MANUAL ON
GOOD PRACTICES IN EXTENSION RESEARCH & EVALUATION

Regression Data Interpretation
Qualitative Extension
Methods Thematic Analysis Designs
Standard Deviation Descriptive Validity
Learning Management Reliability Tools
Ethics Research Margin of Error
Journals Diffusion
Modeling Monitoring Constructs
Policy Implications Impact Assessment
Experimental Variance Evaluation
Case Study Variables
Inference
Replication Psychometrics

नार्म naarm CRISP aesa Sharing, learning and networking for innovation Agricultural Extension in South Asia CTCRI MANAGE मैनेज

PUBLISHED BY

Agricultural Extension in South Asia (AESA)

www.aesa-gfras.net

C/o Centre for Research on Innovation and Science Policy (CRISP)
Road No. 10, Banjara Hills, Hyderabad-500 034, INDIA
www.crispindia.org

In collaboration with

Manual on Good Practices in Extension Research & Evaluation

Editors: P Sethuraman Sivakumar, B S Sontakki, Rasheed Sulaiman V, Saravanan R and Nimisha Mittal

P Sethuraman Sivakumar, Senior Scientist, ICAR-Central Tuber Crops Research Institute (ICAR-CTCRI), Thiruvananthapuram, India http://www.ctcri.org

B S Sontakki, Head-EXM Division, ICAR-National Academy of Agricultural Research Management (ICAR-NAARM), Hyderabad, India https://naarm.org.in

Rasheed Sulaiman V, Director, Centre for Research on Innovation and Science Policy (CRISP), Hyderabad, India http://www.crispindia.org

Saravanan R, Director (Agricultural Extension), National Institute of Agricultural Extension Management (MANAGE), Hyderabad, India http://www.manage.gov.in/

Nimisha Mittal, Program Manager, Centre for Research on Innovation and Science Policy (CRISP), Hyderabad, India http://www.crispindia.org

February 2017

19

Modeling Adoption of Agricultural Technologies

Sivaramane N

Objectives

1. Introduce Modeling adoption of agricultural technologies
2. Explain the multinomial and ordinal logit models
3. Suggest future work on adoption modelling

19.1 Introduction

- Measuring agricultural technology adoption is a multi-faceted process involving several choices and decisions to be made based on the research questions, resource availability, expertise of the researchers, etc.
- In general, adoption is assessed at the individual, institutional and societal levels. Various theories and processes of measuring adoption are discussed elsewhere (Sivakumar and Sulaiman, 2015).
- The most popular method employed in adoption study is Logit model. Logit model is a limited dependent variable model where the dependent variable is binary. This model is used in identifying the factors influencing adoption. However, this model is not applicable in cases where the dependent variable has multiple options or where it is of ordinal type.
- In this module, two econometric approaches, namely, multinomial logit and ordinal logit model employed in modeling adoption, are discussed in detail.

19.2 Discussion

- Multinomial model is an extension of the binary logistic regression model where the dependent variable is nominal and has multiple options.
- A farmer has the option of selecting a particular variety of crop for this season, or he/she takes the decision to dispose farm produce in an accessible market. Here the dependent variable is discrete and one choice is selected ultimately by the farmer based on several factors taken up for study.
- In case the dependent variable of the regression model is a rank or ordinal variable, ordinal logit/probit model will be more appropriate. For instance, the category of adoption of a particular technology (early adopters, late adopters and non-adopters) as a function of socio-economic variable falls under ordinal logit/probit category. The models can be selected based on the nature of dependent variable (Box. 19.1).

Box 19.1:Suitability of econometric models for adoption

Adoption as categorical dependent variable

- Response categories – Adopted/Non-adopted; Adopted/Partially Adopted; Choice of variety;
- While logit model is a non-parametric model, it enjoys wider acceptability in social science research;
- Probit which is based on normal distribution is widely used for experimental data.

Adoption as a composite dependent variable

- Composite variable – Adoption Index
- Censored regression model, such as Tobit, can be used

A. Multinomial logit model

This model is similar to logit model except that the dependent variable has more categories. In this case, the independent variable adoption has qualitative categories which are not ordered. Examples are:

- Choice of varieties sown;
- Selection of brands of pesticides, fertilizers, etc.;
- Choice of occupations;
- Choice of brands of soap, television, newspaper, etc.;
- Choice of color of shirt, bag, cycle, vehicles, etc., within the brands.

Logit model generally requires more samples or data points unlike linear regression model. Multinomial logit model also requires a relatively large data set. The multinomial model generates j-1 sets of parameter estimates, comparing different levels of the dependent variable to a reference level. This makes the model considerably more complex, but also much more flexible. The general form of the model is

$$P\left(Y_i = \frac{1}{X_i}\right) = \frac{1}{1+\sum_{j=2}^{J} exp(X_i\beta_j)} \text{ for m=1}$$
$$P\left(Y_i = \frac{m}{X_i}\right) = \frac{exp(X_i\beta_j)}{1+\sum_{j=2}^{J} exp(X_i\beta_j)} \text{ for m>1}$$

1. Ordinal logistic regression

The ordinal logit model is very similar to the multinomial model except that the values or classes are ordered. Variables containing scores, ranks or ratings can be used in ordinal logit model. Examples are:

- Different levels of adoption of varieties – fully adopted, partially adopted, non-adoption;
- Rating of nine varieties of rice crop based on its adaptation to climate change.

The ratings in ordinal logit are modeled as

Rating=1, log $\frac{P_1}{P_2+P_3+\cdots+P_9}$

Rating=1 or 2, log $\frac{P_1+P_2}{P_3+\cdots+P_9}$ ......

Rating=1 or 2 or .... or 8, log $\frac{P_1+P_2+\cdots+P_8}{P_9}$

For example: The following data set containing the preference ratings of varieties of rice (Table 19.1) chosen in case of an extreme event (drought).

Table 19.1: Varieties of rice crops suitable for climate change; An option survey of farmers' selected particular option

	Ratings 1	Ratings 2	Ratings 3	Ratings 4	Ratings 5	Ratings 6	Ratings 7	Ratings 8	Ratings 9
Variety 1	0	0	1	7	8	8	19	8	1
Variety 2	6	9	12	11	7	6	1	0	0
Variety 3	1	1	6	8	23	7	5	1	0
Variety 4	0	0	0	1	3	7	14	16	11

Note: The figures in the table indicate frequency of farmers' selected particular option which is used as input for fitting ordinal logistic model.

2. Interpretation of results

(i) Odds ratio

An odds ratio (OR) is a measure of association between an exposure and an outcome. The OR represents the odds that an outcome will occur given a particular exposure, compared to the odds of the outcome occurring in the absence of that exposure.

$$OR= \exp(b)$$

Where bi are estimates of logistic model. Odds ratio describes the odds of selecting an option over the reference option.

(ii) Pseudo Rsquare

This measure is used to compare among the models. SPSS outputs three Rsquare measures, viz., Cox and Snell, Nagelkerke, and McFadden. All these tests are used for comparisons, however, varying in the degree of rejection of null hypotheses.

(iii) Reference category

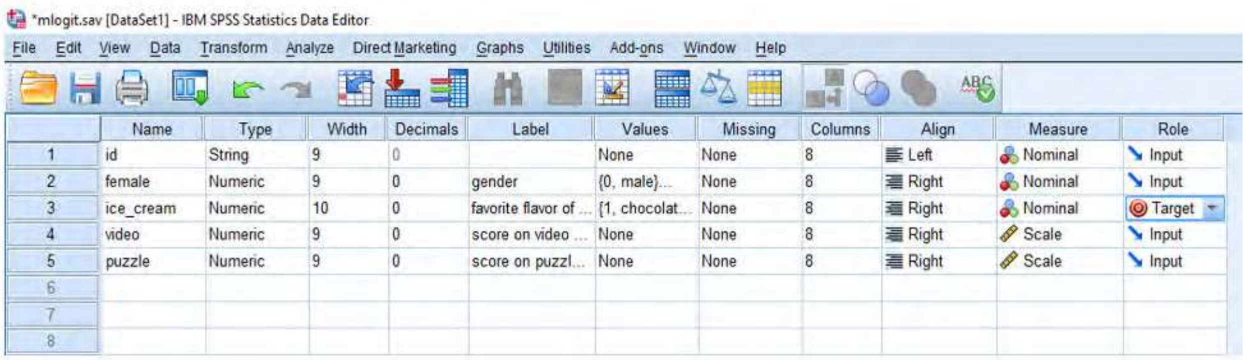
Reference category is used for comparison with other selected categories. Care should be taken to choose the reference category so that the logistic estimates are positive which helps in proper interpretation.

(iv) Receiver Operating Characteristic curve (ROC)

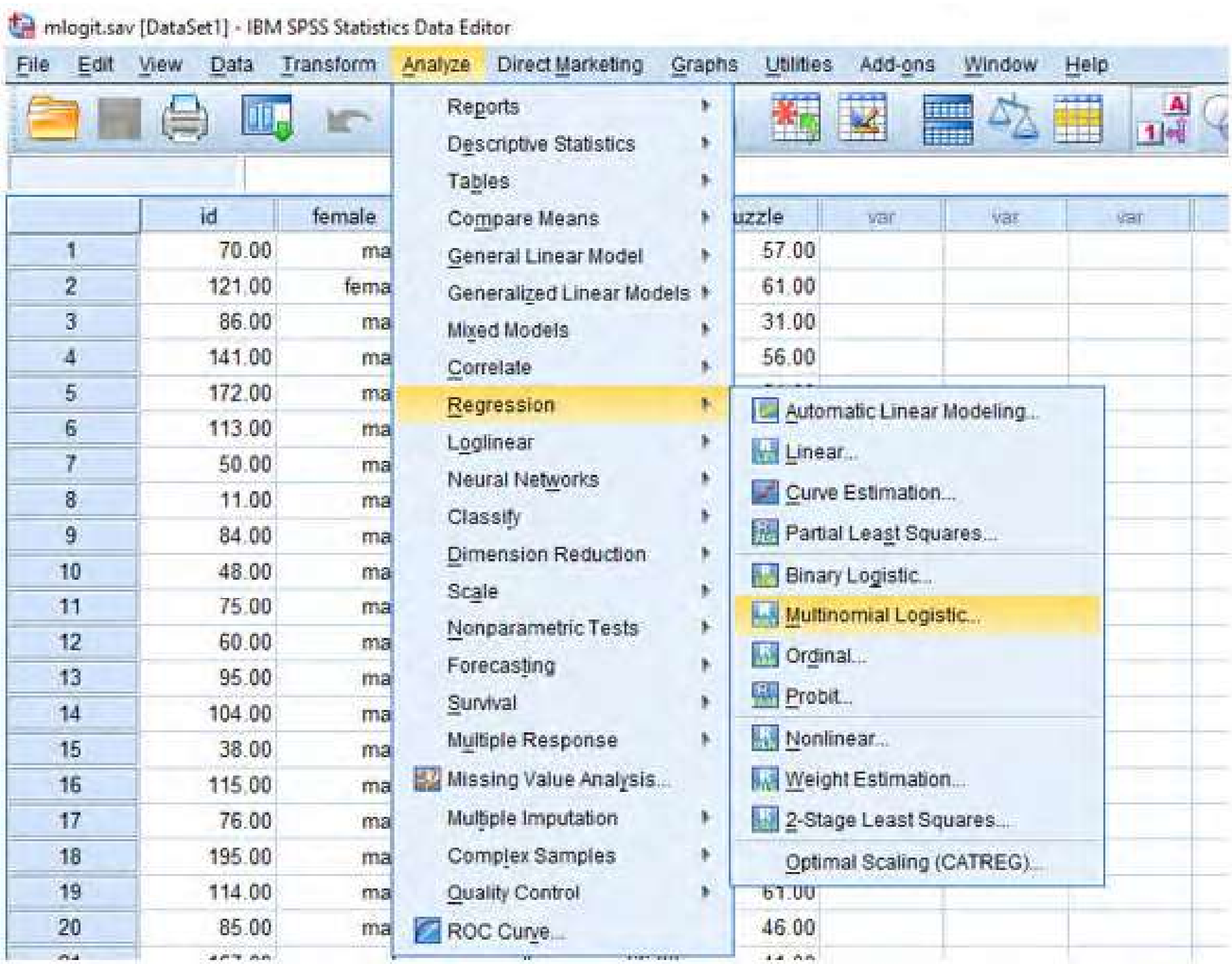
ROC curve is a graphical plot drawn by plotting the true positive rate (known as sensitivity) against the false positive rate (1 - specificity) at various threshold settings. This analysis provides tools to select possibly optimal models and to discard suboptimal ones.

B. Multinomial logit using SPSS

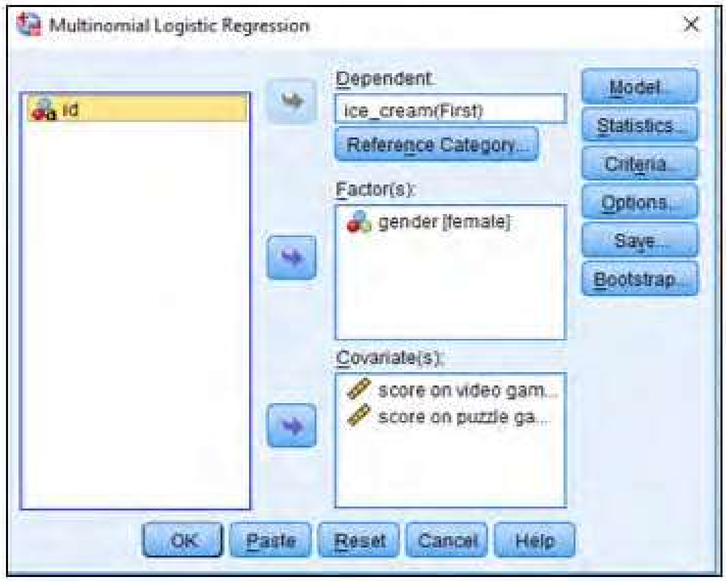
The model dataset mlogit.sav, which is provided along with SPSS software is used in this example. The same can be downloaded from University of California, Los Angeles website. This data set explains an experiment involving 200 high school students and their scores on various tests, including a video game and a puzzle. The outcome measure in this analysis is the students' favorite flavor of ice cream – vanilla, chocolate or strawberry. The screenshot below shows the type of variable used in multinomial logit model. The dependent variable (ice cream) is measured as nominal, and similarly the gender variable (FEMALE).



For employing multinomial logit model, select the appropriate option: ANALYZE> REGRESSION > MULTINOMIAL LOGISTIC … (refer to the screenshot below).



Ice cream is chosen as the dependent variable and chocolate is used as the reference category by specifying the option FIRST. Though it is not mandatory, the reference category is chosen in such a way that the odds-ratio is greater than one.



The output is given below:

Pseudo R-Square
Cox and Snell .153
Nagelkerke .174
McFadden .079

Parameter Estimates

Favorite flavor of ice cream[a]		**B**	**Std. Error**	**Wald**	**df**	**Sig.**	**Exp (B)**
Vanilla	Intercept						
	video	.024	.021	1.262	1	.261	1.024
	puzzle	.039	.020	3.978	1	.046	1.040
	[female=0]	.817	.391	4.362	1	.037	2.263
	[female=1]	0[b]	.	.	0	.	.
Strawberry	Intercept	-6.819	1.442	22.351	1	.000	
	video	.046	.025	3.430	1	.064	1.048
	puzzle	.082	.024	11.816	1	.001	1.085
	[female=0]	.849	.448	3.592	1	.058	2.338
	[female=1]	0[b]	.	.	0	.	.

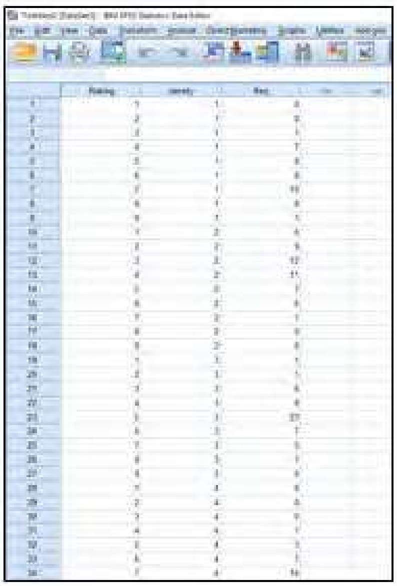
[a]The reference category is chocolate.
[b]This parameter is set to zero because it is redundant.

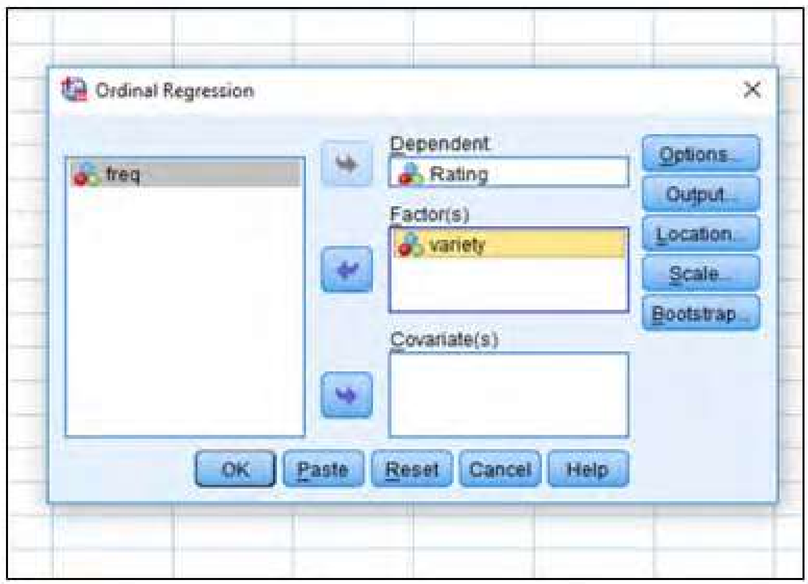
Pseudo R-square should be interpreted with caution and at least it should not be used as a measure of capture of variation in Y as the case in linear regression. It may be used only for comparison among models. The higher the pseudo R-square, the better the fit is. The third flavor, chocolate, is taken as the reference category by selecting the FIRST option in SPSS. See the dialog box where FIRST option is specified next to Ice cream under dependent variable. The parameter estimates showed that in vanilla equation, both the variables video and puzzle are insignificant. While in strawberry equation, the choice of the ice cream is dependent on video and puzzle scores. This reveals that as the score of video and puzzle increases, strawberry flavor is preferred over chocolate flavor. Exp(B) shows the odds-ratio of respective variables. Female=0 indicates category 'MALE'.

A detailed discussion on this example is given in IDRE-UCLA Webpage (http://www.ats.ucla.edu/stat/spss/output/mlogit.htm)

C. Ordinal logit using SPSS

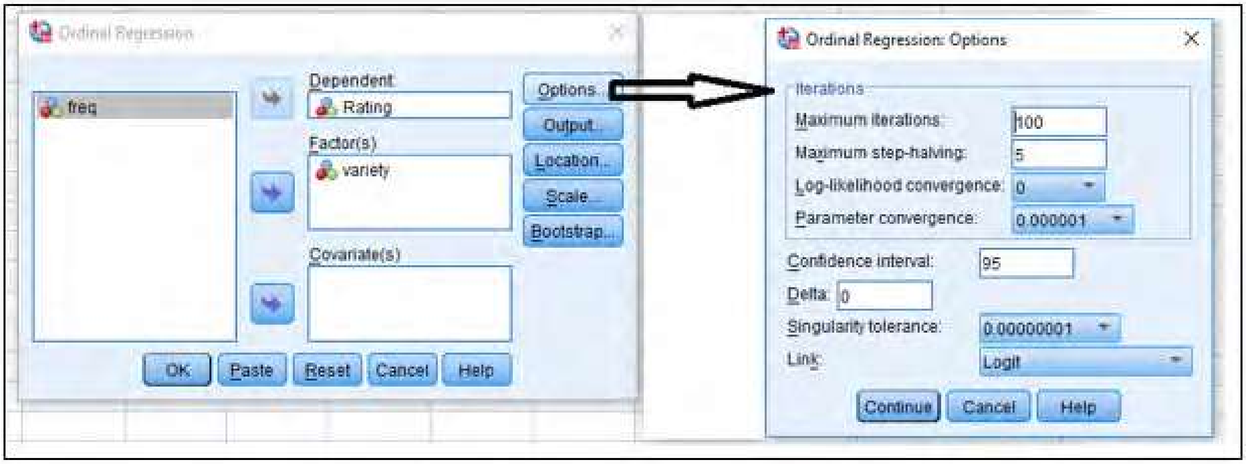
The model dataset orice.sav is used for this demonstration. The dataset is fed into SPSS in the format as shown in the screenshot. Then to call Ordinal Logit dialog box, click ANALYZE > REGRESSION > ORDINAL.





The variable VARIETY is nominal, hence included under FACTOR(S). Don't forget to select the option LOGIT link after clicking OPTIONS… as shown in the figure below.

		Estimate	Std. Error	Wald	df	Sig.
Threshold	[Rating = 1]	-7.080	.562	158.486	1	.000
	[Rating = 2]	-6.025	.475	160.551	1	.000
	[Rating = 3]	-4.925	.427	132.948	1	.000
	[Rating = 4]	-3.857	.390	97.709	1	.000
	[Rating = 5]	-2.521	.343	53.971	1	.000
	[Rating = 6]	-1.569	.309	25.838	1	.000
	[Rating = 7]	-.067	.266	.063	1	.801
	[Rating = 8]	1.493	.331	20.344	1	.000
Location	[variety=1]	-1.613	.378	18.226	1	.000
	[variety=2]	-4.965	.474	109.645	1	.000
	[variety=3]	-3.323	.425	61.094	1	.000
	[variety=4]	0[a]	.	.	0	.



The output of ordinal logistic model is given below.

Parameter estimates

Link function: **Logit.**

a. This parameter is set to zero because it is redundant.

Here the rating 9 is taken as the reference category, if rating=1 was the reference category, then the coefficients would have been positive and easy to interpret. There is no option of specifying the reference category in SPSS dialog box other than FIRST and LAST. It may either be done through syntax mode or by reversing the rating scale. Finally, the ordinal logit yields n-1 intercepts where n=number of ratings. The odds ratio may be computed by taking exponential of the estimate.

19.3 Key Points

- Measuring adoption of a technology is a multi-faceted process, which involves various methods of analysis and employing various analytical tools as envisaged by researchers/managers/extension functionaries.
- In advanced modelling, for identifying the factors influencing the adoption of a technology, models such as binary logistic modeling or logit, multinomial logit, and ordinal logit models are more pertinent.
- For the analysis where adoption variable is binary, say, adopted or not adopted, logit model is used. However, very little literature is available where advanced models such as multinomial or ordinal model are employed in adoption studies even though they are more pertinent. For instance, in cases where the dependent model is having more choices or categories, such as a few varietal options for farmers, choice from a few brands of pesticides available, etc., the multinomial model is appropriate.
- The multinomial model generates j-1 sets of parameter estimates, comparing different levels of the dependent variable to a reference level. This makes the model considerably more complex, but also much more flexible.
- The ordinal logit model is very much similar to the multinomial model except that the values or classes are ordered. Variables containing scores, ranks or ratings can be used in ordinal logit model.
- In this module, an illustration is provided where Rating of nine varieties of rice crop based on its adaptation to climate change were used to fit ordered logit model. Measures, such as Pseudo Rsquare and rationale in selecting the reference category, are discussed in this module along with reference about testing the fit using ROC curve.

- The screenshot of SPSS showing the steps involved in performing multinomial logit model and ordinal logit model is provided in the module.

References

- Agresti, A. (1996.) An introduction to categorical data analysis. New York, NY: John Wiley and Sons.
- Agresti, A. (2002.) Categorical data analysis. Second edition. New York, NY: John Wiley & Sons.
- Belsley, D. A. (1991.) Conditioning diagnostics: Collinearity and weak data in regression. New York, NY: Wiley.
- Burnham, K. P. and Anderson, D. R. (2002.) Model selection and multimodel inference. New York, NY: Springer.
- Doss C. (2006.) Analyzing technology adoption using micro studies: limitations, challenges, and opportunities for improvement. Agri. Econ. 34:207–219.
- Greene, W. (2000.) Econometric Analysis Fourth edition. New York, NY: Wiley.
- Hosmer, D. W. and Lemeshow, S. (2000.) Applied logistic regression. Second edition. New York, NY: John Wiley & Sons.
- Long, J. S. (1997.) Regression models of categorical and limited dependent variables. Thousand Oaks, CA: Sage.
- Madsen, M. (1976.) Statistical analysis of multiple contingency tables; Two examples. Scand. J. Statist. 3:97-106.
- SAS Institute, Inc. (2004.) SAS/STAT 9.1 user's guide. Cary, NC: SAS Institute Inc.
- Stokes, M. E., Davis, C. S., and Koch, G. G. (2000.) Categorical data analysis using the SAS system. Cary, NC: SAS Institute Inc.