Available online at www.sciencedirect.com
SCIENCE DIRECT
ELSEVIER
Journal of Forest Economics 10 (2004) 37–53
Journal of FOREST ECONOMICS
www.elsevier.de/jfe

Valuing special forest products harvesting: a two-step travel cost recreation demand analysis☆

C. Meghan Starbuck[a], Susan J. Alexander[b], Robert P. Berrens[c,*], Alok K. Bohara[c]

[a] *Department of Agricultural and Applied Economics, University of Georgia, Athens, GA 30602, USA*
[b] *Regional Economist Alaska Region, USDA Forest Service, P.O. Box 21628, Juneau, AK 99802, USA*
[c] *Department of Economics, University of New Mexico, Albuquerque, NM 87131, USA*

Received 30 December 2003; accepted 14 January 2004

Abstract

Currently, the empirical literature on outdoor recreation demand lacks estimates of the benefits of special forest products harvesting. This paper provides a recreation demand analysis of non-commercial huckleberry and mushroom picking on the Gifford Pinchot National Forest in southwestern Washington State (USA). Using available survey data and a two-step structural equations model of harvesting and travel cost recreation demand, with a Murphy–Topel (J. Bus. Econ. Stat. 3 (1985) 88) standard error correction, we estimate the consumer surplus associated with special forest products harvesting. Per recreation visitor day consumer surplus is estimated at \$30.82 in 1996 dollars (\$36.06 in 2003 dollars). Estimated values for the full range of non-timber values are becoming increasingly important as public lands management agencies expand their focus to include consideration of alternative forest uses.

JEL classification: Q260

Keywords: Special forest products; Structural equations; Recreation demand; Travel cost method

☆This research was conducted under Joint Venture Agreement PNW 01-JV-11261955-227 between the University of New Mexico and USDA Forest Service, Pacific Northwest Research Station. The authors are responsible for all errors and opinions.

*Corresponding author.
E-mail address: rberrens@unm.edu (R.P. Berrens).

1104-6899/\$ - see front matter
doi:10.1016/j.jfe.2004.03.004

Introduction

Public forestlands provide a wide variety of beneficial uses, including numerous opportunities for recreational activities (Loomis, 1993). To the extent possible, quantifying these recreational benefits will facilitate resource allocation decisions. Estimates of non-market economic values, often referred to as "non-timber values," can be important inputs for evaluating tradeoffs in forest management and policy. Estimated values relating to non-timber uses of forested areas are becoming increasingly important as public lands management agencies expand their planning focus to include full consideration of alternative forest uses. The objective of this paper is to provide a structural equations approach to modeling recreation demand of a common, but under-investigated, recreational activity—special forest products harvesting.

The current empirical literature on outdoor recreation demand (see Rosenberger and Loomis, 2001) lacks estimates of wild berry and mushroom harvesting. The literature also lacks development of a structural equations approach (jointly examining harvest and recreation demand) without relying on a full information maximum likelihood (FIML) estimator (see Englin et al. (1997) for an example of the structural equations FIML model). Outside of a recreation demand context, Murphy and Topel (1985) develop the two-step correction method, but do not derive corrections for count data models. Greene (2000) presents the Murphy–Topel method, but also does not provide count data estimators.

Using an available data set, we develop a two-step structural equations model of harvesting and travel cost recreation demand for both a Poisson and truncated Poisson demand equation, using the Murphy–Topel (1985) standard error correction. The application is to the estimation of the non-commercial value of huckleberry and mushroom picking on the Gifford Pinchot National Forest in southwestern Washington state (USA). Per recreation visitor day consumer surplus is estimated at $30.82 in 1996 dollars ($36.06 in 2003 dollars). More generally, extending the recreation demand method to incorporate a jointly determined demand determinant provides a useful analytical tool for a wide variety of non-timber forest products and recreational values including fishing, hunting, mountain biking, rock climbing, and other skill and/or equipment intensive recreation activities.

The paper proceeds as follows. First we provide a brief background discussion on the history and current importance of special forest products. The next section provides a description of the survey data, its origination, limitations, and descriptive statistics. We then describe the structural equations model and the two-step estimation method. The following section then provides the empirical results and estimated consumer surplus measures, and the last section provides some conclusions.

Background

Increasing awareness of environmental effects associated with timber harvesting has created a need for various land management agencies to begin focusing attention on sustainable extraction of special forest products (SFPs).[1] Intact forest areas impact ecological health by providing environmental services such as CO_2 storage, conservation of biological diversity, and the maintenance of regional climatic zones (Gram, 2001, p. 109). In addition to the ecological services provided by forests, there is a wide range of products extracted from natural areas that have both market and non-market values associated with their extraction. Special forest products are materials harvested from forests that are non-timber in nature and include: wild mushrooms and berries; ferns and medicinal plants; Christmas trees; and peat moss (Hansis, 1998, p. 69).

A small number of previous studies have examined the value of special forest products in the United States (Hansis, 1998; Anderson et al., 2000; Markstrom and Donnelly, 1988), while a slightly larger literature has examined the value of SFPs outside of the USA, mostly in South American countries (Arnold and Perez, 2001; Simpson, 1999; Robinson et al., 2002; Godoy and Lubowski, 1992; Mattson and Li, 1994). One study in Finland measures the commercial value of wild berry harvesting and the contribution to household income that harvesting activities generate (Kangas, 2001). Kangas (2001) uses a Tobit model to estimate a value of FIM 4.8 million earned by wild berry harvesters in the Suomussalmi area of Finland. Markstrom and Donnelly (1988) use the TCM to estimate the demand for Christmas tree cutting in Colorado (USA). They estimate an average value of $9.37 per standing Christmas tree, with an average consumer surplus estimate of $4.37. Anderson et al. (2000) study the value of fern gathering on the San Bernardino National Forest in California (USA), but focus on the characteristics and motivations of harvesters. They collect detailed survey information, but do not estimate any direct value from their data. Godoy and Lubowski (1992) compile estimates of the value of caimans, elephants, game meat, flora, rattan, palms, and wild *camu camu* from the literature on NTFPs from India, Zaire, the Amazon, Peru, Brazil, and Mexico. Estimates range from $0.75 per hectare per year for caimans (in US dollars) to $420 per hectare per year for the flora inventory of Iquitos, Peru. Most of the Godoy and Lubowski values are derived from ethno-botanical surveys and anthropological-based studies. In reviewing the literature concerning special forest products, it is notable that only one study employs a recreation demand or travel cost method (Markstrom and Donnelly, 1988) to derive value estimates, and this was for Christmas tree cutting.

Data description

To estimate our models, we use a data set collected in 1996 from the Gifford Pinchot-National Forest (GPNF) in southwestern Washington State (USA). Several special forest products are routinely harvested on the GPNF, including wild mushrooms, wild blackberries and huckleberries, and plants such as ferns. The

[1] Special Forest Products (SFPs) are also called: non-timber forest products (NTFPs), non-wood forest products, minor forest products, or wild harvested products (Hansis, 1998, p. 67).

GPNF represents one of only a handful of premier wild berry and mushroom harvesting locations in the United States, and both commercial and recreational harvesting permits are issued to harvesters.[2] Currently, each recreational harvester is allowed three gallons of huckleberries and mushrooms free of charge per year. If larger quantities are wanted, or if the harvester intends to sell the harvest, a commercial permit is required which can be purchased from the ranger district offices.[3] Hansis (1998) collected information on the number of permits issued by the GPNF for huckleberry and mushroom harvesting between 1992 and 1994; the GPNF issued a total of 2620 personal mushroom harvesting permits and 8342 commercial mushroom permits. The GPNF issued 25,621 personal huckleberry permits, as well as 73 commercial huckleberry permits.

The data used for this analysis come from a 1996 survey of special forest products harvesting permit holders. The target population was individuals holding recreational use permits to harvest berries and mushrooms from the Packwood and Randall districts in the Gifford Pinchot National Forest in Washington. The permits were for personal use (non-commercial) only and were required for harvesting from the area. The permits were available for no cost from the forest ranger district offices by filling out a simple self-administered form including the applicant's name and address. For the mail survey, all the names and addresses obtained from the recreational use permits issued in 1996 by the Packwood and Randall district offices were used. This resulted in exactly 1000 surveys being mailed out. The self-administered survey packet contained a map and a hand-signed cover letter. Postcard reminders were mailed two weeks after the first mailing, and complete packages were mailed again to non-respondents a month after the first mailing. This resulted in 485 returned surveys (48.5%), with 462 usable responses (46.2%).

The survey collected information on the 1996 mushroom and berry harvest season in the Packwood and Randall districts of the Gifford Pinchot National Forest. There are approximately seven harvesting areas within the general area, but we have no information on variations in site quality, or the number of trips to individual areas, so we treat them as a single site. This is also consistent with the permitting process and the fact that individuals will move across harvesting areas. Questions included transportation type, type of special forest products gathered, quantity gathered, number of trips taken, and a standard set of socio-demographic questions.

[2] The Gifford Pinchot National Forest is one of the oldest National Forests in the United States. Included as part of the Mount Rainier Forest Reserve in 1897, this area was set aside as the Columbia National Forest in 1908. It was renamed the Gifford Pinchot National Forest in 1949. Located in southwestern Washington State, the Forest contains 1,312,000 acres and includes the 110,000-acre Mount St. Helens National Volcanic Monument established by Congress in 1982. Since the 1994 Northwest Forest Plan was enacted, logging on the GPNF has declined to approximately 2 million board feet per year down from the 350 million board fee of lumber during the mid-1980s (Gordon, 2002). This decrease in logging has increased interest in special forest products harvesting on the GPNF.

[3] A commercial berry permit costs $25 and allows a harvest of 250 pounds of berries. There are four commercial mushroom permits: a 2-day permit costs $25 and allows a maximum of 62.5 pounds; a 10-day permit costs $40 and allows 100 pounds; a 20-day permit costs $60 and allows 150 pounds; and a biannual permit costs $125 and allows 312.5 pounds of mushrooms to be harvested.

The average annual household income in our sample was $42,390; the average age was 55 years; the sample was evenly split between men and women; and 52% of the sample held at least a bachelors degree. Ninty-one percent of the sample was white, and 60% were employed full time. The mean distance traveled to harvest at the Randall-Packwood area was 96 miles, and the mean number of trips per respondent for the 1996 season was 1.58. The average respondent harvested 2.48 gallons of berries and 2.48 gallons of mushrooms for the season. The average mushroom and berry harvest is within the three gallons allowed per year for the recreational harvester. However, the data contained individuals who harvested more than the three gallons allowed by the recreational permit. For the berry harvesters 74% of the sample harvested 3 or fewer gallons, while 23% of the sample harvested between 4 and 10 gallons of berries. 99% of the mushroom harvesters in the sample harvested between 0 and 3 gallons of mushrooms, and only one harvested 6 gallons of mushrooms.[4] The majority of the harvesters were within the permit requirements indicating that the sample consists mostly of recreational, not commercial, harvesters. Descriptive statistics are provided in Table 1.

Methods

Numerous authors have estimated travel cost models and detailed the theoretical, empirical, and econometric aspects of recreation demand modeling, including various applications to forest management (Loomis, 1993). Bockstael et al. (1987) develop the utility-theoretic aspects of recreation demand models; Hellerstein (1993) and Hellerstein and Mendelsohn (1993) developed the econometric reasoning and tools for count data and panel data estimators. Ovaskainen et al. (2001) provided count data estimators for truncated and endogenously stratified data; and Chakraborty and Keith (2000) presented a recent count data travel cost model that is similar in design to the sampling methodology of the survey used in this study. A variety of sources (Shaw, 1988; Creel and Loomis, 1990; Grogger and Carson, 1991; Englin and Shonkwiler, 1995; Cameron and Trivedi, 1986) discuss the issues associated with truncated, and endogenously stratified count data. The reader is referred to Ovaskainen et al. (2001) for a recent overview.

In considering the recreational demand for special forest products harvesting on the GPNF, we can specify an individual's general trip demand model as a function of travel costs, household income, harvesting level or quality, and a set of socioeconomic characteristics:

$$\text{Visits} = f(\text{TC}, \text{TC}*\text{MUSH},\ \text{MUSHDUM},\ \text{INC},\ \text{GENDER},\ \text{AGE},\ \text{HARVEST}), \quad (1)$$

[4] In the survey, answers to the harvesting questions could be given in either pounds or gallons. All values were converted to gallons using the following estimates provided by the Bureau of Land Management: 2 lb mushrooms/1gallon of mushrooms, 5 lb of berries/1gallon of berries. This conversion is consistent with the USFS recreational harvest permit units of 3 gallons of mushrooms and 3 gallons of berries allowed per harvester per year. The conversion factors were found at: http://www.blm.gov/nhp/efoia/mt/2000/im/00mtm, accessed 5/12/2003.

Table 1. Variable descriptions and descriptive statistics

Variable	Description	Mean	Std.dev.
ONSITE	Length of stay on site, measured in hours	14.84	27.49
ONEWAY	How far from home to site, measured in hours	2.36	1.37
COMPDIST	Calculated distance from home to site, using center of home Zip Code area to site (latitude/longitude) using USDA's ZIPFIP Program	94.28	177.01
AGE	Age in years	55.26	14.81
VISITS	Total number of visits made to the Gifford Pinchot in 1996 for harvesting mushrooms and berries	1.63	1.37
GENDER	Coded 1 if male, 0 if female	0.52	0.49
INC	Household income in ranges from [min.]$10,000–[max.]$100,000 or greater; in increments of $10,000, median of range used	42,412	22,463
RACE	Binary variable, coded 0 if non-white, 1 if white/Caucasian	0.92	0.27
EDUC	Binary variable, coded 1 if number of years in school greater than 12, 0 else	0.53	0.49
EMPL	Binary variable, coded 1 if employed fulltime, 0 else	0.61	0.48
MUSH	Gallons of mushrooms gathered in 1996 season	2.48	15.02
BERRY	Gallons of berries gathered in 1996 season	2.48	3.03
EXPEND	Reported expenditures for camping or lodging, gasoline, and equipment for the 1996 harvest season	72.37	156.16
GRPSZ	Harvesting group size	3.87	2.11
HARVEST	Aggregate harvest variable, HARVEST = MUSH + BERRY, measured as total harvest for the season	4.96	14.99
MUSHDUM	Coded 1 if harvested mushrooms, 0 else	0.18	0.38
TCMUSH	Interaction term, TCMUSH = TC∗MUSH	6.73	25.81
TC	$TC = (0.31*2*COMPDIST)/GRPSZ + 0.33*(INC/2000)*2*ONEWAY$	55.92	51.24

Note: [$N = 392$].

where TC is the individual's constructed travel cost measure, which serves as the price proxy, TC∗MUSH is an interaction term between travel cost and whether or not the observation came from a mushroom harvester (with huckleberry-only harvester as the reference case), INC is the household income variable, and HARVEST is the combined quantity of mushrooms and berries harvested per season.[5] With this basic demand specification in mind, the rest of this section discusses various issues in developing a more complete demand model.

First, as Hellerstein (1993) and others have shown, the number of trips taken is not a continuous distribution. Trip data occur as positive integers. This results in linear estimates of recreation demand being misspecified; they are inefficient and biased indicators of true demand (Freeman, 1993). The non-negative integer nature of the data results in a count data estimator such as Poisson or Negative Binomial being the appropriate distribution. The Poisson distribution is defined as

$$P(Y_i = y_i) = \frac{e^{-\lambda_i}\lambda_i^{y_i}}{y_i!}, \quad y_i = 0, 1, 2 \ldots T, \tag{2}$$

where Y is the dependent variable (VISITS). The Poisson distribution defines the probability of event occurrence (VISITS) for each individual i. The distribution is defined over all observations from $i = 1 \ldots T$. For notational simplicity, in what follows the subscript i will be omitted. Typically a log-linear specification for λ is used, which can then be defined by letting, $\ln \lambda = \beta' X$. Where X is the vector of independent variables, in this case as we have defined in Eq. (1). This then implies that the expected number of events per period is given by $E[y|X] = \mathrm{Var}[y|X] = \lambda = e^{\beta' x}$ (Greene, 2000, p. 880). If the mean and the variance are equal, then Poisson is an appropriate distribution, but if the data are dispersed such that the mean and the variance are not equal then the Negative Binomial distribution should be used (Greene, 2000, p. 880). The Negative Binomial relaxes the assumption that the expected value must equal the variance, but uses the same log-linear specification for λ.

In addition to the choice between Negative Binomial and Poisson, there is also the issue of truncation and endogenous stratification that can arise from the sample selection process. A sample is truncated if it does not include non-users, and a sample is endogenously stratified if the likelihood of certain persons being sampled depends on the frequency of visits to the site (Englin and Shonkwiler, 1995, p. 104).

Our sampling involves all recreational permit holders and is not derived from an on-site sample and should not be endogenously stratified. While the sample process allows for zero trip response, in our usable sample set there are no observations of zero, and we effectively observe $y_i = y_i^*$ if $y_i^* > 0$. To account for possible truncation in our data we estimate both the standard Poisson and the truncated Poisson as given in Englin and Shonkwiler (1995). The reader is referred to Ovaskainen et al. (2001) for a full exposition of truncation, endogenous stratification, and over-dispersion of recreation trip data.

Another potential issue is the cost allocation for individuals for whom the trip to the site was not the main purpose. In our sample, only special forest products permit holders were surveyed and the question regarding number of trips taken to the GPNF in 1996 specifically asks respondents to report the number of *harvesting* trips

[5] This combined construction of HARVEST is used in order to develop the structural equations model, and is reasonable because there are so few mushroom harvesters in our sample (only 17 reported collecting any quantities of mushrooms); additionally, the per gallon prices of mushrooms and berries are in the same order of magnitude, so the measurement error introduced by using this construction should be within acceptable limits.

taken. Thus, our entire sample is the target population and we do not need to reject observations based on reported trip objectives.

To construct the individual travel cost (TC) variable the following method was used:

$$\mathrm{TC} = ((0.31*2*\mathrm{COMPDIST})/\mathrm{GRPSZ}) + 0.33\left(\frac{\mathrm{INC}}{2000}\right)*2*\mathrm{ONEWAY}, \tag{3}$$

where TC is the travel cost variable, and ONEWAY is the respondent's self-reported one-way distance from home (multiplied by two to get round trip distance), COMPDIST is the ZIPFIP calculated distance from the respondent's home zip code to the harvesting site, and GRPSZ is the number of individuals in the harvesting group.[6] The survey did not ask for the wage rate, so income was divided by 2000 (40 h/week*50 weeks/year) to construct the wage rate. As is historically common in TCM studies, one-third the wage rate was used to approximate the opportunity cost of travel time. From the US Government Services Administration, the standard mileage reimbursement rate for 1996 was $0.31 per mile.[7]

For exposition, in Eq. (1) we included HARVEST (seasonal quantity in gallons) in the vector of regressors. However, the use of the level of harvest as an explanatory variable in recreation demand is problematic, as it may not be an exogenous factor of demand. If the individual chooses HARVEST internally or jointly with the total trip decisions, then HARVEST may be endogenous to trip demand. Englin et al. (1997) examine the demand for fishing trips taken by anglers, and point out that anglers prefer to catch more fish to less and thus some estimate of expected fish catch enters the demand for recreational fishing trips. Modeling the catch rate as an exogenous demand factor will lead to biased and/or inefficient parameter estimates. If catch rate is endogenous, then the problem must be modeled as one of joint determination of catch and trip demand. This leads to a structural equations model where catch and trip demand is estimated jointly (Englin et al., 1997, pp. 33–34). Much like the problem of modeling the catch rate and its effect on trip demand, the demand for recreational harvesting trips is likely to be jointly determined by the harvest rate. In order to model special forest products harvesting on the Gifford Pinchot, a structural equations approach must be used due to the potential endogeneity of harvest to trip demand.

The structural equations approach then defines the following jointly determined recreation demand model:

$$\mathrm{HARVEST} = f(\mathrm{ONSITE},\ \mathrm{GRPSZ},\ \mathrm{EXPEND},\ \mathrm{MUSHDUM},\ \mathrm{RACE},\ \mathrm{EMPL}), \tag{4}$$

[6] TC was based on the respondent's reported trip distance. Missing responses on ONEWAY were imputed using the mean sample value. COMPDIST used the respondent's zip code and the latitude/longitude of each of the sites to construct a distance traveled to the site. The program that computes these distances uses a centric-road distance method for calculating the distance, which can introduce measurement error, but the survey did not provide for self-reported distances from home to the harvesting sites.

[7] See URL: http://www.gsa.gov; site accessed on 10/15/2002.

$$\text{VISITS} = f(\text{TC, TCMUSH, MUSHDUM, INC, GENDER, AGE, EXPHRV}), \quad (5)$$

where (4) is the harvesting production function, and includes a measure of harvesting effort or time spent (ONSITE), number of harvesters in the group (GRPSZ), trip expenditures (EXPEND), if they are mushroom harvesters (MUSHDUM), a dummy variable for white versus non-white harvesters (RACE), and employment status (EMPL). Trip demand (5) is then modified from our basic specification (1) to replace the harvest quantity (HARVEST) with the *predicted* or expected harvest quantity (EXPHRV), which is an estimable, but unobservable variable.[8]

The two-step structural equations approach

Typically, an unobserved, but estimable variable is replaced by its predicted value from an auxiliary regression or estimated jointly with the model. Two approaches can be taken when modeling these types of endogenous variables problems. The first is the FIML approach, and the second is the two-step method. Englin et al. (1997) use the FIML joint estimation approach, but as Murphy and Topel (1985) discuss the use of FIML is not always preferred. The use of FIML dictates strict assumptions concerning the joint distribution of errors between the auxiliary equation and the equation of interest. The conditions under which FIML leads to consistent estimators are limited. In many cases the ability to appropriately define the joint distribution is problematic, which dictates the use of a two-step estimator. Additionally, FIML frequently has convergence problems. The two-step estimator can be used under more general conditions because there is no need to determine a joint-density function for the errors, and as long as each separate function is estimable, does not have difficulty converging (Greene, 2002).

However, a two-step method fails to account for the unobserved regressors used to calculate the second step parameters and standard errors. The imputed values applied in the second step are thus measured with sampling error. If it is assumed that the auxiliary (first step) model produces consistent estimates of both first step parameters and their asymptotic covariance matrix, then the sampling error of the imputed values vanishes in the limit. Thus, the second step parameters are consistently estimated. Further, under fairly general conditions the estimated limiting distribution of this error may be used to consistently estimate the variances of the second step parameters (Murphy and Topel, 1985, p. 89).

To model the recreational trip demand associated with wild berry and mushroom harvesting on the GPNF, it is necessary to jointly model harvest and trip demand, as shown in Eqs. (4) and (5). Thus, following Murphy and Topel (1985), let $Y =$ VISITS, $W =$ HARVEST, and let:

$$F_1(W; Z, q_1) = W = f\begin{pmatrix} \text{CONSTANT, ONSITE, GRPSZ,} \\ \text{EXPEND, MUSHDUM, RACE, EMPL} \end{pmatrix}, \quad (6)$$

[8] Eqs. (4) and (5) represent total special forest products harvesting trips taken to the GPNF for 1996.

$$F_2(Y; X, Z, q_1, q_2) = Y = f\begin{pmatrix}\text{CONSTANT, TC, TCMUSH, MUSHDUM}\\ \text{INC, GENDER, AGE, } E[W|Z]\end{pmatrix}. \tag{7}$$

Eqs. (6) and (7) define the structural equations model. Let F_1 be the auxiliary model that predicts the harvest of wild berries and mushrooms, and let F_2 be the model that predicts the number of trips taken to the GPNF for recreational harvesting, which includes the predicted harvest level. Then let W be the harvest and let Y be the number of trips taken in 1996 to the Gifford Pinchot. Let Z be the vector of independent variables associated with F_1 and let θ_1 be the estimated parameters from the auxiliary model. Then let X be the vector of independent variables for F_2 and let θ_2 be the estimated parameters of the main equation. The marginal distributions can both be estimated using maximum likelihood methods (Murphy and Topel, 1985, pp. 94–95). Then θ_1 is estimated, yielding a prediction for W, which is then used in F_2. After estimating θ_2, the variance–covariance matrix for the second step must be corrected using the Murphy–Topel correction. This approach allows us to treat HARVEST as endogenous and use the prediction for harvest in the second step. The set of regressors used in the auxiliary equation are not contained within the set of regressors for the main equation, thus differentiating this structural equations approach from an instrumental variables method.

Then the models (F_1, F_2) in a log-linear/Poisson two-step estimator for harvest and trips are defined, respectively, as

$$P(W = w) = \left(\frac{1}{\sqrt{2\pi\sigma^2}}\right)\mathrm{e}^{(1/2(\ln w - Z'\theta_1)/\sigma^2)(1/w)}, \quad \text{where } \lambda_1 = E[w|Z] = \mathrm{e}^{Z'\theta_1 + \sigma^2/2}, \tag{8}$$

$$P(Y = y|y > 0) = \frac{\mathrm{e}^{-\lambda_2}\lambda_2^y}{[1 - \mathrm{e}^{-\lambda_2}]y!}, \quad \text{where } \lambda_2 = \mathrm{e}^{X'\theta_2 + \gamma E[w|Z]} = \mathrm{e}^{X'\theta_2 + \gamma \exp(Z'\theta_1 + \sigma^2/2)}. \tag{9}$$

Eqs. (8) and (9) define the marginal distributions of two random vectors Y and W. Note that the conditional mean function for F_2 contains the expected value of F_1 in the definition of the λ_2 (Greene, 2000, p. 435). Thus, the predicted number of trips is defined by the expression:

$$E[y|X, Z, \theta_1, \theta_2, \gamma, |y > 0] = \frac{\mathrm{e}^{-\lambda_2}\lambda_2^y}{[1 - \mathrm{e}^{-\lambda_2}]y!}. \tag{10}$$

As (10) shows, the predicted number of trips is jointly determined by the expected amount of harvest. From the density functions, we form the log-likelihood functions as

$$\ln L_1 = -\frac{n}{2}\ln(2\pi) - \frac{n}{2}\ln\sigma^2 - \frac{1}{2}\left(\frac{[\ln w - Z'\theta_1]}{\sigma^2}\right) - \frac{1}{w}, \tag{11}$$

$$\ln L_2 = -\left(e^{X'\theta_2+\gamma\ \exp(Z'\theta_1+\sigma^2/2)}\right) + y\ \ln\left(e^{X'\theta_2+\gamma\ \exp(Z'\theta_1+\sigma^2/2)}\right) - \ln y! - \ln(1 - e^{-e^{X'\theta_2+\gamma\ \exp(Z'\theta_1+\sigma^2/2)}}). \quad (12)$$

After estimating L_1 using MLE, the predicted harvest (λ_1) is inserted into L_2 and then the second model is estimated. The prediction for number of recreational trips now contains the information from the first step, but also contains the error associated with the prediction for harvest. Thus, the standard errors resulting from the estimation of Eq. (12) must be corrected. Finally, let V_1 be the variance–covariance for the auxiliary model (Eqs. (9) and (11)) and let V_2 be the variance–covariance matrix for the main equation (Eqs. (10) and (12)). Then the corrected variance–covariance matrix for the second step can be defined in terms of the asymptotic distribution for the two-step MLE.

As long as the standard regularity conditions are met for (11) and (12), then the second-step maximum likelihood estimator of θ_2 is consistent and asymptotically normally distributed with an asymptotic covariance matrix defined by

$$V_2^* = V_2 + V_2[CV_1C' - RV_1C' - CV_1R']V_2, \quad (13)$$

where V_1 is the asymptotic variance matrix of L_1 (Eq. (11)), and V_2 is the asymptotic variance matrix of L_2 given θ_1 (Eq. (12)). We can then define C and R as

$$C = E\left[\left(\frac{\partial \ln L_2}{\partial \ln \theta_2}\right)\left(\frac{\partial \ln L_2}{\partial \theta_1'}\right)\right], \quad R = E\left[\left(\frac{\partial \ln L_2}{\partial \theta_2}\right)\left(\frac{\partial \ln L_1}{\partial \theta_1'}\right)\right]. \quad (14)$$

The estimates for C and R can be obtained by summing the individual observations on the cross products of the derivatives (Greene, 2000, p. 135). In order to calculate V_2^* the partials and cross partials of L_1 and L_2 were taken, yielding:

$$\frac{\partial \ln L_1}{\partial \ln \theta_1} = \frac{1}{\sigma^2} Z'(\ln w - Z'\theta_1), \quad (15)$$

$$\frac{\partial \ln L_1}{\partial \sigma^2} = -\frac{n}{2\sigma^2} + \frac{1}{2\sigma^4}(\ln w - Z'\theta_1)'(\ln w - Z'\theta_1), \quad (16)$$

$$\frac{\partial \ln L_2}{\partial \theta_2} = -\lambda_2 X' + y\left(\frac{1}{\lambda_2}\right)\lambda_2 X' - \left(\frac{1}{(1 - e^{-\lambda_2})}\right)(-e^{-\lambda_2})(-\lambda_2)X', \quad (17)$$

$$\frac{\partial \ln L_2}{\partial \theta_1} = -\lambda_2 Z'\gamma\lambda_1 + y\left(\frac{1}{\lambda_2}\right)\lambda_2 Z'\gamma\lambda_1 - \left(\frac{1}{(1 - e^{-\lambda_2})}\right)(-e^{-\lambda_2})(-\lambda_2)Z'\gamma\lambda_1. \quad (18)$$

With the estimates for C and R, the corrected variance–covariance matrix for the second step model can be calculated. The Murphy–Topel Correction for the log-linear/truncated Poisson is thus:

$$R = \left[\frac{\partial \ln L_2}{\partial \theta_2}\right]\left[\frac{\partial \ln L_1}{\partial \theta_1}\right] = \left\{X'\left[-\lambda_2 + y\left(\frac{1}{\lambda_2}\right)\lambda_2 - \left(\frac{1}{(1-\mathrm{e}^{-\lambda_2})}\right)(-\mathrm{e}^{-\lambda_2})(-\lambda_2)\right]\right\} \times \left\{Z'\left[\frac{1}{\sigma^2}(\ln w - \theta_1)\right]\right\}, \quad (19)$$

$$C = \left[\frac{\partial \ln L_2}{\partial \theta_2}\right]\left[\frac{\partial \ln L_2}{\partial \theta_1}\right] = \left\{X'\left[-\lambda_2 + y\left(\frac{1}{\lambda_2}\right)\lambda_2 - \left(\frac{1}{(1-\mathrm{e}^{-\lambda_2})}\right)(-\mathrm{e}^{-\lambda_2})(-\lambda_2)\right]\right\} \times \left\{Z'\left[\gamma\lambda_1\left(-\lambda_2 + y\left(\frac{1}{\lambda_2}\right)\lambda_2 - \left(\frac{1}{(1-\mathrm{e}^{-\lambda_2})}\right)(-\mathrm{e}^{-\lambda_2})(-\lambda_2)\right)\right]\right\}. \quad (20)$$

We modify the above two-step method defined by Eqs. (6)–(20) to produce log-linear/Poisson, log-linear/truncated Poisson (as detailed above), Poisson/Poisson, and Poisson/truncated Poisson models.[9] The models are presented in the next section.

Results

Using the two-step derivations above, a total of eight models were created. Four of the eight models are presented in this section. We evaluate four different versions of the two-step models discussed above, each with a main (trips) and auxiliary equation (harvest): the log-linear/Poisson model with TCMUSH (LL-PM); log-linear/Poisson model without TCMUSH (LL-P); log-linear/truncated Poisson with TCMUSH (LL-TPM) model; and the log-linear/truncated Poisson without the TCMUSH variable (LL-TP).[10] The four models are presented in Table 2. In this section we examine the econometric results of our modeling.

The trip data in our sample have a mass centered around 1, with the frequency falling off significantly as the trip count increases; and by examining the mean (1.63) and the variance (1.61) of the variable VISITS the mean and variance do not exhibit over-dispersion.[11] This confirms the use of the Poisson count data model for

[9] The details of these derivations are provided in a separate appendix, available upon request from the authors.

[10] In estimating the two-step models, a total of eight versions were estimated. Only the four models using the log-linear harvest equation are presented here because of specification difficulties with the P-TP and P-P models, as discussed later in this section.

[11] A truncated negative binomial model was run in order to examine the effect of relaxing the Poisson assumption that the mean and variance of the dependent variable are equal. The parameter α, which tests whether the variance is significantly different from the mean, was estimated at 0.366, with a t-statistic of 1.985 (significant at the 0.05 level). Per trip consumer surplus was calculated at \$48.32, with a confidence interval of \$6.58. The variables EXPHRV, AGE, and INC were not statistically significant. However, TC, and TCMUSH were significant (0.01 level). The truncated negative binomial model has difficulty converging, and while several different algorithms and step-levels were tried, the LIMDEP likelihood

Table 2. Two-step estimation results (LHS = VISITS)

Independent variable	Models			
	LL-PM	LL-P	LL-TPM	LL-TP
CONSTANT	0.524***	0.612***	0.280	0.377*
	(2.498)	(3.047)	(0.865)	(1.289)
TC	−0.002**	−0.003***	−0.016***	−0.018***
	(−1.816)	(−2.530)	(−4.345)	(−5.451)
TCMUSH	−0.005*	—	−0.004	—
	(−1.568)	—	(−0.745)	—
MUSHDUM	0.660***	0.462***	0.879***	0.744***
	(4.339)	(4.654)	(3.948)	(5.595)
INC	0.001	−0.003	0.049*	0.047*
	(0.025)	(−0.164)	(1.511)	(1.452)
GENDER	−0.161**	−0.168**	−0.307***	−0.313***
	(−1.934)	(−2.029)	(−2.428)	(−2.482)
AGE	0.001	−0.003	0.022	0.017
	(0.049)	(−0.124)	(0.510)	(0.412)
EXPHRV	0.009*	0.010*	0.011**	0.011**
	(1.466)	(1.628)	(1.662)	(1.673)
Pearson's R^2	0.404	0.394	0.454	0.454
Deviance-based R^2	0.269	0.257	0.301	0.300
Log-likelihood	−541.797	−543.225	−387.315	−387.599
χ^2	67.181	64.325	376.144	375.576
Akaike's information criteria	0.112	0.112	0.080	0.080
Per trip consumer surplus	422.542	313.948	62.353	57.139
95% confidence interval	(190–655)	(190–438)	(48–77)	(47–68)
Recreational visitor day CS	227.939	169.358	33.636	30.823

Notes: [$N = 392$], with t-statistics, using corrected standard errors, in parentheses. *, **, ***, indicate significance at the 0.10, 0.05, 0.01 levels, respectively. LL-PM = log-linear/Poisson with TCMUSH; LL-P = log-linear/Poisson; LL-TPM = Log-linear/truncated Poisson with TCMUSH; LL-TP = Log-linear/Truncated Poisson. Consumer Surplus values are in 1996 US Dollars.

examining trip demand. However, it is not so clear what distribution approximates the HARVEST data. While the HARVEST data appear to be of a discrete nature, the process of harvesting should yield a continuous distribution of harvest. Thus, the

(*footnote continued*)
procedure did not converge. Additionally, the χ^2 value was 15.08 compared with 375.58 for the LL-TP model. The SFP data have a mean VISITS value of 1.63, and a variance of 1.61, indicating that the Poisson assumption that the mean and the variance are the same is appropriate. Given the difficulty with convergence for truncated negative binomial models, the statistical lack of significance on EXPHRV, a CS measure higher than the LL-TP model, a χ^2 value of only 15.08, and the mean-variance of VISITS—the truncated negative binomial distribution was rejected in favor of the Poisson.

apparent count nature of the data may be due to measurement error. In looking at the distribution of the HARVEST data, we find that the mean (4.96) and variance (224.96) are significantly dispersed, suggesting that a Poisson model is not appropriate for our harvest data. Examining the results of the auxiliary models, we find the log-linear predicts an average harvest of 4.96 gallons of mushrooms and berries while the Poisson predicts a value of 5.05 gallons; this compares with the sample mean of 4.96 gallons of harvest. Both the log-linear and Poisson harvest models are significant (0.01 level) and yield significant estimates of the coefficients; however, the Poisson has inflated standard errors due to the over-dispersion of the harvest data. For the log-linear model, the estimated coefficients for ONSITE, GRPSZ, EXPEND, MUSHDUM and EMPL are significant (0.01 level), while the RACE coefficient is statistically insignificant. The full results for the auxiliary regression for the log-linear harvest model are presented in Table 3. Based on the goodness-of-fit statistics, predicted harvest, the over-dispersion of the harvest data, and the significance levels of the estimated coefficients, we reject the Poisson harvest model in favor of the log-linear harvest model.[12]

Of the four models presented here, the log-linear/truncated Poisson (LL-TP) yields the most conservative estimate of per trip consumer surplus (CS) at $57.14. In comparison, the log-linear/Poisson (LL-P) yields a CS of $313.95. The difference between the LL-P and the LL-TP is $256.81, well outside the 95% confidence intervals for the regressions ($124.05 and $10.86, respectively). As these results indicate, using the standard Poisson distribution in the recreation demand model produces CS estimates considerably larger than those from the comparable truncated Poisson model.

In terms of estimated coefficients, the results are consistent across the models. All four models have estimated coefficients on travel costs (TC) that are negative and significant at the 0.05 level or better. In all models, the estimated coefficient on MUSHDUM is positive and significant (0.01 level), and the estimated coefficient on GENDER is negative and significant (0.05 level). Finally, the estimated coefficient on the predicted harvest EXPHRV is positive and significant (0.05 level) in all the models.

In terms of goodness-of-fit statistics, for all four models the chi-squared values from a likelihood ratio test indicate that the results are significant at the 0.01 level; Pearson's R^2 values range from a low of 0.394 for the LL-P model to a high of 0.454 for the LL-TP model.

After examining the goodness-of-fit statistics, the predicted seasonal trips and predicted harvest levels, the estimated consumer surplus values, and the significance of the estimated coefficients based upon the corrected standard errors, our results indicate that the log-linear/truncated Poisson (LL-TP) model provides the best fit to our data and yields consistent and efficient estimates of consumer surplus. These results are consistent with the data since our usable sample does not contain any zero trip values.

[12] The results of the four models using the Poisson auxiliary equation can be obtained from the authors upon request.

Table 3. Two-step estimation model results: auxiliary equation (LHS = HARVEST)

Variable	Coefficient
ONE	−0.013
	(−0.064)
ONSITE	0.007***
	(4.419)
GRPSZ	0.126***
	(5.439)
EXPEND	0.001***
	(2.447)
MUSHDUM	0.828***
	(6.637)
RACE	0.027
	(0.162)
EMPL	0.202**
	(2.212)
R^2	0.302
Adjusted R^2	0.291
F-statistic	27.790***
Predicted harvest	4.962

Notes: [$N = 392$], *t*-statistics in parentheses. *, **, ***, indicate significance at the 0.10, 0.05, and 0.01 levels, respectively. For the Poisson harvest model (not presented here) the predicted harvest was 5.05 gallons.

Thus, based on the LL-TP model we estimate a per trip consumer surplus value of \$57.14 from recreational harvesting trips to the Gifford Pinchot National forest in 1996 dollars. To convert our seasonal values to the more standard recreational visitor day (RVD) values, we calculate

$$\mathrm{CS_{RVD}} = \left(\frac{\mathrm{CS}}{\frac{14.84}{8}} \right) \tag{21}$$

for each of our models, where 14.84 is the mean time in hours on site, and 8 hours is the standard length of an RVD. Using the LL-TP model we estimate a $\mathrm{CS_{RVD}}$ of \$30.82 in 1996 dollars, or \$36.06 when adjusted for inflation to 2003 dollars.

In summary, the structural equations model provides an econometric correction of the endogeneity stemming from the inclusion of the predicted harvest in the recreation demand equation. The two-step procedure produces corrected standard errors that improve the efficiency of the estimated coefficients.

Discussion

Public forestlands provide a wide variety of beneficial uses, including numerous opportunities for recreational activities. To the extent possible, quantifying these recreational benefits will facilitate resource allocation decisions. Thus, estimates of non-timber values can be important inputs for evaluating tradeoffs in forest management policy. The current empirical literature lacks recreation demand estimates of berry and mushroom harvesting in the Pacific Northwest (USA) and elsewhere. In estimating the demand for non-commercial harvesting, the expected harvest level must be explicitly modeled. As Englin et al. (1997) have shown, the expected harvest can be endogenous to trip demand and failure to model trip demand and harvest explicitly can lead to biased and/or inefficient demand estimates.

In closing, to improve the efficiency of our estimates of the non-commercial wild berry and mushroom harvest from the GPNF, this study uses a structural equations approach with a Murphy–Topel standard error correction. From our preferred model, we estimate an RVD consumer surplus of $30.82 for wild berry and mushroom harvesting on the Gifford Pinchot National Forest in 1996 dollars, or $36.06 when adjusted for inflation to 2003 dollars. By way of comparison with previous Pacific Coast Area (USA) forest recreation demand studies (Rosenberger and Loomis, 2001, p. 13), our estimated RVD value of $30.82 (1996$) falls between the mean value for camping of $86.96 (1996$, with 4 observations) and the mean value for picnicking of $53.52 (1996$, with 3 studies).

References

Anderson, J.A., Blahna, D., Chavez, D., 2000. Fern gathering on the San Bernardino National Forest: cultural versus commercial values among Korean and Japanese participants. Society and Natural Resources 13, 747–762.

Arnold, M.J.E., Perez, M.R., 2001. Can non-timber forest products match tropical forest conservation and development objectives. Ecological Economics 39, 437–447.

Bockstael, N., Strand, I., Hanemann, M., 1987. Time and the recreational demand model. American Journal of Agricultural Economics 69, 293–302.

Cameron, C.A., Trivedi, P.K., 1986. Econometric models based on count data: comparisons and applications of some estimators and tests. Journal of Applied Econometrics 1, 29–53.

Chakraborty, K., Keith, J.E., 2000. Estimating the recreation demand and economic value of mountain biking in Moab, Utah: an application of count data models. Journal of Environmental Planning and Management 43, 461–469.

Creel, M.D., Loomis, J.B., 1990. Theoretical and empirical advantages of truncated count data estimators for analysis of deer hunting in California. American Journal of Agricultural Economics 72, 434–441.

Englin, J., Shonkwiler, J.S., 1995. Estimating social welfare using count data models: an application to long-run recreation demand under conditions of endogenous stratification and truncation. The Review of Economics and Statistics 77, 104–112.

Englin, J., Lambert, D., Shaw, W.D., 1997. A structural equations approach to modeling consumptive recreation demand. Journal of Environmental Economics and Management 33, 33–43.

Freeman III, M.A., 1993. The Measurement of Environmental and Resource Values: theory and methods. Resources for the Future, Washington, DC.

Godoy, R., Lubowski, R., 1992. Guidelines for the economic valuation of nontimber tropical-forest products. Current Anthropology 33, 423–433.

Gordon, S., 2002. Ecosystem comes first. News Tribune—Tacoma Washington, Sunday November 17th, 2002.
Gram, S., 2001. Economic valuation of special forest products: an assessment of methodological shortcomings. Ecological Economics 36, 109–117.
Greene, W.H., 2000. Econometric Analysis. Prentice-Hall, Englewood Cliffs, NJ.
Greene, W.H., 2002. Personal Communication (email).
Grogger, J.T., Carson, R.T., 1991. Models for truncated counts. Journal of Applied Econometrics 6, 225–238.
Hansis, R., 1998. A political ecology of picking: non-timber forest products in the Pacific Northwest. Human Ecology 26, 67–86.
Hellerstein, D., 1993. Intertemporal data and travel cost analysis. Environmental and Resource Economics 3, 193–207.
Hellerstein, D., Mendelsohn, R., 1993. A theoretical foundation for count data models. American Journal of Agricultural Economics 75, 604–611.
Kangas, K., 2001. Commercial wild berry picking as a source of income in northern and eastern Finland. Journal of Forest Economics 7, 53–68.
Loomis, J.B., 1993. Integrated Public Lands Management. Columbia University Press, New York.
Markstrom, D.C., Donnelly, D.M., 1988. Christmas tree cutting: demand and value as determined by the travel cost method. Western Journal of Applied Forestry 3, 83–86.
Mattson, L., Li, C.Z., 1994. How do different forest management practices affect the non-timber value of forests?—An economic analysis. Journal of Environmental Management 41, 79–88.
Murphy, K.M., Topel, R.H., 1985. Estimation and inference in two-step econometric models. Journal of Business and Economic Statistics 3, 88–97.
Ovaskainen, V., Mikkola, J., Pouta, E., 2001. Estimating recreation demand with on-site data: an application of truncated and endogenously stratified count data models. Journal of Forest Economics 7, 125–144.
Robinson, E.J.Z., Williams, J., Albers, H., 2002. The influence of markets and policy on spatial patterns of non-timber forest product extraction. Land Economics 78, 260–271.
Rosenberger, R.S., Loomis, J.B., 2001. Benefit transfer of outdoor recreation use values: a technical document supporting the forest service strategic plan. General Technical Report RMRS-GTR-72, US Department of Agriculture, Forest Service, Rocky Mountain Research Station, pp. 1–59.
Shaw, D., 1988. On-site samples' regression, problems of non-negative integers, truncation, and endogenous stratification. Journal of Econometrics 37, 211–223.
Simpson, D., 1999. The price of biodiversity. Issues in Science and Technology 15, 65–70.