Modeling Site Specific Heterogeneity in an On-Site Stratified Random Sample of Recreational Demand

Kavita Sardana[1]
John C. Bergstrom

*Selected Paper prepared for presentation at the Agricultural and Applied Economics Association's 2011 AAEA and NAREA Joint Annual Meeting in Pittsburgh,
July 24-26, 2011.*

[1]Kavita Sardana (`sardanak@iastate.edu`) is a Postdoctoral Research Associate, Department of Economics, Iowa State University, Ames, IA, and John C. Bergstrom (`jberg@uga.edu`) is a Professor, Department of Agricultural and Applied Economics, University of Georgia, Athens, GA.

Modeling Site Specific Heterogeneity in an On-Site Stratified Random Sample of Recreational Demand

Kavita Sardana John C. Bergstrom

Abstract

Using estimation of demand for the George Washington/Jefferson National Forest as a case study, it is shown that in a stratified/clustered on-site sample, latent heterogeneity needs to be accounted for twice: first to account for dispersion in the data caused by unobservability of the process that results in low and high frequency visitors in the population, and second to capture unobservable heterogeneity among individuals surveyed at different sites according to a stratified random sample (site specific effects). It is shown that both of the parameters capturing latent heterogeneity are statistically significant. It is therefore claimed in this paper, that the model accounting for site-specific effects is superior to the model without such effects. Goodness of fit statistics show that our empirical model is superior to models that do not account for latent heterogeneity for the second time. The price coefficient for the travel cost variable changes across model resulting in differences in consumer surplus measures. The expected mean also changes across different models. This information is of importance to the USDA Forest Service for the purpose of consumer surplus calculations and projections for budget allocation and resource utilization.

Introduction

In order to reduce survey costs, on-site survey samples are either clustered or stratified. Random samples are drawn within these clusters to make inferences about the relevant populations. According to Cameron and Trivedi (1986), survey data are usually dependent. This may be due to the use of cluster samples to reduce survey costs. In such cases the data may be correlated within a cluster owing to a presence of a common unobserved cluster-specific term. According to Pepper (2002), whenever a group of sample observations share a common factor, any theoretical and empirical analysis not accounting for clustering effects would give inconsistent parameter estimates. This points to the need to account for cluster-specific effects in the modeling data generated from on-site samples, where individuals are surveyed at various sites in a given stratum across the National Forest.

In NVUM surveys, individuals are sampled at various sites within a National Forest which are stratified according to site type and site use. A group of individuals surveyed at a particular site share common factors, the observed and unobserved site specific attributes. For example, individuals surveyed at a fishing site have a common recreational use-value for fishing. Statistically, there is a strong reason to believe that individuals intercepted at the same site are somehow correlated rather than independent. According to Galwey (2006), the relationship of the outcome variable, which is visits to a recreation site, may be perfectly replicated for each site, but most likely there will be some differences in this relationship. These differences, or between-site variations, could be ascribed to chance or to some observed or unobserved characteristics or attributes. Therefore to capture the within-site correlation, it is important to introduce site-specific heterogeneity.

Count outcomes are modeled as discrete outcomes and not continuous quantities using a poisson or a negative binomial distribution. The latter is a more flexible and reasonable assumption for empirical data because it drops the assumption of equidispersion. A negative binomial distribution is derived by introducing heterogeneity resulting from unobserved individual taste and preference. Greene (2005) points out that heterogeneity can be introduced the second time if a negative binomial is the base model. We exploit this idea to introduce heterogeneity for the second time. But unlike Greene, we introduce site-specific heterogeneity instead of individual-specific heterogeneity, to explain correlation among individuals sampled at the same site.

Introducing heterogeneity in a poisson model to derive a negative binomial distribution causes heteroskedasticity in estimation of standard errors. Espiñeira and Tuffour(2008) use a more flexible specification for the overdispersion parameter to correct for heteroskedasticity in modeling demand for Gros Morne National Park. They make the overdispersion parameter a function of individual characteristics and show that doing so improves the goodness of fit. Greene (2005) also recommends this specification to correct for heteroskedasticity.

In this paper it has been hypothesized that in a stratified on-site sample, there is a strong reason to believe that individuals sampled at the same site are correlated rather than independent. The hypothesis is tested by modeling demand for outdoor recreation at the George Washington/Jefferson National Forest, where individuals are sampled at 88 sites clustered under four settings types. It is shown that the site-specific effects are significant and there is a strong theoretical and empirical reason to introduce such site-specific effects. By estimating design effects, it is shown that the asymptotic standard errors for the travel cost variable are significantly different under the assumption of clustered sampling rather than random sampling. It is also shown that the expected mean estimates, which are often used for the purpose of projections, is significantly different in each model and so is the estimate of the overdispersion parameter.

This paper is organized as follows. In the second section we give details about the data used for our analysis. In the third section we explain our theoretical model. In the fourth section we specify our empirical model and summary statistics. In the fifth section we estimate six models: a poisson model accounting for stratification and truncation(TSP); a negative binomial model accounting for stratification and truncation(TNB); A poisson model accounting for truncation, stratification and site-specific effects(TSP2); a negative binomial model accounting for stratification, truncation, and an overdispersion parameter to vary by individual characteristics(TNB1); a negative binomial accounting for stratification, truncation, and site-specific effects(TNB2); and finally a negative binomial model accounting for stratification, truncation, and accounting for site-specific effects and an overdispersion parameter to vary by individual characteristics(TNB3). Conclusions are presented at the end of the chapter.

Data

The empirical model will be estimated using NVUM data collected for the George Washington/Jefferson National Forest in the southeastern region of the U.S. The NVUM was conducted at 88 sites stratified by settings within the National Forest. The settings include Wilderness (WILD), Day Use Developed Sites (DUDS), Overnight Use Developed Sites (OUDS) and General Forest Area (GFA). There are 781 sample observations. The data was collected for four sample years, 2000-2003. More detail was provided on NVUM in the previous chapter. For our analysis, we have only included observations for which recreational trips to the National Forest are less than 52. Following Bowker et al. (2009) we also deleted observations with travel cost greater than 720 and people in the vehicle greater than 10.

Theoretical Model

According to Haab and McConell (1996),

> "estimation of single site demand models begins with an assessment of the data generating process which is governed by the assumed stochastic structure of the demand functions and the sampling procedure."

In this chapter we discuss modeling the stochastic structure of the demand functions. The stochastic structure of demand depends on whether the dependent variable, which is an individual's trips to a site, is assumed to be distributed continuously or as a count variable. For the travel cost model the dependent variable is a count variable. Count data for number of visits to a recreational site is not available in continuous quantities.Under this scenario poisson distribution results in an asymptotic outcome, according to Hellerstein (1996). This is because a binomial distribution approaches a poisson distribution as the number of draws approaches infinity. However, when the dependent variable is a count outcome, equidispersion of data is rarely a realistic empirical assumption. A negative binomial distribution is statistically derived by introducing an unobserved individual specific effect in the poisson distribution. The effect is random and each effect is independent of each other and follows a gamma distribution with a dispersion parameter.

Unobserved individual effects are consistent with utility theory. These unobserved effects are attributed to an individual's taste and preferences which are known by the individual but are unobserved by the analyst. One common phenomenon with any travel cost study is that the high frequency visitors who live close to the site make numerous low cost visits, whereas the low frequency visitors who live far away from the site make a few high cost visits. Combining high frequency and low frequency visitors does not account for differences in these individuals, leading to observed over-dispersion in the data. Therefore, we have reasons to believe that the base model for travel cost is a negative binomial with the introduction of unobserved individual-specific effects in the poisson model. We use a negative binomial with a quadratic variance function (NB2) as our base model which is a good approximation in many empirical situations. Also, maximum likelihood estimation of NB2 is robust to misspecification of the conditional mean (Cameron and Trivedi, 1986).

In this chapter, it has been hypothesized that there are reasons to believe that the stochastic process includes unobserved site-specific effects which account for the differences across various sites where the on-site sampling is conducted. Therefore, according to our hypothesis, unobserved effects are introduced in the model. But these are not individual specific unobserved effects but site-specific effects. In the previous chapter issues of weighting to control for choice-based survey design were discussed. In this chapter limitations of the independence assumption in survey data is discussed and econometric techniques are suggested to correct for such limitations.

In microeconometrics, an individual's choice between various sites is treated as a separate estimation equation, logit or nested logit. Because it is conditional on choice, the dependent variable is estimated as a count process. Various applications include site-specific effects in the choice equation. However, the sampled site data for all the sites is extremely costly and in most cases is not available. In this case it becomes even more important to introduce site-specific effects in the count equation. This model can be used to estimate demand for a given National Forest where a random sample is selected at various sites within a setting. When non-negativity, stratification and truncation are included this model would also account for correlation in the variance parameter among various individuals going to the same settings.

The random negative binomial model (RNBM) used by Greene (2005) in a panel data setting is generalized to a cluster of sites in the George Washington/Jefferson National Forest to capture intra-cluster correlation in the variance-covariance matrix. Greene (2005) shows that heterogeneity can be introduced twice if a negative binomial is the base model. A random model is chosen over a fixed effect model to capture the intra-cluster correlation which implies from relaxing the independence assumption within a given cluster.
The log-likelihood of poisson correcting for truncation and stratification is given by (TSP),

$$logl = (y_{ij} - 1)(X'_{ij}\beta) - Exp(Xij'\beta) - log(\Gamma(y_{ij})) \tag{1}$$

and expected mean is given by,

$$E(y \mid x) = EXP(x\beta) + 1 \tag{2}$$

Site-specific effects are added in the mean statement, to derive the poisson distribution model correcting for truncation and stratification with site-specific effects(TSP2). In recreational demand models these site-specific effects could be attributes about a particular site which are unobserved. According to Murdock (2006),

> "one obvious way to address unobserved heterogeneity is to simply include a full set of alternative specific constants. The proposed approach will be useful when there are important characteristics that only vary across recreation locations and not also across time or individuals."

He mentions such site characteristics for fishing such as regulations, water quality, fish consumption advisories, physical characteristics, adjacent land use, and the presence of facilities.

$$Xij'\beta + \sigma b_j \tag{3}$$

where,

$$b_j \sim N(0,1) \tag{4}$$

The negative binomial correcting for truncation and endogenous stratification can be derived by introducing individual-specific heterogeneity which follows a one parameter gamma distribution (TNB),

$$= log(y_{ij}) + log\Gamma(yij + \alpha^{-1}) + y_{ij}log(\alpha) + (y_{ij} - 1)(x'ij\beta)$$
$$-(y_{ij} + \alpha^{-1})log(1 + \alpha exp(x'_{ij}\beta)) - log\Gamma(\alpha^{-1})$$

and the expected mean is given by,

$$E(y \mid x) = EXP(x\beta) + 1 + \alpha EXP(x\beta) \tag{5}$$

Subject-specific effects in 3.6 are similar to 3.3 and 3.4.
The overdispersion parameter in the TNB1 and TNB3 models is specified as,

$$\alpha = \frac{1}{exp(Z_i'\gamma)} \tag{6}$$

The nlmixed procedure in SAS is used to maximize the unconditional likelihood given by,

$$Prob[Y = y_{ij}|x_{ij}] = \int_{b_j} Prob[Y = y_{ij}|x_{ij}, b_j] f(b_j) db_j \tag{7}$$

where,

$$Prob[Y = y_{ij}|x_{ij}, b_j]$$

is given by 3.5, and $f(b_j)$ is given by 3.4,

Empirical Model

For the purpose of estimation, we have scaled our data by dividing explanatory variables by it's mean. Summary statistics are tabulated in Table 1.
The empirical model is specified as,

$$NFV12MO_j^i = f(INCE^i, AGE^i, PEOPVEH^i, GENDER^i, TC^i) \tag{8}$$

Table 1: Summary Statistics for George Washington/Jefferson National Forest NVUM Data, 2000-2003

	Mean1	Min	Max
INCE[a]	23388.03	12647.48	105597.62
AGE[b]	41.571	18	75
GENDER[c]	0.213	0	1
PEOPVEH[d]	2.469	1	9
TC[e]	42.136	0.469	1103.84
NFV12MO1[f]	13.147	1	51
STYPE1[g]	0.117	1	0
STYPE2[h]	0.097	1	0
STYPE3[i]	0.445	1	0
STYPE4[j]	0.342	1	0
NOBS	600		

[a]IRS reported average after tax income for an individual's ZIP Code
[b]Age
[c]A dummy for Gender equals 1 if female
[d]No.of people in the vehicle
[e]As a function of one way travel distance and income foregone (Refer to footnote 1, Chapter2)
[f]Number of annual recreation visits per group
[g]An indicator for Wilderness visits
[h]An indicator for Day Used Developed Site visits
[i]An indicator for Overnight Used Developed Site visits
[j]An indicator for General Forest Area visits

where,

$$i = 1, 2, ...N$$

are the number of individuals

$$j = 1, 2, ...88$$

are the number of sites in the sample The dependent variable in the empirical model is the number of annual recreation visits to the George Washington/Jefferson National Forest per group. Demand for visits is a function of six variables: own price or cost of the trip (TC), number of people in the vehicle

(PEOPVEH), annual income (INCOME), gender (GENDER1), and age (AGE). Site-specific effects are included in the mean statement additively.

The following example illustrates the motivation of the introduction of site-specific effects to capture the correlation between individuals sampled at the same site. In the Chattahoochee National Forest, Brasstown Bald is a popular visitor attraction. Rising 4,784 feet above sea level, Georgia's highest mountain allows clear views of four southern states, Georgia, Tennessee, North Carolina and South Carolina. This site has four hiking trails: Brasstown Bald Trail, the Arkaquah Foot Trail, Jack Knob Foot Trail, and Wagon Train Foot Trail. The observatory also provides facilities for picnicking and nature viewing. The view from the 4,784 feet peak is a popular attraction and most visits to the site are of short duration and usually involve nature viewing and relaxing as the primary activities.

The 5.5 mile long Arkaquah Foot Trail near the observatory is a wilderness trail that attracts a wide variety of hikers and nature viewers. The duration of visits to this trail is usually longer than the duration of visits to the observatory and the site draws both locals and non-locals. The trailhead connects to Track Rock Gap, one of the best known of the petroglyph, or marked stone sites, in Georgia. The Jack Knob Foot Trail is about 4.5 miles and leads to the famous Appalachian Trail. The Wagon Train Foot Trail is 5.8 miles and leads to the Wagon Train Road which ends at Young Harris College. The trail is traditionally hiked by graduating students and their families, the evening before graduation. Thus, it mostly draws locals for a short duration of time.

The above example suggests dependence between individuals surveyed at a particular site due to some observed or unobserved site-specific effects. In a survey sample of this nature, random sampling might not be the most reasonable assumption about the data. Dependence between individuals visiting the same site to estimate the demand for a single National Forest is modeled. It is most likely that individuals surveyed at a given site are correlated rather than independent. The above argument is used to motivate a mixture model where site-specific random effects follow a standard normal distribution. In this model, the overdispersion parameter is modeled as,

$$\alpha = f(INTERCEPT, STYPE) \tag{9}$$

where,
Site types (STYPE) or settings is a dummy variable for each settings type. Settings include Day Used Developed Sites (DUDS), Overnight Used Developed Sites (OUDS), General Forest Area (GFA) and Wilderness (WILD). For estimation, the dummy for Overnight Used Developed Sites is dropped. Thus, the OUDS setting serves as base.

Results

Similar to Pepper (2002), design effects for the variables are constructed in the mean statement for two models, TNB1 and TNB3. Design effects are defined as the ratio of asymptotic variance under the assumption of random sampling to asymptotic variance under the assumption of clustered sampling. The sampling scheme has negligible effects on the asymptotic variance for most of the variables.

Table 2: Design Effects for TNB1 and TNB3 model

TNB1[1]	TNB3[2]
TC	1.020

For the travel cost variable,the design effect is, around 1.020, implying that the estimated standard error in the clustered sample exceeds that of random sample by 19%. Also, the sample size of our data is fairly small and number of clusters are fairly large(88 sites) where the survey was conducted. The design effects tend to grow as more observations are made within a cluster.

Also, in the TSP2, TNB2, and TNB3 models, significant site-specific effects are found. This can be seen from the significance of the variance parameter, given by sigma in the results. The parameter estimates are 0.663, 0.6693, and 0.317 for the TSP2, TNB2, and TNB3 models respectively, each significant at the 1% significance.

In comparing four negative binomial models, negative binomial model (TNB2) additionally accounting for site-specific effects perform better than the simple negative binomial model accounting for stratification and truncation (TNB), with log-likelihoods of 13108.5 and 13086.5 respectively and BIC criteria are -26183 and -26136 respectively .

Now we compare the two negative binomial models including overdispersion as a function of individual characteristics: TNB1 does not account for site-specific effects, while TNB3 does. The log-likelihood for TNB3 is higher than TNB1 (13111 and 13092.5 respectively). TNB3 also does better than TNB1 in terms of the BIC

Table 3: Estimation Results of Outdoor Recreation Demand for George Wahington/Jefferson National Forest: NVUM DATA: 2000-2003

	TSP[a]	TNB[b]	TNB1[c]	TSP2[d]	TNB2[e]	TNB3[f]
Intercept[g]	3.0655	*not estimated*	-9.7918	2.7756	-0.00305	0.3088
	(0.0676)*	-	(119.01)	(0.1134)*	(1.1076)	(0.8226)
INCE	-0.2635	-.335707	-0.3090	-0.3432	-0.2736	-0.2706
	(0.0538)*	(0.1282)*	(0.1331)**	(0.06921)*	(0.1364)**	(0.1356)**
AGE	0.2732	.199571	0.2817	0.2922	0.2795	0.2929
	(0.0363)*	(0.1353)	(0.1427)**	(0.04196)*	(0.1504)***	(0.1487)**
GENDER	-0.3732	-.437244	-0.4708	-0.2817	-0.3444	-0.3586
	(0.0364)*	(0.1182)*	(0.1218)*	(0.04190)*	(0.1273)*	(0.1263)*
PEOPVEH	-0.2150	-.217506	-0.1785	-0.1356	-0.1096	-0.09735
	(0.0251)*	(0.07847)*	(0.08244)**	(0.02915)*	(0.08721)	(0.08702)
TC	-0.3471	-.241602	-0.2402	-0.2915	-0.2169	-0.2169
	(0.0179)*	(0.03444)*	(0.03480)*	(0.02155)*	(0.03421)*	(0.03411)*
ALPHA	-	20.4326	-	-	14.8816	-
	-	(3.79603)*	-		(17.2892)	-
sigma	-	-	-	0.663	0.6693	0.317
	-	-	-	(0.04015)*	(0.05472)*	(0.05684)*
log(a)						
Intercept	-	-	-12.7354	-	-	-2.3860
	-	-	(119.01)	-	-	(0.8855)*
WILD	-	-	0.1837	-	-	0.2318
	-	-	(0.1542)	-	-	(0.3070)
OUDS	-	-	0.3788	-	-	0.5674
	-	-	(0.1666)**	-	-	(0.3288)***
DUDS	-	-	-0.07110	-	-	-0.07396
	-	-	(0.1063)	-	-	(0.1981)
NOBS	600					
LOGL	-3701.1900	13086.5	13092.5	3234.8	13108.5	13111
BIC	7440.0286	-26136	-26122	6499.6	-26183	-26175

*1% significance
**5% significance
***10% significance

[a] Truncated Stratified Poisson
[b] Truncated Stratified Negative Binomial
[c] Truncated Stratified Negative Binomial; Modeling Overdispersion Parameter
[d] Truncated Stratified Poisson Accounting For Site-Specific Effects
[e] Truncated Stratified Negative Binomial Accounting For Site Specific Effect
[f] Truncated Stratified Negative Binomial; Modeling Overdispersion Parameter and Accounting For Site Specific Effects
[g] Coefficient estimates reported in the first row and standard error reported in parentheses

criterion. The BIC criterions are -26183 and -26122 respectively for TNB1 and TNB3. In modeling mean and overdispersion, the parameter estimates for the intercept are very different in the two models. We can see this in Table 4. The expected mean for TNB1 and TNB3 are 11.338 and 4.5539 respectively.

Table 4: Estimation Results of Expected Mean and Overdispersion Parameter

	TSP[a]	TNB[b]	TNB1	TSP2	TNB2	TNB3
E(Y)	12.4046	42.79	11.338	10.365	32.35	4.5539
alpha	-	20.4326	0.000003	-	14.8816	0.097

[a]For all TSP models,

$$E(y \mid x) = EXP(X'\beta) + 1$$

[b]For all TNB models,

$$E(y \mid x) = EXP(X'\beta) + 1 + \alpha EXP(X'\beta)$$

Conclusions and Implications

It is shown that there is a theoretical and empirical reason to account two times for heterogeneity in modeling recreational demand for National Forests, where individuals are sampled at various sites which are stratified or clustered according to their use. The first time, heterogeneity accounts for dispersion in the data due to unobservability of the process which results in existence of two different types of visitors in the population, high-frequency and low-frequency. The second time heterogeneity is accounted for in order to capture dependence between individuals sampled at similar sites according to a stratified random sample. Positive results for our hypothesis are found. Both in the poisson and the negative binomial model, the model accounting for site-specific effects performs better than the one not accounting for site-specific effects, with statistically significant results. The results are of particular interest in deriving consumer surplus per person trip, since the coefficient for the price variable changes across most of the models. Also, model differences would be important for the purpose of deriving future projections of demand. This is because the expected mean changes across different models. This can be clearly seen from the expected mean calculations in Table 4. Therefore, in this paper a case for treating individuals within a given stratum as dependent rather than independent is made.

References

[1] Bowker, J.M., C.M. Starbuck, D.B.K. English, J.C. Bergstrom, R.S. Rosenberger, and D.W. McCollum, "Estimating the Net Economic Value of National Forest Recreation: An Application of the National Visitor Use Monitoring Database" *Faculty Series Working Paper, FS 09-02, September 2009, Department of Agricultural and Applied Economics, The University of Georgia, Athens*.

[2] Cameron, A. C., and P. K. Trivedi, "Econometric Models Based on Count Data: Comparisons and Applications of Some Estimators and Tests ", *Journal of Applied Econometrics*, Vol 1(1) (Jan. 1986): 29-53.

[3] Egan, K., and J. A. Herriges, "Multivariate Count Data Regression Models with Individual Panel Data From an On-Site Sample", *Journal of Environmental Economics and Management*, Vol. 52 Issue 2 (Sep. 2006), 567-581.

[4] Englin, J., and J.S. Shonkwiler, "Estimating Social Welfare Using Count Data Models: An Application to Long Run Recreation Demand under Conditions

of Endogenous Stratification and Truncation", *Review of Economics and Statistics* 77 (1995): 104-112.

[5] English, D.B.K, et al., “Forest Service National Visitor Use Monitoring Process: Research Method Documentation." *United States Department of Agriculture, Forest Service, General Technical Report SRS-57* 2002.

[6] Galwey, N.W., “Introduction to Mixed Modelling, Beyond Regression and Analysis of Variance", *John Wiley and Sons, Ltd.* 2006.

[7] Greene, W.H., “Functional Form and Heterogeneity in Models for Count Data, *Foundations and Trends in Econometrics”*, Vol. 1, No. 2 (2005) 113-218.

[8] Haab, T.C. and K.E. McConnell, “Count Data Models and the Problem of Zeros in Recreation Demand Analysis", *American Journal of Agricultural Economics*, Vol. 78 (Feb 1996): 89-102.

[9] Hanemann, W. N., “Discrete/Continuous Models of Consumer Demand" *Econometrica* Vol. 52, No.3 (May, 1978): 541-561.

[10] Hellerstein, D.M. “Using Count Data Models in Travel Cost Analysis with Aggregate Data", *American Journal of American Economics*, Vol. 73 (Aug 1991): 860-66.

[11] Karan, V., and D.P. Cram, “Review of Choice-Based and Matched Sample Studies in Management Accounting Research," July 2006. http://aaahq.org/mas/MASPAPERS2007/research_forum/Karan%20and%20Cram.pdf (Accessed July 15, 2010).

[12] Lawless, J.F., “Negative Binomial and Mixed Poisson Regression", *The Canadian Journal of Statistics*, Vol. 15, No. 3,(Sep. 1987), 209-25.

[13] Leamer, E.E.,"Specification Searches Ad Hoc Inference with Non Experimental Data" *Wiley Series in Probability and Mathematical Statistics, John Wiley and Sons, Inc.*, 1978.

[14] Longford, N.T., "Random Coefficient Models", *Oxford Science Publications, Clarendon Press*, 1993.

[15] Manski, C.F., and D.L. McFadden, "Chapter 1: Alternative Estimators and Sample Designs for Discrete Choice Analysis, Structural Analysis of Discrete Data and Econometric Applications", *Cambridge: The MIT Press*, 1981.

[16] Manski, C.F., and S.R. Lerman, "The Estimation Of Choice Probabilities From Choice Based Samples", *Econometrica*, Vol.45, No.8,(Nov 1977): 1977-1988.

[17] Martínez-Espiñeira R., and J. Amoako-Tuffour, "Recreation Demand Analysis Under Truncation, Overdispersion and Endogenous Stratification: An Application to Gros Morne National Park", *Journal of Environmental Management* 88, 2008, 1320-32.

[18] Morey, E.R., "The Demand for Site-Specific Recreational Activities: A Characteristic Approach", *Journal of Environmental Economics and Management*, 8 (1981): 345-71.

[19] Murdock, J., "Handling Unobserved Site Characteristics in Random Utility Models of Recreation Demand", *Journal of Environmental Economics and Management*, 51 (2006): 1-25.

[20] Pepper, J.V., "Robust Inferences From Random Clustered Samples: An Application Using Data From Panel Data Study of Income Dynamics", *Economic Letters*, 75 (2002): 341-345.

[21] Shaw, D., “On-Site Samples’ Regression: Problems of Non-Negative Integers, Truncation, and Endogenous Stratification", *Journal of Econometrics*, Vol. 37 No. 2,(1988): 211-223.

[22] Wang, X., C.V. Eden, and J.V. Zidek, “Asymptotic Properties of Estimators in Maximum Likelihood Estimation", *Journal of Statistical Planning and Inference*, 119, (2004) 37-54.

[23] Winklemann, R., “Seemingly Unrelated Negative Binomial Regression," *Oxford Bull. Econ. Statist.*, 62(4), (2000): 553-560.

[24] Woolodridge J.M., “Asymptotic Properties of Weighted M-Estimators for Standard Stratified Samples," *Economic Theory*, Vol. 17, (2001): 451-70.