma^{-1} + y_{ii} - 1}{\pi_i \gamma^{-1} - 1}_{\mathbb{R}}$$

Giving $f(y_t) = \binom{\gamma^{-1} + n_t - 1}{n_t}_{\mathbb{R}} \prod_{i=1}^K \binom{\pi_i \gamma^{-1} + y_{ii} - 1}{\pi_i \gamma^{-1} - 1}_{\mathbb{R}}$

APPENDIX A.13

MOMENTS OF Y_i

$$E[Y_{it}] = E_{\gamma, \pi} [E_{\omega} [Y_{it} | \omega_t]] = E_{\gamma, \pi} [n_t \omega_{it}] = n_t \pi_i$$

$$\begin{aligned} \text{Var}(Y_{it}) &= E_{\gamma, \pi} [\text{Var}_{\omega} (Y_{it} | \omega_t)] + \text{Var}_{\gamma, \pi} (E_{\omega} [Y_{it} | \omega_t]) \\ &= n_t \pi_i (1 - \pi_i) \frac{1 + n_t \gamma}{1 + \gamma} \end{aligned}$$

$$\begin{aligned} \text{Cov}(Y_{it}, Y_{jt}) &= E_{\gamma, \pi} [\text{Cov}_{\omega} (Y_{it}, Y_{jt} | \omega_t)] + \text{Cov}_{\gamma, \pi} (E_{\omega} [Y_{it} | \omega_t], E_{\omega} [Y_{jt} | \omega_t]) \\ &= -n_t \pi_i \pi_j \frac{1 + n_t \gamma}{1 + \gamma} \end{aligned}$$

APPENDIX A.14

CALCULATION AND REDUCTION OF THE LOGLIKELIHOOD, $\ell(\gamma, \boldsymbol{\pi} | \mathbf{y})$

Calculating the log-likelihood is simplified by first writing out the likelihood in a longer form:

$$L(\gamma, \boldsymbol{\pi} | \mathbf{y}) = \prod_{t=1}^T \left[\frac{n_t! (\gamma^{-1} - 1) \cdots (\gamma^{-1} - \lfloor \gamma^{-1} \rfloor)}{(\gamma^{-1} + n_t - 1) \cdots (\gamma^{-1} + n_t - \lfloor \gamma^{-1} \rfloor - n_t)} \prod_{i=1}^K \frac{(\pi_i \gamma^{-1} + y_{ti} - 1) \cdots (\pi_i \gamma^{-1} + y_{ti} - \lfloor \pi_i \gamma^{-1} \rfloor - y_{ti})}{y_{ti}! (\pi_i \gamma^{-1} - 1) \cdots (\pi_i \gamma^{-1} - \lfloor \pi_i \gamma^{-1} \rfloor)} \right]$$

then

$$\ell(\gamma, \boldsymbol{\pi} | \mathbf{y}) = \sum_{t=1}^T \left\{ \left[\ln(n_t!) + \left(\sum_{j=\gamma^{-1}-\lfloor \gamma^{-1} \rfloor}^{\gamma^{-1}-1} \ln j \right) - \left(\sum_{j=\gamma^{-1}-\lfloor \gamma^{-1} \rfloor}^{\gamma^{-1}+n_t-1} \ln j \right) \right] + \sum_{i=1}^K \left[\left(\sum_{j=\pi_i \gamma^{-1}-\lfloor \pi_i \gamma^{-1} \rfloor}^{\pi_i \gamma^{-1}+y_{ti}-1} \ln j \right) - \left(\sum_{j=\pi_i \gamma^{-1}-\lfloor \pi_i \gamma^{-1} \rfloor}^{\pi_i \gamma^{-1}-1} \ln j \right) - \ln(y_{ti}!) \right] \right\}$$

Removing factors not involving $\boldsymbol{\pi}$ or γ and reducing within the summations gives:

$$\begin{aligned}
\ell(\gamma, \boldsymbol{\pi} | \mathbf{y}) &\propto \sum_{t=1}^T \left\{ \left[- \sum_{j=\gamma^{-1}}^{\gamma^{-1} + n_t - 1} \ln j \right] + \sum_{i=1}^K \left[\sum_{j=\pi_i \gamma^{-1}}^{\pi_i \gamma^{-1} + y_{ti} - 1} \ln j \right] \right\} \\
&= - \sum_{t=1}^T \sum_{j=1}^{n_t} \ln(\gamma^{-1} + j - 1) + \sum_{t=1}^T \sum_{i=1}^K \sum_{j=1}^{y_{ti}} \ln(\pi_i \gamma + j - 1) \\
&= - \sum_{t=1}^T \sum_{j=1}^{n_t} (\ln(1 + [j-1]\gamma) - \ln \gamma) + \sum_{t=1}^T \sum_{i=1}^K \sum_{j=1}^{y_{ti}} (\ln(\pi_i + [j-1]\gamma) - \ln \gamma) \\
&= - \sum_{t=1}^T \sum_{j=1}^{n_t} \ln(1 + [j-1]\gamma) + \sum_{t=1}^T \sum_{i=1}^K \sum_{j=1}^{y_{ti}} \ln(\pi_i + [j-1]\gamma) + \underbrace{\sum_{t=1}^T \sum_{j=1}^{n_t} \ln \gamma - \sum_{t=1}^T \sum_{i=1}^K \sum_{j=1}^{y_{ti}} \ln \gamma}_{*}
\end{aligned}$$

$$* = \sum_{t=1}^T n_t \ln \gamma - \sum_{t=1}^T \sum_{i=1}^K y_{ti} \ln \gamma = \sum_{t=1}^T n_t \ln \gamma - \sum_{t=1}^T n_t \ln \gamma = 0 \quad \left(\text{since } n_t = \sum_{i=1}^K y_{ti} \right)$$

Therefore,

$$\ell(\gamma, \boldsymbol{\pi} | \mathbf{y}) \propto - \sum_{t=1}^T \sum_{j=1}^{n_t} \ln(1 + [j-1]\gamma) + \sum_{t=1}^T \sum_{i=1}^K \sum_{j=1}^{y_{ti}} \ln(\pi_i + [j-1]\gamma)$$

APPENDIX A.15

INITIAL ESTIMATE OF $\hat{\gamma}$

Consider: $\text{Var}(\pi_i) = \frac{\pi_i(1-\pi_i)}{n_*} \frac{1+n_*\gamma}{1+\gamma}$, where n_* is some measure of central

$n_t : t = 1, \dots, T$, giving $\widehat{\text{Var}}(\pi_i) = \frac{\hat{\pi}_{0i}(1-\hat{\pi}_{0i})}{n_*} \frac{1+n_*\hat{\gamma}}{1+\hat{\gamma}}$.

However, $\widehat{\text{Var}}(\hat{\pi}_{0i}) = \frac{\sum_{t=1}^T (\hat{\pi}_{ti} - \hat{\pi}_{0i})^2}{T-1}$.

Let $Q = \frac{\sum_{t=1}^T (\hat{\pi}_{ti} - \hat{\pi}_{0i})^2 n_*}{(T-1)\hat{\pi}_{0i}(1-\hat{\pi}_{0i})}$. Then the initial estimate of $\hat{\gamma}$ is $\frac{Q-1}{n_*-Q}$.

APPENDIX A.16

CHECK FOR DECREASING LOGLIKELIHOOD AT $\gamma = 0$

Calculate $\left. \frac{\partial \ell}{\partial \gamma} \right|_{\gamma=0}$ to determine if likelihood is decreasing at $\gamma = 0$; in which case, set

MLE $\hat{\pi}$ to that under the assumption of homogeneity – i.e. under the IMS model

$$\begin{aligned}
 \left. \frac{\partial \ell}{\partial \gamma} \right|_{\gamma=0} &= - \sum_{t=1}^T \sum_{j=1}^{n_t} (j-1) + \sum_{t=1}^T \sum_{i=1}^K \sum_{j=1}^{y_{ti}} \frac{j-1}{\pi_i} \\
 &= - \sum_{t=1}^T \left[\left(\sum_{j=1}^{n_t} j \right) - n_t \right] + \sum_{t=1}^T \sum_{i=1}^K \frac{1}{\pi_i} \left[\left(\sum_{j=1}^{y_{ti}} j \right) - y_{ti} \right] \quad (\text{II.1}) \\
 &= - \sum_{t=1}^T \frac{n_t (n_t - 1)}{2} + \sum_{t=1}^T \sum_{i=1}^K \frac{y_{ti} (y_{ti} - 1)}{2\pi_i} \quad \text{since } \sum_{j=1}^{n_t} j = \frac{n_t (n_t + 1)}{2}
 \end{aligned}$$

If this is negative at $\hat{\pi}_0$ then the likelihood is decreasing at $\gamma = 0$, hence set $\hat{\gamma} = 0$ and

$$\hat{\pi} = \hat{\pi}_0.$$

APPENDIX A.17

HETEROGENEITY MULTINOMIAL LOGIT MODEL DERIVATIVES

$$\ell(\gamma, \boldsymbol{\pi} | \mathbf{y}) \propto - \sum_{t=1}^T \sum_{j=1}^{n_t} \ln(1 + [j-1]\gamma) + \sum_{t=1}^T \sum_{i=1}^K \sum_{j=1}^{y_{it}} \ln(\pi_i + [j-1]\gamma)$$

$\frac{\partial \pi_i}{\partial \beta_m}$ is given in equation (I.7) and $\frac{\partial^2 \pi_i}{\partial \beta_i \partial \beta_m}$ is given in equation (I.8).

The first partial derivative of $\ell(\gamma, \boldsymbol{\pi} | \mathbf{y})$ with respect to γ and β_m are given below.

$$\frac{\partial \ell}{\partial \beta_m} = \sum_{t=1}^T \sum_{i=1}^{K-1} \sum_{j=1}^{y_{it}} \frac{\frac{\partial \pi_i}{\partial \beta_m}}{[\pi_i + (j-1)\gamma]} - \sum_{t=1}^T \sum_{j=1}^{y_{tK}} \frac{\sum_{i=1}^{K-1} \frac{\partial \pi_i}{\partial \beta_m}}{[\pi_K + (j-1)\gamma]}$$

$$\frac{\partial \ell}{\partial \gamma} = - \sum_{t=1}^T \sum_{j=1}^{n_t} \frac{(j-1)}{[1 + (j-1)\gamma]} + \sum_{t=1}^T \sum_{i=1}^K \sum_{j=1}^{y_{it}} \frac{(j-1)}{[\pi_i + (j-1)\gamma]}$$

The second partial derivatives of $\ell(\gamma, \boldsymbol{\pi} | \mathbf{y})$ are:

$$\frac{\partial^2 \ell}{\partial \beta_m \partial \beta_r} = - \sum_{t=1}^T \sum_{i=1}^{K-1} \sum_{j=1}^{y_{it}} \frac{\left(\frac{\partial \pi_i}{\partial \beta_r} \right) \left(\frac{\partial \pi_i}{\partial \beta_m} \right)}{[\pi_i + (j-1)\gamma]^2} - \sum_{t=1}^T \sum_{j=1}^{y_{tK}} \frac{\left(\sum_{i=1}^{K-1} \frac{\partial \pi_i}{\partial \beta_r} \right) \left(\sum_{i=1}^{K-1} \frac{\partial \pi_i}{\partial \beta_m} \right)}{[\pi_K + (j-1)\gamma]^2} + \sum_{t=1}^T \sum_{i=1}^{K-1} \sum_{j=1}^{y_{it}} \frac{\left(\frac{\partial^2 \pi_i}{\partial \beta_r \partial \beta_m} \right)}{[\pi_i + (j-1)\gamma]} - \sum_{t=1}^T \sum_{j=1}^{y_{tK}} \frac{\left(\sum_{i=1}^{K-1} \frac{\partial^2 \pi_i}{\partial \beta_r \partial \beta_m} \right)}{[\pi_K + (j-1)\gamma]}$$

$$\frac{\partial^2 \ell}{\partial \gamma \partial \beta_m} = - \sum_{t=1}^T \sum_{i=1}^{K-1} \sum_{j=1}^{y_{it}} \frac{(j-1)}{[\pi_i + (j-1)\gamma]^2} \frac{\partial \pi_i}{\partial \beta_m} + \sum_{t=1}^T \sum_{j=1}^{y_{tK}} \frac{(j-1)}{[\pi_K + (j-1)\gamma]^2} \left(\sum_{i=1}^{K-1} \frac{\partial \pi_i}{\partial \beta_m} \right)$$

$$\frac{\partial^2 \ell}{\partial \gamma^2} = \sum_{t=1}^T \sum_{j=1}^{n_t} \frac{(j-1)^2}{[1 + (j-1)\gamma]^2} - \sum_{t=1}^T \sum_{i=1}^K \sum_{j=1}^{y_{it}} \frac{(j-1)^2}{[\pi_i + (j-1)\gamma]^2} \quad (\text{II.2})$$

APPENDIX A.18

ASYMPTOTIC TESTS OF $H_0 : \gamma = 0$ vs $H_1 : \gamma > 0$

A.18.1 Asymptotic Likelihood Ratio Statistic and Test (LRT)

$$H_0 : \gamma = 0 \text{ vs } H_1 : \gamma > 0$$

The loglikelihood under the saturated heterogeneity model is:

$$\ell(\gamma, \boldsymbol{\pi} | \mathbf{y}) \propto -\sum_{t=1}^T \sum_{j=1}^{n_t} \ln(1 + [j-1]\gamma) + \sum_{t=1}^T \sum_{i=1}^K \sum_{j=1}^{y_{ti}} \ln(\pi_i + [j-1]\gamma),$$

where $\pi_K = 1 - \sum_{i=1}^{K-1} \pi_i$. Under H_0 , this loglikelihood becomes:

$$\ell(\gamma = 0, \boldsymbol{\pi} | \mathbf{y}) \propto \sum_{t=1}^T \sum_{i=1}^K y_{ti} \ln \pi_i,$$

which is identical to loglikelihood under the IMS model. The MLE's are obtained under both H_0 and H_1 , and the usual asymptotic LRT chi-squared statistic:

$$\text{LRT}_{\text{stat}} = -2(\ell_0(\gamma = 0, \hat{\boldsymbol{\pi}}_0 | \mathbf{y}) - \ell(\hat{\gamma}, \hat{\boldsymbol{\pi}} | \mathbf{y}))$$

is calculated. Under H_0 , LRT_{stat} is distributed as a chi-squared statistic on one degree of freedom, $\text{LRT}_{\text{stat}} \sim \chi_1^2$.

A.18.2 Score Statistic and Test

$$H_0 : \gamma = 0 \quad \text{vs} \quad H_1 : \gamma > 0$$

The first partial derivative of $\ell(\gamma, \boldsymbol{\pi} | \mathbf{y})$ with respect to γ at $\gamma = 0$ is:

$$\left. \frac{\partial \ell}{\partial \gamma} \right|_{\gamma=0} = - \sum_{t=1}^T \frac{n_t (n_t - 1)}{2} + \sum_{t=1}^T \sum_{i=1}^K \frac{y_{ti} (y_{ti} - 1)}{2\pi_i} \quad \text{by equation (II.1)}$$

The negative second partial derivatives of $\ell(\gamma, \boldsymbol{\pi} | \mathbf{y})$ at $\gamma = 0$ are:

$$- \left. \frac{\partial^2 \ell}{\partial \pi_i^2} \right|_{\gamma=0} = \pi_i^{-2} y_{\cdot i} + \pi_k^{-2} y_{\cdot k}$$

$$- \left. \frac{\partial^2 \ell}{\partial \pi_i \partial \pi_j} \right|_{\gamma=0} = \pi_k^{-2} y_{\cdot k} \quad \forall i \neq j$$

$$- \left. \frac{\partial^2 \ell}{\partial \pi_i \partial \gamma} \right|_{\gamma=0} = \sum_{t=1}^T \frac{(y_{ti} - 1) y_{ti}}{2\pi_i^2} - \sum_{t=1}^T \frac{(y_{tK} - 1) y_{tK}}{2\pi_K^2}$$

$$- \left. \frac{\partial^2 \ell}{\partial \gamma^2} \right|_{\gamma=0} = - \sum_{t=1}^T \frac{(n_t - 1) n_t (2n_t - 1)}{6} + \sum_{t=1}^T \sum_{i=1}^K \frac{(y_{ti} - 1) y_{ti} (2y_{ti} - 1)}{6\pi_i^2} \quad \text{by equation (II.2)}$$

Consider the score, $U(\gamma = 1, \hat{\boldsymbol{\pi}}_0) = \frac{\partial \ell(\gamma, \boldsymbol{\pi} | X)}{\partial \gamma} \Big|_{\gamma=0, \boldsymbol{\pi}=\hat{\boldsymbol{\pi}}_0}$ and the observed information

matrix $J(\gamma = 0, \hat{\boldsymbol{\pi}}_0) =$

$$\begin{bmatrix} -\frac{\partial^2 \ell}{\partial \pi_1^2} \Big|_{\gamma=0, \hat{\boldsymbol{\pi}}_0} & -\frac{\partial^2 \ell}{\partial \pi_1 \partial \pi_2} \Big|_{\gamma=0, \hat{\boldsymbol{\pi}}_0} & \dots & -\frac{\partial^2 \ell}{\partial \pi_1 \partial \pi_{K-1}} \Big|_{\gamma=0, \hat{\boldsymbol{\pi}}_0} & -\frac{\partial^2 \ell}{\partial \pi_1 \partial \gamma} \Big|_{\gamma=0, \hat{\boldsymbol{\pi}}_0} \\ -\frac{\partial^2 \ell}{\partial \pi_1 \partial \pi_2} \Big|_{\gamma=0, \hat{\boldsymbol{\pi}}_0} & -\frac{\partial^2 \ell}{\partial \pi_2^2} \Big|_{\gamma=0, \hat{\boldsymbol{\pi}}_0} & & \vdots & -\frac{\partial^2 \ell}{\partial \pi_2 \partial \gamma} \Big|_{\gamma=0, \hat{\boldsymbol{\pi}}_0} \\ \vdots & & \ddots & & \vdots \\ -\frac{\partial^2 \ell}{\partial \pi_1 \partial \pi_{K-1}} \Big|_{\gamma=0, \hat{\boldsymbol{\pi}}_0} & \dots & & -\frac{\partial^2 \ell}{\partial \pi_{K-1}^2} \Big|_{\gamma=0, \hat{\boldsymbol{\pi}}_0} & -\frac{\partial^2 \ell}{\partial \pi_{K-1} \partial \gamma} \Big|_{\gamma=0, \hat{\boldsymbol{\pi}}_0} \\ -\frac{\partial^2 \ell}{\partial \pi_1 \partial \gamma} \Big|_{\gamma=0, \hat{\boldsymbol{\pi}}_0} & -\frac{\partial^2 \ell}{\partial \pi_2 \partial \gamma} \Big|_{\gamma=0, \hat{\boldsymbol{\pi}}_0} & & -\frac{\partial^2 \ell}{\partial \pi_{K-1} \partial \gamma} \Big|_{\gamma=0, \hat{\boldsymbol{\pi}}_0} & -\frac{\partial^2 \ell}{\partial \gamma^2} \Big|_{\gamma=0, \hat{\boldsymbol{\pi}}_0} \end{bmatrix}$$

where $\hat{\boldsymbol{\pi}}_0$ is the MLE for $\boldsymbol{\pi}$ under H_0 .

The score statistic, $\text{Score}_{\text{stat}} = U(\gamma = 0, \hat{\boldsymbol{\pi}}_0)^2 * J^{-1}_{K,K}$ has a chi-squared distribution on one degree of freedom under H_0 , $\text{Score}_{\text{stat}} \sim \chi_1^2$. $J^{-1}_{K,K}$ is the K^{th} diagonal entry of the inverse of $J(\gamma = 0, \hat{\boldsymbol{\pi}}_0)$.

A.18.3 Wald Statistic and Test

$\text{Wald}_{\text{stat}} = \frac{(\hat{\gamma} - 0)^2}{J^{-1}_{K,K}}$ has a chi-squared distribution on one degree of freedom under H_0 ,

$\text{Wald}_{\text{stat}} \sim \chi_1^2$. Here $\hat{\gamma}$ is the MLE for γ under the heterogeneity model.

APPENDIX A.19

TABLES OF SIZE FOR $H_0 : \gamma = 0$ VS $H_1 : \gamma > 0$

Level $\alpha = 0.10$ test.				Size x 100		
T	n_t	Pi	SS	score	Wald	LRT
50	50	.25, .25, .25, .25	2000	1.85	6.55	3.20
50	30	.25, .25, .25, .25	2000	2.00	5.45	2.95
50	10	.25, .25, .25, .25	2000	1.75	6.80	2.95
30	50	.25, .25, .25, .25	2000	1.40	5.95	3.00
30	30	.25, .25, .25, .25	2000	1.65	5.95	2.90
30	10	.25, .25, .25, .25	2000	1.35	6.30	2.65
10	50	.25, .25, .25, .25	2000	0.65	6.25	2.15
10	30	.25, .25, .25, .25	2000	0.70	6.35	1.90
10	10	.25, .25, .25, .25	2000	0.40	5.95	1.65
50	50	.1, .2, .3, .4	2000	2.10	6.70	3.85
50	30	.1, .2, .3, .4	2000	2.15	6.60	3.55
50	10	.1, .2, .3, .4	2000	2.05	6.70	3.55
30	50	.1, .2, .3, .4	2000	1.35	6.20	2.85
30	30	.1, .2, .3, .4	2000	1.40	7.10	3.35
30	10	.1, .2, .3, .4	2000	1.00	6.25	2.50
10	50	.1, .2, .3, .4	2000	0.45	5.55	1.75
10	30	.1, .2, .3, .4	2000	0.90	5.65	2.25
10	10	.1, .2, .3, .4	2000	0.35	6.95	2.05
50	50	.15, .15, .35, .35	2000	2.45	7.00	3.75
50	30	.15, .15, .35, .35	2000	1.85	6.15	3.35
50	10	.15, .15, .35, .35	2000	1.80	6.50	3.35
30	50	.15, .15, .35, .35	2000	1.70	7.15	3.50
30	30	.15, .15, .35, .35	2000	1.75	6.35	3.10
30	10	.15, .15, .35, .35	2000	1.20	6.85	2.75
10	50	.15, .15, .35, .35	2000	0.75	6.20	2.35
10	30	.15, .15, .35, .35	2000	0.45	6.00	1.70
10	10	.15, .15, .35, .35	2000	0.55	6.05	2.00
50	50	.2, .2, .2, .4	2000	1.50	5.95	2.65
50	30	.2, .2, .2, .4	2000	2.25	5.45	3.50
50	10	.2, .2, .2, .4	2000	1.90	6.90	3.10
30	50	.2, .2, .2, .4	2000	1.70	6.60	3.10
30	30	.2, .2, .2, .4	2000	1.05	6.10	2.20
30	10	.2, .2, .2, .4	2000	1.50	7.30	3.05
10	50	.2, .2, .2, .4	2000	0.45	5.10	1.55
10	30	.2, .2, .2, .4	2000	0.45	6.00	1.80
10	10	.2, .2, .2, .4	2000	0.85	6.75	2.90

Level $\alpha = 0.05$ test.				Size x 100		
T	n_t	Pi	SS	score	Wald	LRT
50	50	.25, .25, .25, .25	2000	0.55	4.05	1.45
50	30	.25, .25, .25, .25	2000	0.80	3.60	1.65
50	10	.25, .25, .25, .25	2000	0.50	4.20	1.60
30	50	.25, .25, .25, .25	2000	0.40	4.05	1.30
30	30	.25, .25, .25, .25	2000	0.55	3.75	1.60
30	10	.25, .25, .25, .25	2000	0.15	4.10	1.35
10	50	.25, .25, .25, .25	2000	0.15	4.50	0.80
10	30	.25, .25, .25, .25	2000	0.15	4.40	1.00
10	10	.25, .25, .25, .25	2000	0.30	4.25	0.70
50	50	.1, .2, .3, .4	2000	0.65	4.70	1.65
50	30	.1, .2, .3, .4	2000	0.45	4.50	1.65
50	10	.1, .2, .3, .4	2000	0.65	4.40	1.85
30	50	.1, .2, .3, .4	2000	0.45	3.95	1.30
30	30	.1, .2, .3, .4	2000	0.35	4.40	1.30
30	10	.1, .2, .3, .4	2000	0.30	3.85	1.10
10	50	.1, .2, .3, .4	2000	0.15	3.85	0.70
10	30	.1, .2, .3, .4	2000	0.20	4.30	1.10
10	10	.1, .2, .3, .4	2000	0.05	5.10	0.80
50	50	.15, .15, .35, .35	2000	1.00	4.35	1.95
50	30	.15, .15, .35, .35	2000	0.60	4.00	1.35
50	10	.15, .15, .35, .35	2000	0.60	4.20	1.30
30	50	.15, .15, .35, .35	2000	0.55	4.80	1.60
30	30	.15, .15, .35, .35	2000	0.45	4.20	1.55
30	10	.15, .15, .35, .35	2000	0.40	4.45	1.15
10	50	.15, .15, .35, .35	2000	0.25	4.85	1.15
10	30	.15, .15, .35, .35	2000	0.15	4.20	0.75
10	10	.15, .15, .35, .35	2000	0.15	4.65	0.85
50	50	.2, .2, .2, .4	2000	0.50	3.30	1.20
50	30	.2, .2, .2, .4	2000	0.85	4.00	1.75
50	10	.2, .2, .2, .4	2000	0.60	4.05	1.45
30	50	.2, .2, .2, .4	2000	0.60	4.30	1.45
30	30	.2, .2, .2, .4	2000	0.60	3.95	1.00
30	10	.2, .2, .2, .4	2000	0.35	4.65	1.50
10	50	.2, .2, .2, .4	2000	0.15	3.50	0.80
10	30	.2, .2, .2, .4	2000	0.05	4.05	0.60
10	10	.2, .2, .2, .4	2000	0.15	5.30	1.45

Level $\alpha = 0.01$ test.				Size x 100		
T	n_t	Pi	SS	score	Wald	LRT
50	50	.25, .25, .25, .25	2000	0.00	1.55	0.25
50	30	.25, .25, .25, .25	2000	0.05	1.70	0.25
50	10	.25, .25, .25, .25	2000	0.00	1.65	0.20
30	50	.25, .25, .25, .25	2000	0.00	1.75	0.25
30	30	.25, .25, .25, .25	2000	0.05	2.00	0.25
30	10	.25, .25, .25, .25	2000	0.00	1.75	0.00
10	50	.25, .25, .25, .25	2000	0.00	2.75	0.15
10	30	.25, .25, .25, .25	2000	0.00	2.15	0.15
10	10	.25, .25, .25, .25	2000	0.00	2.25	0.30
50	50	.1, .2, .3, .4	2000	0.05	1.90	0.25
50	30	.1, .2, .3, .4	2000	0.10	1.90	0.30
50	10	.1, .2, .3, .4	2000	0.05	2.10	0.35
30	50	.1, .2, .3, .4	2000	0.05	1.70	0.15
30	30	.1, .2, .3, .4	2000	0.00	2.25	0.25
30	10	.1, .2, .3, .4	2000	0.05	1.95	0.25
10	50	.1, .2, .3, .4	2000	0.00	2.05	0.15
10	30	.1, .2, .3, .4	2000	0.00	2.40	0.25
10	10	.1, .2, .3, .4	2000	0.00	2.85	0.05
50	50	.15, .15, .35, .35	2000	0.00	1.95	0.50
50	30	.15, .15, .35, .35	2000	0.00	1.45	0.20
50	10	.15, .15, .35, .35	2000	0.05	1.75	0.25
30	50	.15, .15, .35, .35	2000	0.05	2.30	0.30
30	30	.15, .15, .35, .35	2000	0.05	2.00	0.25
30	10	.15, .15, .35, .35	2000	0.00	2.00	0.30
10	50	.15, .15, .35, .35	2000	0.00	2.45	0.35
10	30	.15, .15, .35, .35	2000	0.00	2.05	0.15
10	10	.15, .15, .35, .35	2000	0.00	2.60	0.15
50	50	.2, .2, .2, .4	2000	0.00	1.20	0.10
50	30	.2, .2, .2, .4	2000	0.15	1.80	0.30
50	10	.2, .2, .2, .4	2000	0.10	1.80	0.35
30	50	.2, .2, .2, .4	2000	0.00	2.20	0.45
30	30	.2, .2, .2, .4	2000	0.05	1.25	0.50
30	10	.2, .2, .2, .4	2000	0.00	2.15	0.35
10	50	.2, .2, .2, .4	2000	0.05	1.90	0.15
10	30	.2, .2, .2, .4	2000	0.00	2.05	0.15
10	10	.2, .2, .2, .4	2000	0.00	3.40	0.25

APPENDIX A.20

TABLES OF POWER FOR $H_0 : \gamma = 0$ VS $H_1 : \gamma > 0$

K = 3				Power x 100		
T	n_t	π	γ	Score	Wald	LRT
50	50	1/3, 1/3, 1/3	1/3	100.00	100.00	100.00
50	30	1/3, 1/3, 1/3	1/3	100.00	100.00	100.00
50	10	1/3, 1/3, 1/3	1/3	100.00	100.00	100.00
30	50	1/3, 1/3, 1/3	1/3	100.00	100.00	100.00
30	30	1/3, 1/3, 1/3	1/3	100.00	100.00	100.00
30	10	1/3, 1/3, 1/3	1/3	100.00	100.00	100.00
10	50	1/3, 1/3, 1/3	1/3	100.00	100.00	100.00
10	30	1/3, 1/3, 1/3	1/3	100.00	100.00	99.90
10	10	1/3, 1/3, 1/3	1/3	87.89	98.90	95.30
50	50	1/3, 1/3, 1/3	1/15	100.00	100.00	100.00
50	30	1/3, 1/3, 1/3	1/15	100.00	100.00	100.00
50	10	1/3, 1/3, 1/3	1/15	82.90	95.20	89.70
30	50	1/3, 1/3, 1/3	1/15	100.00	100.00	100.00
30	30	1/3, 1/3, 1/3	1/15	100.00	100.00	100.00
30	10	1/3, 1/3, 1/3	1/15	54.40	84.10	71.30
10	50	1/3, 1/3, 1/3	1/15	92.40	99.00	97.00
10	30	1/3, 1/3, 1/3	1/15	69.30	93.00	83.90
10	10	1/3, 1/3, 1/3	1/15	8.40	47.80	26.10
50	50	1/3, 1/3, 1/3	1/30	100.00	100.00	100.00
50	30	1/3, 1/3, 1/3	1/30	98.40	99.70	99.10
50	10	1/3, 1/3, 1/3	1/30	31.10	62.60	45.3
30	50	1/3, 1/3, 1/3	1/30	99.50	100.00	99.90
30	30	1/3, 1/3, 1/3	1/30	87.70	97.50	93.80
30	10	1/3, 1/3, 1/3	1/30	14.30	46.60	28.80
10	50	1/3, 1/3, 1/3	1/30	58.70	88.00	76.90
10	30	1/3, 1/3, 1/3	1/30	26.00	66.70	44.40
10	10	1/3, 1/3, 1/3	1/30	1.50	24.00	10.40
50	50	1/3, 1/3, 1/3	1/75	85.60	95.10	90.70
50	30	1/3, 1/3, 1/3	1/75	46.50	74.80	59.70
50	10	1/3, 1/3, 1/3	1/75	5.10	21.20	10.30
30	50	1/3, 1/3, 1/3	1/75	64.30	87.60	76.80
30	30	1/3, 1/3, 1/3	1/75	25.20	56.60	37.60
30	10	1/3, 1/3, 1/3	1/75	2.60	16.20	7.40
10	50	1/3, 1/3, 1/3	1/75	9.90	50.80	28.00
10	30	1/3, 1/3, 1/3	1/75	4.20	31.40	13.70
10	10	1/3, 1/3, 1/3	1/75	0.30	12.50	3.80

K = 4				Power x 100		
T	n_t	π	γ	Score	Wald	LRT
50	50	.25, .25, .25, .25	1/4	100.00	100.00	100.00
50	30	.25, .25, .25, .25	1/4	100.00	100.00	100.00
50	10	.25, .25, .25, .25	1/4	100.00	100.00	100.00
30	50	.25, .25, .25, .25	1/4	100.00	100.00	100.00
30	30	.25, .25, .25, .25	1/4	100.00	100.00	100.00
30	10	.25, .25, .25, .25	1/4	100.00	100.00	100.00
10	50	.25, .25, .25, .25	1/4	100.00	100.00	100.00
10	30	.25, .25, .25, .25	1/4	100.00	100.00	100.00
10	10	.25, .25, .25, .25	1/4	91.50	99.20	96.50
50	50	.25, .25, .25, .25	1/20	100.00	100.00	100.00
50	30	.25, .25, .25, .25	1/20	100.00	100.00	100.00
50	10	.25, .25, .25, .25	1/20	78.30	92.60	86.20
30	50	.25, .25, .25, .25	1/20	100.00	100.00	100.00
30	30	.25, .25, .25, .25	1/20	100.00	100.00	100.00
30	10	.25, .25, .25, .25	1/20	50.50	81.50	66.30
10	50	.25, .25, .25, .25	1/20	95.80	99.50	98.20
10	30	.25, .25, .25, .25	1/20	67.90	93.30	83.70
10	10	.25, .25, .25, .25	1/20	6.30	41.60	20.70
50	50	.25, .25, .25, .25	1/40	100.00	100.00	100.00
50	30	.25, .25, .25, .25	1/40	99.20	99.70	99.60
50	10	.25, .25, .25, .25	1/40	26.80	56.20	39.60
30	50	.25, .25, .25, .25	1/40	99.90	100.00	100.00
30	30	.25, .25, .25, .25	1/40	86.60	96.90	91.60
30	10	.25, .25, .25, .25	1/40	13.20	38.80	24.10
10	50	.25, .25, .25, .25	1/40	58.00	88.60	76.70
10	30	.25, .25, .25, .25	1/40	20.80	62.70	40.30
10	10	.25, .25, .25, .25	1/40	1.40	18.90	6.70
50	50	.25, .25, .25, .25	1/100	86.30	96.40	91.70
50	30	.25, .25, .25, .25	1/100	45.10	67.10	53.40
50	10	.25, .25, .25, .25	1/100	5.40	19.40	8.70
30	50	.25, .25, .25, .25	1/100	59.10	82.40	69.90
30	30	.25, .25, .25, .25	1/100	22.20	49.50	35.00
30	10	.25, .25, .25, .25	1/100	2.40	16.10	6.00
10	50	.25, .25, .25, .25	1/100	9.90	45.90	23.50
10	30	.25, .25, .25, .25	1/100	3.20	23.20	9.30
10	10	.25, .25, .25, .25	1/100	0.10	8.30	2.20

K = 5				Power x 100		
T	n_t	π	γ	Score	Wald	LRT
50	50	.2, .2, .2, .2, .2	1/5	100.00	100.00	100.00
50	30	.2, .2, .2, .2, .2	1/5	100.00	100.00	100.00
50	10	.2, .2, .2, .2, .2	1/5	100.00	100.00	100.00
30	50	.2, .2, .2, .2, .2	1/5	100.00	100.00	100.00
30	30	.2, .2, .2, .2, .2	1/5	100.00	100.00	100.00
30	10	.2, .2, .2, .2, .2	1/5	100.00	100.00	100.00
10	50	.2, .2, .2, .2, .2	1/5	100.00	100.00	100.00
10	30	.2, .2, .2, .2, .2	1/5	100.00	100.00	100.00
10	10	.2, .2, .2, .2, .2	1/5	89.70	98.80	95.60
50	50	.2, .2, .2, .2, .2	1/25	100.00	100.00	100.00
50	30	.2, .2, .2, .2, .2	1/25	100.00	100.00	100.00
50	10	.2, .2, .2, .2, .2	1/25	75.80	90.70	84.30
30	50	.2, .2, .2, .2, .2	1/25	100.00	100.00	100.00
30	30	.2, .2, .2, .2, .2	1/25	100.00	100.00	100.00
30	10	.2, .2, .2, .2, .2	1/25	7.80	76.00	61.40
10	50	.2, .2, .2, .2, .2	1/25	95.70	99.70	98.60
10	30	.2, .2, .2, .2, .2	1/25	69.20	92.90	84.70
10	10	.2, .2, .2, .2, .2	1/25	7.70	35.20	18.50
50	50	.2, .2, .2, .2, .2	1/50	100.00	100.00	100.00
50	30	.2, .2, .2, .2, .2	1/50	98.50	99.30	98.90
50	10	.2, .2, .2, .2, .2	1/50	26.70	47.70	35.70
30	50	.2, .2, .2, .2, .2	1/50	99.50	99.90	99.70
30	30	.2, .2, .2, .2, .2	1/50	83.80	94.80	89.90
30	10	.2, .2, .2, .2, .2	1/50	9.10	33.20	18.20
10	50	.2, .2, .2, .2, .2	1/50	56.60	87.30	73.80
10	30	.2, .2, .2, .2, .2	1/50	20.40	59.80	37.70
10	10	.2, .2, .2, .2, .2	1/50	0.80	13.70	4.50
50	50	.2, .2, .2, .2, .2	1/125	83.20	94.40	88.50
50	30	.2, .2, .2, .2, .2	1/125	39.30	64.80	50.30
50	10	.2, .2, .2, .2, .2	1/125	3.80	14.50	6.80
30	50	.2, .2, .2, .2, .2	1/125	51.90	79.30	65.60
30	30	.2, .2, .2, .2, .2	1/125	18.30	44.60	28.70
30	10	.2, .2, .2, .2, .2	1/125	1.60	10.50	4.00
10	50	.2, .2, .2, .2, .2	1/125	8.20	36.80	21.00
10	30	.2, .2, .2, .2, .2	1/125	3.10	19.70	8.70
10	10	.2, .2, .2, .2, .2	1/125	0.10	5.80	1.50

APPENDIX A.21

TABLES OF POWER AND LOF FOR $H_0 : \beta_1 = 0$ VS $H_1 : \beta_1 \neq 0$

T	n_t	π	γ	Power x 100		LOF	
				Het.	IMS	Het.	IMS
50	50	.192, .231, .269, .308	10/52	88.25	98.80	0.85	59.00
50	30	.192, .231, .269, .308	10/52	85.25	97.60	0.65	43.25
50	10	.192, .231, .269, .308	10/52	72.00	88.10	1.25	13.65
50	50	.192, .231, .269, .308	1/26	100.00	100.00	1.05	17.90
50	30	.192, .231, .269, .308	1/26	99.80	99.95	0.60	10.35
50	10	.192, .231, .269, .308	1/26	91.50	95.00	1.15	2.60
50	50	.192, .231, .269, .308	1/52	100.00	100.00	1.45	8.60
50	30	.192, .231, .269, .308	1/52	100.00	100.00	0.95	4.50
50	10	.192, .231, .269, .308	1/52	93.95	95.35	0.45	1.20
30	50	.192, .231, .269, .308	10/52	68.55	94.35	0.90	59.45
30	30	.192, .231, .269, .308	10/52	63.40	90.60	0.85	45.50
30	10	.192, .231, .269, .308	10/52	52.55	75.20	0.65	13.60
30	50	.192, .231, .269, .308	1/26	98.70	99.85	0.85	18.05
30	30	.192, .231, .269, .308	1/26	96.20	99.25	1.35	10.70
30	10	.192, .231, .269, .308	1/26	72.25	81.20	0.75	2.45
30	50	.192, .231, .269, .308	1/52	99.80	99.90	0.65	8.40
30	30	.192, .231, .269, .308	1/52	98.35	99.75	1.20	4.50
30	10	.192, .231, .269, .308	1/52	78.30	83.00	1.15	1.75

T	n_t	π	γ	Power x 100		LOF	
				Het.	IMS	Het.	IMS
50	50	.143, .171, .200, .229, .257	10/70	98.50	99.85	2.45	73.65
50	30	.143, .171, .200, .229, .257	10/70	97.65	99.80	2.55	57.45
50	10	.143, .171, .200, .229, .257	10/70	89.85	95.95	2.85	22.15
50	50	.143, .171, .200, .229, .257	1/35	100.00	100.00	2.60	26.05
50	30	.143, .171, .200, .229, .257	1/35	99.80	99.95	2.80	15.40
50	10	.143, .171, .200, .229, .257	1/35	98.60	99.20	2.60	6.20
50	50	.143, .171, .200, .229, .257	1/70	100.00	100.00	1.80	12.30
50	30	.143, .171, .200, .229, .257	1/70	100.00	100.00	2.10	7.60
50	10	.143, .171, .200, .229, .257	1/70	99.15	99.30	3.05	4.65
30	50	.143, .171, .200, .229, .257	10/70	89.05	98.40	3.25	72.30
30	30	.143, .171, .200, .229, .257	10/70	86.75	97.45	2.30	57.90
30	50	.143, .171, .200, .229, .257	1/70	100.00	100.00	2.50	12.20
30	30	.143, .171, .200, .229, .257	1/70	100.00	100.00	2.25	7.45
30	50	.143, .171, .200, .229, .257	1/35	99.95	100.00	2.95	26.65
30	30	.143, .171, .200, .229, .257	1/35	99.65	100.00	2.45	14.65