Journal of Environmental Economics and Management 101 (2020) 102317

Contents lists available at ScienceDirect

Journal of Environmental Economics and Management

journal homepage: www.elsevier.com/locate/jeem

The impact of wildfires on the recreational value of heathland: A discrete factor approach with adjustment for on-site sampling☆

Anne Nobel [a,*], Sebastien Lizin [a], Nele Witters [a], Francois Rineau [b], Robert Malina [a,c]

[a] Research Group Environmental Economics, Hasselt University, Martelarenlaan 42, 3500, Hasselt, Belgium
[b] Center for Environmental Sciences, Hasselt University, Martelarenlaan 42, 3500, Hasselt, Belgium
[c] Laboratory for Aviation and the Environment, Department of Aeronautics and Astronautics, Massachusetts Institute of Technology, 77 Massachusetts Avenue, Cambridge, MA 02139, USA

ARTICLE INFO

Article history:
Received 13 June 2019
Received in revised form 11 March 2020
Accepted 12 March 2020
Available online 28 March 2020

Keywords:
Joint RP-SP
Revealed and stated preference
Travel cost
Contingent behavior

ABSTRACT

This study extends the finite mixture model proposed by Landry and Liu (2009) to control for on-site sampling in the context of a system of recreation demand equations. We apply the model to a quasi-panel of recreational trip data under current and hypothetical wildfire conditions. We use the estimated model to provide the first estimate of recreational value for heathland in a European protected area and for how wildfires affect this value. Although we find statistical support for our proposed on-site sampling correction, we find that the correction generates no significant impact on welfare measures in our application.

1. Introduction

1.1. Wildfire impacts on recreation demand

Uncontrolled wildfires pose an increasingly serious threat to a range of ecosystems (Depicker et al., 2020). The increasing severity and frequency of droughts, predicted throughout Europe, are expected to increase the frequency and spread of wildfires (Flannigan et al., 2016). Ecosystems that are susceptible to wildfires include those in temperate climate zones in Northwest Europe, of which European dry heathland (hereafter, heathland) is particularly susceptible (Schepers et al., 2014). Heathland is an important ecosystem that provides a variety of ecosystem goods and services (Fagúndez, 2013) and is classified as a very rare and valuable habitat type under the Habitats Directive (European Commission, 1992). Heathland areas provides valuable recreational opportunities because they are considered aesthetically attractive. Although periodic controlled burning of small patches of heathland is historically used to promote heathland conservation (Fagúndez, 2013), large and uncontrolled wildfires may affect their overall attractiveness.

The available evidence on the impact of wildfires on recreation demand is ambiguous. Wildfires have been found to increase recreation for two reasons: (1) Some visitors are curious to see the effects of a recent wildfire from up close (Sánchez

☆ AN, SL, and NW are all funded by fellowships of the Research Foundation Flanders (respectively 11G2319N, 12G5418N and 12B2917N).
* Corresponding author.
E-mail address: anne.nobel@uhasselt.be (A. Nobel).

https://doi.org/10.1016/j.jeem.2020.102317
0095-0696/

et al., 2016); and (2) other visitors want to see how the vegetation recovers in the years following a wildfire (Hilger and Englin, 2009). Other studies found that wildfires can result in lower visitation rates because they make natural areas less attractive for recreational activities (e.g. Hesseln et al., 2003). Several studies have also found that the wildfire intensity and the time since a wildfire have non-linear impacts on recreation demand (Englin et al., 2001; Vaux et al., 1984). In addition to the mixed evidence of wildfire impacts on recreation, the existing studies are strictly limited to recreation in forest ecosystems in the United States. The recreational values and value changes provided by these studies may not be accurate for applications in other geographical regions where ecosystems and recreational use are quite different (Schägner et al., 2016).

It is challenging to measure the impact of actual wildfires on recreation for two reasons. First, uncontrolled wildfires are extreme events that occur infrequently. Second, notwithstanding that wildfire risk can be estimated with historical data (Depicker et al., 2020), the exact timing and location of any given wildfire is not known upfront. Hence, it is difficult to observe actual changes of recreation demand after actual wildfires. One way to overcome this challenge is to collect revealed preference (RP) and stated preference (SP) data, where the latter provide anticipated visitation rates after hypothetical wildfires. In this case, recreation demand changes can be estimated by combining RP data about current recreation demand with stated preference SP data about recreation demand after a hypothetical wildfire. This approach can be implemented cost-effectively through on-site sampling of recreational users (Bateman et al., 2002). However, on-site sampling may result in biased and inconsistent parameter estimates if zero-truncation and endogenous stratification are not properly accounted for in estimation (Shaw, 1988).

1.2. Contribution of this study

This study makes two main contributions. First, this study combines a particular finite mixture model − the Discrete Factor Method−Generalized Negative Binomial (DFM−GNB) model − with on-site survey data in order to estimate a system of recreation demand equations while controlling for on-site sampling. Landry and Liu (2009) were the first to apply the DFM−GNB model to recreation demand data but did not correct for on-site sampling in estimation. Several studies have developed on-site sampling corrections for various model specifications (e.g. Egan and Herriges, 2006; Hynes and Greene, 2013; Moeltner; Shonkwiler, 2005). However, only Egan and Herriges (2006) incorporated on-site sampling adjustments in a (quasi-)panel data context. The main advantage of DFM-GNB compared to parametric models, such as the one used by Egan and Herriges (2006), is its computational tractability, which enables the inclusion of a larger set of hypothetical scenarios. Second, although some studies have estimated wildfire impacts on recreation for study areas outside the United States using SP data (e.g. Molina et al., 2019), none of these studies used a joint RP-SP approach. Hence, we are the first to estimate a joint RP-SP model for wildfire impacts in a study area outside the United States. We are also the first to provide estimates of the recreational value for heathland in a European protected area, and estimates of how wildfires of varying magnitudes affect this value, both immediately after a wildfire and in subsequent years.

The remainder of this paper is organized as follows. Section 2 introduces the theoretical framework used to quantify recreational value changes. Section 3 discusses the econometric challenges that arise from estimating recreational value changes with on-site survey data. This section also introduces a specification of the DFM-GNB model that controls for on-site sampling bias. Section 4 describes the collection of the on-site survey data and provides summary statistics. Section 5 presents the main regression results and discusses how controlling for on-site sampling bias affects the parameter and welfare results. Section 6 concludes.

2. Theoretical model

Welfare impacts caused by quality changes of public goods, such as recreational sites, cannot be inferred from markets, because markets for public goods generally do not exist. However, it is possible to derive welfare changes due to quality changes of public goods based on the demand for related private goods, as long as two necessary conditions are met (Bockstael et al., 1989): First, the consumption of a private good is an essential input for individuals to receive value from the quality of the public good. Second, changes of the quality of the public good have no effect on individuals if they do not consume the private good. Within the context of demand for public recreational sites, Bockstael et al. (1989) showed that the demand for quality changes of recreational sites can be inferred from visitation rates and the travel cost incurred by recreational visitors, because travel is a prerequisite for recreational experiences at a recreational site. Furthermore, individuals receive no value from quality changes of a recreational site if they do not use it for recreation.

In order to model the demand for quality of a recreational site, we start with the assumption that individuals maximize utility, denoted as U, within an income constraint, denoted as Y (Pudney, 1991). Furthermore, as we are interested in modeling the demand for recreational quality, we assume that two subsets of goods are available to individuals. The first subset consists of recreational trips x with a travel cost of p^x to a recreational site with quality Q, and the second consists of a vector of all other goods z with prices p^z. Formally, individuals are assumed to solve the following optimization problem:

$$\max_{x,z}[U(x, Q, z)|p^x x + p^z z = Y]. \tag{1}$$

The solution to the first-order conditions of this optimization problem yields the Marshallian demand function:

$$x^m = x^m(p^x, Q, z, Y, \beta), \tag{2}$$

where x^m is the Marshallian (uncompensated) demand for the recreational site of interest and β is a vector of parameters to be estimated. We do not explicitly model the demands for all other goods z and assume quasi-fixed values for p^z, so that an incomplete demand system arises. The advantage of this approach is that the demand for the quality of a good of interest can be modeled while being consistent with consumer behavior theory (von Haefen, 2002).

On the condition that the specification of x^m is integrable, the impact of quality moving from current level Q_0 to new level Q_1 on consumer surplus can be approximated by calculating the difference between the integrals of the Marshallian demand function before and after the quality change (Bockstael and McConnell, 1993).[1] In other words:

$$\Delta CS^m = \int_{p_0^x}^{p_c^x} x^m(p^x, Q_1, Y, \beta)dp - \int_{p_0^x}^{p_c^x} x^m(p^x, Q_0, Y, \beta)dp, \tag{3}$$

where ΔCS^m denotes the approximation of consumer surplus change based on the Marshallian, uncompensated demand function,[2] and the interval bounds p_0^x and p_c^x indicate the initial price and the price that causes demand to fall to zero, respectively.

3. Econometric estimation

3.1. Estimating quality change impacts on recreation demand

The nature of the data used to estimate recreation demand x^m for a recreational site of interest introduces three potential issues for the econometric estimation of x^m: (1) Heterogeneity across individuals in the sample; (2) Inconsistency between the preferences implied by different data sources; (3) On-site sampling bias.

The first issue arises from the fact that recreation demand x^m is estimated based on observations of recreation demand in function of varying prices, or travel cost. To obtain sample data with sufficient variation over a relevant price range, it is common to use a cross-sectional dataset in which travel cost varies *across* individuals. This introduces individual level observed and unobserved heterogeneity. Observed heterogeneity can be accounted for by including individuals' characteristics in the demand specification. However, any remaining unobserved heterogeneity across individuals implies that the demand curve cannot be estimated deterministically. Instead, the demand curve needs to be estimated using a probability density function that, in turn, can be linked to expected demand (Hellerstein and Mendelsohn, 1993). This probability density function should be defined for non-negative integer values only, because recreation demand is expressed in terms of trip frequencies (Hellerstein and Mendelsohn, 1993).

The issue of inconsistency between different data sources stems from the difficulty to obtain data about recreation demand for a recreational site under quality conditions other than the historically observed quality conditions. In particular, researchers increasingly rely on combinations of revealed preference (RP) data to estimate recreation demand under current quality conditions and stated preference (SP) data to estimate recreation demand under hypothetical conditions (Adamowicz et al., 1994). Combining RP and SP data has attracted considerable interest in the literature, partly because this approach enables researchers to ground SP behavior in actual (RP) behavior, while extending the range of environmental quality beyond what can be observed in real life. An advantage of employing RP and SP data in a (quasi-)panel setting is that it increases the efficiency of parameter estimates (Ben-Akiva et al., 1994). However, approaches that combine RP and SP data are based on the assumption that both data sources describe the same preferences (Cherchi and Ortúzar, 2011). Thus, we would expect that preferences stated by individuals are consistent with the preferences they revealed in the real world, ceteris paribus. However, many studies that combined RP and SP sources have found that parameters calibrated on RP data are significantly different from parameters calibrated on SP data (Cherchi and Ortúzar, 2011). On one hand, the inconsistency between RP and SP parameters is sometimes treated as proof that RP data are more valid than SP data because the former involve actual behavior. In particular, preferences implied by anticipated (SP) trip behavior may be biased compared to preferences implied by actual (RP) trip behavior, among other reasons because respondents are not familiar with the hypothetical quality change, or because respondents ignore their budget constraint (Whitehead et al., 2008). Another explanation is that respondents anticipate an income increase in the future, resulting in a different marginal utility of income (Whitehead et al., 2011). The inconsistency can also be caused by uncertainties about the future, such as uncertainty about family size or place of residence. On the other hand, RP data may be biased due to measurement error or recall error. Hence, it is unknown whether RP or SP data are more valid in the case of parameter inconsistency. We deal with the potential

[1] This difference equals the area between the before and after quality-change uncompensated demand curves bounded by the current price and the price that lets demand fall to zero.

[2] Willig (1978) derived the conditions under which observable demand can also be used to evaluate Hicksian (compensated) welfare measures. Hicksian demand differs from Marshallian demand because real income is held constant in its calculation.

inconsistency between RP and SP data by defining a system of equations with separate RP and SP parameters and error terms, resulting in the following quasi-panel of recreational demand[3]:

$$x_{it}^{rp} = f\left(S_{it}, \phi_{rp}, \varepsilon_{it}^{rp}\right), \tag{4}$$

$$x_{it}^{sp} = f\left(S_{it}, \phi_{sp}, \varepsilon_{it}^{sp}\right), \tag{5}$$

where x_{it} is the annual number of trips of individual i under treatment t, S_{it} represents a vector of observed individual and site quality characteristics, ϕ_{rp} and ϕ_{sp} parameter vectors, and ε_{it}^{rp} and ε_{it}^{sp} are equation-specific error terms. The consistency between the preferences implied by RP and SP data sources can be tested by comparing a constrained model with parameters that are equal across RP and SP equations with an unconstrained model in a likelihood ratio test (LRT). Importantly, accurate estimation of the system of equations in Eqs. (4) and (5) requires that the econometric model should allow for demand equation-specific error terms, while allowing for correlation between the error terms because revealed and stated trips are likely to be correlated for a given individual (Whitehead et al., 2011).

A common way to deal with count data from two data sources is to start with two separate count distributions − that is, Poisson or negative binomial distributions − and add a form of randomness to the conditional mean of each of the distributions (Herriges et al., 2008). Correlation across the conditional means of these distributions is induced by allowing the randomness to come from a common source, such as a variance-covariance matrix or a latent class framework. An example of the first approach is the Multivariate Poisson Log-Normal (MPLN) model (Egan and Herriges, 2006). The MPLN model consists of a Poisson distribution, in which a normally distributed error term is included in the conditional mean. The resulting model is an example of a mixing distribution that can be estimated via simulation. To avoid likelihood optimization problems, it is necessary to impose restrictions on the number of treatments or the correlation structure of such mixture models (Egan and Herriges, 2006). However, an advantage of MPLN is that it may provide more efficient estimates if its distributional assumptions hold (Landry and Liu, 2009). An example of the second approach is the Discrete Factor Method−Generalized Negative Binomial (DFM−GNB) model. As we will explain below, DFM−GNB belongs to the class of finite mixture models, which are now widely used for recreation demand analysis (e.g. Thiene et al., 2007). DFM−GNB assumes a multivariate generalized negative binomial distribution in which the demand equations include an additive, discrete random error term that is composed of two factors. In particular, the error term (ε_{it}) is modeled as the product of a demand equation-specific factor γ that is identical across individuals, and a factor λ_k that is identical across demand equations, but specific for the individuals who belong to heterogeneity class k. In other words, the randomness added to the conditional mean has the form $\varepsilon_{i,RP} = \gamma_{RP}\lambda_k$ for the RP demand equation, and $\varepsilon_{i,SP} = \gamma_{SP}\lambda_k$ for the SP demand equation.[4] To date, the DFM−GNB model has been applied to beach recreation data only (Landry and Liu, 2009). In this paper, the DFM−GNB model is used because it enables us to include a relatively large number of treatments with different quality change magnitudes to cover a range of possible wildfire outcomes.[5] The drawback of the DFM−GNB model compared to normality-based estimators is that it may produce less efficient estimates if data are normally distributed (Whitehead et al., 2011). We discuss the implications of this drawback, and of choosing DFM−GNB for climate change impacts on recreation in general, in the discussion.

3.2. Discrete Factor Method−Generalized Negative Binomial model

The DFM−GNB model consists of a generalized negative binomial distribution combined with the discrete factor method proposed by Heckman and Singer (1984). The probability mass function of the univariate DFM−GNB is specified as follows (Landry and Liu, 2009):

$$g(x_{it}|S_{it}) = \sum_{k=1}^{K}\left(\Pr(\lambda_k)\,\frac{\Gamma\left(x_{it}+a_t^{-1}\mu_{itk}^{2-p}\right)a_t^{x_{it}}\mu_{itk}^{(px_{it}-2x_{it})}\left(1+a_t\mu_{itk}^{p-1}\right)^{-\left(x_{it}+a_t^{-1}\mu_{itk}^{2-p}\right)}}{\Gamma(x_{it}+1)\Gamma\left(a_t^{-1}\mu_{itk}^{2-p}\right)}\right), \tag{6}$$

where $\Pr(\lambda_k)$ denotes the probability that an individual belongs to class k, Γ denotes the gamma distribution, x_{it} is the demand for trips of individual i under treatment t, μ_{itk} is the expected conditional demand for trips of individual i under treatment t if he or she belongs to heterogeneity class k, a_t is a treatment-specific parameter, and p is a dispersion parameter. This distribution can be extended to a multivariate distribution in which RP trip counts under current conditions are calculated with

[3] For simplicity, we suppress the superscript m in x^m and CS^m in the remainder of this paper.

[4] This approach differs from the latent class approach of Hynes and Greene (2013) in two respects. First, Hynes and Greene incorporate individual heterogeneity by allowing slope coefficients to differ between classes. Second, they pool outcome data from different data sources in a single demand equation. Hence, they implicitly assume that the magnitude of the impact of observed and unobserved heterogeneity on recreation demand is identical across data sources.

[5] To our knowledge, MPLN has been applied in recreation demand studies with up to four treatments (Egan and Herriges, 2006; Landry and Liu, 2009).

an RP equation, and SP trip counts under hypothetical treatments are calculated with an SP equation. Expected demand, conditional upon observed individual and site quality characteristics S_{it} and unobserved heterogeneity related to k class membership is specified as:

$$\mu_{itk} = \exp(S_{it}\beta + \varepsilon_{it}) = \exp(S_{it}\beta + \gamma_t\lambda_k) = \exp(S_{it}\beta)\exp(\gamma_t\lambda_k). \tag{7}$$

Let the deterministic part of the demand equation in Eq. (7) be denoted as $\delta_{it} = \exp(S_{it}\beta_t)$; then, the variance of x_{it} for a generalized negative binomial distribution with random parameter λ and dispersion parameters a and p is given by Landry and Liu (2009) as:

$$var(x_{it}|S_{it}) = \exp(\delta_{it} + \gamma_t\lambda) + a_t\exp(\delta_{it} + \gamma_t\lambda)^p. \tag{8}$$

The treatment-specific factors in Eq. (7) and Eq. (8) are restricted to be equal across SP trip counts, so that the scale of unobserved heterogeneity differs across RP and SP demand equations only and not across different SP trip counts. The number of classes and to which class each individual belongs are not known upfront. Hence, the probability of k class membership $\Pr(\lambda_k)$ is estimated empirically. The probability density function is then summed up over all possible classes. Hence, the resulting probability distribution is a finite mixture model. The likelihood contribution of individual i is:

$$g.(x_i|S_i.) = \sum_{k=1}^{K}(\Pr(\lambda_k)\cdot\prod_{t=1}^{T}\frac{\Gamma\left(x_{it} + a_t^{-1}\mu_{itk}^{2-p}\right)a_t^{x_{it}}\mu_{itk}^{(px_{it}-2x_{it})}\left(1 + a_t\mu_{itk}^{p-1}\right)^{-\left(x_{it}+a_t^{-1}\mu_{itk}^{2-p}\right)}}{\Gamma(x_{it}+1)\Gamma\left(a_t^{-1}\mu_{itk}^{2-p}\right)}. \tag{9}$$

where the estimates of heterogeneity class factors λ_k are confined to the interval [0,1]. Furthermore, the probabilities of individuals belonging to heterogeneity class k, denoted as $f(\theta_k)$, are transformed to a multinomial logit, so that the probabilities sum up to 1. It is also possible to incorporate class membership predictors into $f(\theta_k)$. However, such a specification is less parsimonious than the DFM approach because of the marked increase of the number of parameters to be estimated even with a small number of latent classes. Furthermore, the identification of the parameters of latent class models with class membership covariates can be problematic because the same covariates may show up in both the recreation demand equation and the class probabilities (Landry and Liu, 2009).[6]

3.3. Correcting for on-site sampling bias

The third econometric issue stems from the fact that RP and SP data about recreation demand for a site of interest are often obtained by means of on-site surveys (Egan and Herriges, 2006). This sampling method introduces endogenous stratification and zero-truncation. Endogenous stratification means that frequent visitors are over-represented in the sample, and zero-truncation means that the sample does not include non-visitors because they do not visit the site (Egan and Herriges, 2006). In more general terms, endogenous stratification and zero-truncation arise because the probability of an individual being included in the sample depends on the value of the response variable (that is, the number of trips), and failure to correct for the resulting sampling bias can result in biased and inconsistent welfare estimates (Englin and Shonkwiler, 1995).

Shaw (1988) showed that an on-site probability mass function can be obtained by assuming that the probability of an individual being included in the on-site sample is proportional to his or her actual trip frequency, relative to the conditional mean. In this case, the conditional on-site probability mass function $h(\cdot)$ can be defined as:

$$h(x_i|S_i) = \frac{x_i\cdot g(x_i|S_i)}{\sum_{v=1}^{m}v\cdot g(v|S_i)}, \tag{10}$$

where x_i denotes actual trip frequency of a visitor, v denotes all possible values for x_i and $g(\cdot)$ denotes the probability that an individual drawn from the general population made y_i trips during a specified, historical period. In our case, $g(\cdot)$ is generated by the mixing distribution in Eq. (6). This means that we can obtain the on-site probability mass function conditional on S_i by integrating the probability mass function over λ_k for $k = 1, .., K$, so that Eq. (10) becomes:

[6] The latent class framework can be used to relate latent class membership to individual characteristics. In order to provide a point of comparison for the DFM approach we also estimated regression models using the traditional latent class approach (see Regression results section).

$$h(x_{i,RP}|S_{i,RP}) = \frac{x_{i,RP}\cdot g(x_{i,RP}, S_{i,RP})}{\sum_{k=1}^{K}(f(\theta_k)\cdot\sum_{v=1}^{m}(v\cdot g(v|S_{i,RP})))} = \frac{x_{i,RP}\cdot g(x_{i,RP}, S_{i,RP})}{\sum_{k=1}^{K}(f(\theta_k)\cdot\mu_{i,RP,k})} = \frac{x_{i,RP}\cdot g(x_{i,RP}|S_{i,RP})}{\mu_{i,RP}}, \tag{11}$$

where subscript RP emphasizes that $h(\cdot)$ and $g(\cdot)$ are the probability mass functions of revealed, actual trips. The third line of Eq. (11) shows that $h(\cdot)$ is a function of the general population probability mass function $g(\cdot)$ and conditional expectation $\mu(\cdot)$.

Implementing Eq. (11) in Eq. (6) yields the following likelihood contribution for individual i:

$$h(x_{i,RP}|S_{i,RP}) = \frac{x_{i,RP}}{\mu_{i,RP}}\sum_{k}^{K}(\Pr(\lambda_k)\frac{\Gamma\left(x_{i,RP}+a_{RP}^{-1}\mu_{i,RP,k}^{2-p}\right)a_{RP}^{x_{i,RP}}\mu_{i,RP,k}^{(px_{i,RP}-2x_{i,RP})}\left(1+a_{RP}\mu_{i,RP,k}^{p-1}\right)^{-\left(x_{i,RP}+a_{RP}^{-1}\mu_{i,RP,k}^{2-p}\right)}}{\Gamma(x_{i,RP}+1)\Gamma\left(a_{RP}^{-1}\mu_{i,RP,k}^{2-p}\right)}. \tag{12}$$

Based on Patil and Rao (1978), weighting the population probability mass function $g(\cdot)$ according to Eq. (12) leads to the following relationship between the expected mean of the on-site sample and its population counterpart:

$$E(\tilde{x}_{RP}|S_{RP}) = E(x_{RP}|S_{RP}) + \frac{var(x_{RP}, S_{RP})}{E(x_{RP}|S_{RP})}, \tag{13}$$

where $\tilde{x}$ denotes the on-site sample mean. Eq. (13) shows that the on-site sampling bias of trips depends on the degree of over-dispersion in the population distribution, which is the ratio of the conditional variance and conditional mean.

The above equations can be used to obtain population means for RP trips in the univariate case. However, in some studies, including the present study, multiple RP and SP observations are obtained for each individual. If these observations are independent, then the SP observations are not affected by on-site sampling bias because the sample inclusion probability for SP observations would not depend on RP observations. However, SP trips may be correlated with RP trips, hence potentially carrying over some of the on-site sampling bias from RP trips to SP trips. Intuitively, this makes sense because frequent visitors are also more likely to report higher trip counts in the future.

Analogous to Eq. (13), the relationship between the expected sample mean for SP trips and the population is given by Egan and Herriges (2006):

$$E(\tilde{x}_{RP}|S_{RP}) = E(x_{RP}|S_{RP}) + \frac{var(x_{RP}, x_{RP})}{E(x_{RP}|S_{RP})}. \tag{14}$$

Eq. (14) shows that the magnitude of the on-site sampling bias of SP trips depends on the covariance of SP and RP trips.

In our specification, RP and SP trips are dependent because class membership is fixed over time. Hence, unobserved heterogeneity associated with each class is a common factor in the specification of both conditional means, that is, x_{RP} and x_{SP}. However, conditional on class membership, the covariance between the jointly distributed random variables is zero. This means that the covariance between x_{RP} and x_{SP} comes only from the specification of the conditional means given in Eq. (7) (analogous to Proposition I in Gurmu and Elder (2012)):

$$cov(x_{RP}, x_{SP}|S.) = cov(\delta_{RP}\exp(\gamma_{RP}\lambda), \delta_{SP}\exp(\gamma_{SP}\lambda)). \tag{15}$$

Substituting Eq. (8) and Eq. (15) in Eq. (13) and Eq. (14), respectively, gives the following expression of the on-site sample means:

$$E(\tilde{x}_t|S_t) = \begin{cases} E(x_1|S_1) + \dfrac{\exp(\delta_1+\gamma_{RP}\lambda) + a_1\exp(\delta_1+\gamma_{RP}\lambda)^p}{E(x_1|S_1)}, t=1 \\ E(x_t|S_t) + \dfrac{\delta_1\delta_t cov(\exp(\gamma_{RP}\lambda), \exp(\gamma_{SP}\lambda))}{E(x_t|S_t)}, t\neq 1. \end{cases} \tag{16}$$

The first line of Eq. (16) shows that if $a_1 = 0$, there is no over-dispersion in RP trips; that is, the conditional variance is equal to the conditional mean. In this case, the on-site sample mean becomes one more than the population mean. The second line of Eq. (16) shows that when the correlation between RP and SP trips increases – that is, when $cov(\exp(\gamma_{RP}\lambda), \exp(\gamma_{SP}\lambda))$ increases – the on-site sample mean of SP trips becomes increasingly upward biased from the population mean.

Now that we have shown how on-site sampling bias can affect both RP and SP trips, we extend the correction of the univariate probability mass function to the multivariate case. Similar to Eq. (11), the on-site sample probability mass function becomes:

$$h(x_{i\cdot} \mid S_{i\cdot}) = \frac{x_{i,RP}}{\mu_{i,RP}} g(x_{i\cdot} \mid S_{i\cdot}). \tag{17}$$

Finally, this results in the following joint likelihood function for the on-site sample:

$$L = h(x_{\cdot} \mid S_{\cdot}) = \prod_{i=1}^{N} \prod_{t=1}^{T} (h(x_{it} | S_{it}, \beta, \gamma, \lambda, \theta, \alpha, p)), \tag{18}$$

where β (for observable heterogeneity), γ, λ and θ (for unobserved heterogeneity), and α, p (for model dispersion) are the model parameters to be estimated.

3.4. Model estimation

Expectation maximization (EM) algorithms (Allman et al., 2019) and quasi-Newton algorithms (see e.g. Landry and Liu, 2009), are commonly used to estimate the model parameters of finite mixture models. EM algorithms employ an iterative procedure in which conditional probabilities are constructed in a first expectation step, and the model parameters related to class membership probabilities and the other model parameters are separately maximized in a second step (von Haefen and Domanski, 2018). Quasi-Newton algorithms use an approximation of the Hessian matrix to find a local minimum of an objective function. In the present study, we use the Bayesian Adaptive Direct Search (BADS) algorithm. The BADS algorithm is a general optimizer that estimates all parameters of a model simultaneously by maximizing an objective function. We use the BADS algorithm to directly maximize the unconditional full likelihood function by estimating all model parameters. The BADS algorithm belongs to the class of direct search algorithms. These algorithms sequentially compare candidate solutions with the current best solution in order to determine the next candidate solution (Acerbi and Ma, 2017). BADS performs direct searches by alternating between two steps. First, BADS performs a local optimization step in which solutions in the neighborhood of the current parameter values are randomly selected and evaluated. The algorithm is able to escape locally optimal solutions due to the Gaussian nature of this search process. As soon as the objective value does not improve any further, BADS switches to the polling step. In this step, a set of candidate solutions are selected along a set of polling directions and evaluated. These points lie on a mesh grid that increases in size when a better solution is found and decreases in size when no improvement is found. The algorithm continues to alternate between the two steps until the improvements of the objective function are arbitrarily small for a pre-defined number of iterations. The polling step guarantees theoretical convergence to an optimum (Audet and Dennis, 2006), so that BADS converges to maximum likelihood parameter estimates.

When finite mixture models are concerned, the advantage of using BADS algorithm is that it is theoretically more likely to converge to a global optimum than EM and quasi-Newton algorithms because it continues searching for better solutions after reaching a local optimum or saddle point. EM and quasi-Newton algorithms stop searching when a local optimum or saddle point is found because the zero gradient condition is met in these points. This is important because the unconditional likelihood function of finite mixture models is non-convex and possibly contains multiple local optima. However, depending on the stopping criteria, such as the maximum number of iterations, it is possible that the BADS algorithm does not reach global convergence within the allocated computational time. Hence, the use of the BADS algorithm reduces but not fully eliminates the need to start from different initial parameter values. It should be noted that the benchmark results reported by Acerbi and Ma (2017) show that the BADS algorithm may converge slightly faster than other quasi-Newton and direct search algorithms when the number of unknown model parameters is large. However, speed comparisons of the BADS algorithm and the EM algorithm – which is frequently reported to be slow for models with many parameters (Jollois and Nadif, 2007) – remain an avenue for future research.

The BADS algorithm has two drawbacks when finite mixture models are concerned. First, the BADS algorithm requires the specification of optimization constraints in order to assure the identification of meaningful parameter estimates. In particular, the weights of the mixture distribution need to be constrained to be positive and sum up to one. Hence, the BADS algorithm may be considered less elegant compared to the EM algorithm which automatically produces valid parameters for the mixture distribution in the Expectation step. Second, the stochastic search process of the BADS algorithm causes optimization results to *slightly* differ between runs even if the same starting values are used. In order to assess whether our results are sensitive to the algorithm used in estimation, we compare the log likelihood and parameter results obtained with the BADS algorithm with those obtained with the EM algorithm.

3.5. Model selection

The on-site corrected likelihood function leads to a model with maximum likelihood parameter estimates that are different from those obtained with an uncorrected likelihood function. When comparing non-nested models, the standard likelihood ratio test (LRT) cannot be used because the limiting distribution of the standard LRT will generally not be a chi-squared distribution under the null hypothesis (Lewis et al., 2011). Alternatively, the Akaike Information Criterion (AIC) (Akaike, 1973) and Bayesian Information Criterion (BIC) are often suggested for non-nested model comparison (Huang, 2017). Models with lower AIC and BIC values indicate higher support compared to models with higher AIC and BIC values. Furthermore, a Vuong non-nested likelihood ratio test (Vuong, 1989) can be applied to formally test if one of two non-nested models is more appropriate than the other.[7] In this study we employ the Vuong test to assess whether controlling for on-site sampling bias leads to a significant model improvement. Specifically, under the null hypothesis that the corrected and uncorrected models fit equally well to the data, the contributions of individuals to the log likelihood ratio follow a normal distribution with a mean of zero (Hynes and Greene, 2013). A large positive Vuong statistic beyond a critical t value then indicates that the corrected model is more favorable than the uncorrected model.

4. Data

4.1. Study site

The data used in this study come from an on-site survey among visitors of a heathland area in the Hoge Kempen National Park (hereafter also called HKNP) in Belgium (N 50 58.674, E 5 39.555). The HKNP, which received over 1 million visitors in 2013, hosts a variety of recreational activities such as walking, cycling, jogging, mountain biking, and horseback riding (Agentschap Natuur en Bos, 2016). Within the HKNP, an area of nearly 6000 ha of nature is protected under the EU Habitats Directive. Twenty percent of this protected area is heathland. The largest patch of heathland in the HKNP is located within walking distance from the main access point, which is used by the majority of visitors. This heathland is located in an area with a high-to-very-high wildfire risk (Depicker et al., 2020). The most recent major wildfire in this heathland occurred in 1996, when 200 ha of heathland vegetation was burned.

4.2. Survey design

We define six treatments, two of which are ‘no quality change’ treatments and four are hypothetical quality change treatments. The two ‘no quality change’ treatments refer to the current situation without and with a hypothetical travel cost increase, respectively. To reduce the chance of protest responses caused by an unfamiliar or unrealistic payment vehicle, we choose the payment vehicles for the travel cost increase to be identical to how access fees are collected for entrance to another part of the National Park. In particular, visitors are asked to choose between an entrance fee of EUR 3 per day trip and an annual subscription pass of EUR 5 per person.[8]

The first hypothetical quality change treatment relates to the calendar year that comes after the year in which the survey is conducted. In this treatment, a hypothetical wildfire occurs. As wildfires are most likely to ignite in April of any given year[9] (Depicker et al., 2020), we assume, for this treatment, that a spring wildfire causes 50 percent of the heathland to be burned, leaving an area with blackened *Calluna Vulgaris* plant material.

In the year following a wildfire, grass species such as *Molinia Caerulea* encroach post-burn heathland (Schepers et al., 2014). It can take up to 10 years before the heather plant has fully grown back and the grasses have disappeared (Hobbs and Gimingham, 1984). To measure recreation demand during this recovery period, we define a second hypothetical quality change treatment in which 50 percent of the heathland is fully covered by grasses throughout the year. This treatment takes place in the year after the first treatment.

The third and fourth hypothetical quality change treatments are similar to the first and second hypothetical quality change treatments, respectively, in terms of physical change and timing, but the affected area is 100 percent instead of 50 percent. We have two reasons for defining these treatments. First, the fragment size of heathland areas is important for their recreational value (Cordingley et al., 2015), implying that including affected areas of varying size may reveal scope or contrasting effects on recreation demand.[10] Second, grass encroachment can reduce the aesthetic value of heathland (Van Marwijk et al., 2012), but a 50 percent grass cover can also increase the aesthetic value of the heathland area if it causes an increase in overall landscape diversity (Schägner et al., 2016). Thus, defining 50 percent and 100 percent treatments may potentially reveal non-linear or contrasting effects of the affected area on recreational value.

We include the treatments in a paper-based survey that is structured as follows. First, respondents are asked a general ‘warm-up’ question about the purpose of their trip. Second, respondents are asked for travel cost details. Third, recreation demand under a ‘no quality change’ treatment without hypothetical travel cost increase is elicited by asking the respondents about their actual number of visits to the heathland area during the past 12 months. Respondents are then asked about their

[7] The Vuong test has been applied previously to test for on-site correction of recreation demand models (Beaumais and Appéré, 2010; Landry and Liu, 2009; Moeltner; Shonkwiler, 2005).

[8] We divide the annual amount by the annual number of trips, so that it can be included in the travel cost calculation according to Eq. (19).

[9] More than 30 percent of the wildfires between 1994 and 2016 in Belgium ignited in the month of April (Depicker et al., 2020).

[10] For example, Vaux et al. (1984) found that less intense wildfires have a positive impact on demand for recreation in Californian forests, while intense wildfires have a negative impact.

Table 1
Survey versions.

	Survey version											
	1	2	3	4	5	6	7	8	9	10	11	12
Hypothetical quality change treatment												
50% Grass cover in year+2 (dummy)	A	B	A	B	A	B						
100% Grass cover in year+2 (dummy)	B	A					A	B	A	B		
50% Burned cover in year+1 (dummy)			B	A			B	A			A	B
100% Burned cover in year+1 (dummy)					B	A			B	A	B	A

Table 2
Summary of contingent behavior section.

	Question 1. How often did you visit the *Mechelse Heide* in the past 12 months?	Question 2. How often would you visit the *Mechelse Heide* in 2019 if it looks as described in Scenario 1, in which 50 percent of the area is burned?[a]	Question 3. How often would you visit the *Mechelse Heide* in 2020 if it looks as described in Scenario 2, in which 50 percent of the area is covered with grasses?[a]	Question 4. How often would you visit the *Mechelse Heide* in 2019 if you had to pay for access in the way you indicated above[b]?
Response format	__ visits per week __ visits per month __ visits per year	__ visits per week __ visits per month __ visits per year	__ visits per week __ visits per month __ visits per year	__ visits per week __ visits per month __ visits per year

[a] The hypothetical quality changes are randomly selected and ordered from a total of four hypothetical quality change treatments.
[b] Respondents indicated in a preceding question whether they prefer to pay for a day pass (EUR 3 per visit) or an annual subscription pass (EUR 5 per year).

anticipated annual number of visits under two hypothetical quality change treatments. These treatments, say A and B, are randomly selected from a total of four hypothetical quality change treatments and randomly ordered, leading to a total of 12 survey versions (see Table 1). The random ordering of the hypothetical quality change treatments is intended to reduce ordering effects, if any.[11] Finally, respondents are asked about anticipated number of visits during the next year if an access fee is imposed, without any site quality changes. The contingent behavior section of the survey is summarized in Table 2.[12]

The treatments are presented as textual descriptions, supported with pictures, in order to facilitate the respondent's understanding of the treatments, taking seasonal variations into account (a full example of the treatment descriptions is provided in Online Online Appendix 1). It is explicitly stated that the surroundings of the heathland area and substitute sites are not subject to wildfires or other changes. To increase credibility, respondents are also made aware of the most recent wildfire in the study area, and of the fact that such wildfires may happen again in the future. Fourth, respondents are asked to provide demographic details. Fifth, two debriefing questions are asked in order to identify protest responses to the hypothetical travel cost increase (Loomis et al., 2011).

4.3. Data collection

The survey was pilot-tested with 40 park visitors in April 2018 and subsequently improved.[13] Survey data were collected at the *Mechelse Heide* access point by two interviewers on 25 days between May and September 2018. Individual visitors and groups with up to six members who intended to exit the *Mechelse Heide* were randomly asked to participate in the survey, after ensuring that they are proficient in Dutch. In total, we approached 562 individuals, 364 of whom (64.8 percent) were willing and able to fill out the questionnaire. The completed questionnaires were entered into a database. Individuals were removed from this database for any of four reasons,[14] listed here in chronological order: (1) They did not provide demographic details ($n = 14$) or travel cost details ($n = 3$); it is not possible to calculate recreation demand without these details; (2) They did not provide answers to all the contingent behavior questions ($n = 16$); the inclusion of these responses would make the panel unbalanced, requiring modifications of the likelihood function (the total number of incomplete responses, including the respondents mentioned under the first bullet point, is 26); (3) They visited the heathland area every other day ($n = 11$) or

[11] One may argue that the fixed position of the first (current situation) and last (access fee) treatments can still cause some ordering effects. However, the paper-based survey format allows respondents to see all contingent behavior questions prior to providing the responses. The advance disclosure of the contingent behavior questions may reduce potential ordering effects (Bateman et al., 2002). We test whether visitation rates reported under the first presented hypothetical quality change scenario are different from those reported under the second scenario by pooling these responses and estimating a recreation demand model that includes a "First scenario" dummy variable (see Section 5).
[12] A complete translated version of this survey section is included in Online Appendix 1.
[13] e.g. in the final survey we emphasized the unit of the payment by writing it in bold and we removed two 'warm-up' questions to reduce the total length of the survey.
[14] The excluded individuals do not differ significantly from the individuals remaining in the sample in terms of income ($p = 0.41$), age ($p = 0.25$), membership of environmental organizations ($p = 0.28$), or gender ($p = 0.31$); educational attainment is significantly lower among excluded individuals ($p = 0.06$).

daily (n = 15).[15] We assumed that these visitors derive little or no recreational value from the presence of heathland because they routinely visit the National Park mainly because of its close proximity to their homes. Approximately two-thirds of these visitors indicated "letting the dog out" as the purpose of their visit; (4) Informed consent could not be assumed because the individual had not reached the age of 18 (n = 1). We also removed one respondent, whose inconsistent answers (the respondent stated a one-way travel time by car of 2 h for a distance of less than 5 km) impeded the travel cost calculation. We also excluded 18 individuals because they provided a protest response; that is, they indicated a zero visitation rate if they would have to pay an entrance fee and their motivation for doing do so was identified as a typical protest response (Meyerhoff and Liebe, 2008). Of the remaining respondents, 93 did provide details about their employment status, but not about their household income bracket. Based on their reported employment status (employed, business owner or manager, or unemployed) we replaced missing income values with the sample mean income of the respondents with the same employment status.[16] We investigated whether recreation demand differs between respondents who did not report their income compared to other respondents by including a *missing income* dummy variable in the regression model. Consequently, the econometric analysis was conducted on a final dataset of 1172 observations from 293 individuals.

4.4. Descriptive statistics of sample data

Summary statistics of the travel cost determinants are provided in Table 3. We used these data to calculate the travel cost for each respondent (p_i^x) as follows:

Table 3
Summary statistics of travel cost determinants (n = 293).

	Mean	Std deviation	Minimum	Maximum	Median
Travel cost element[c]					
Stated relative importance of heathland for decision to make the current trip (on a seven-point scale, where one is totally unimportant and seven is most important)	5.27	1.68	1	7	6
Mean group size (number of adults)	3.48	4.05	1	60	2
One-way travel time[a] (minutes)	25.34	24.91	0[b]	210	20
One-way travel distance (fastest route between postal code centroid and *Mechelse Heide* access point, in kilometers)	21.61	31.93	0[b]	200.80	12.15

	Number of respondents	Percentage (%)
Travel mode		
Car	241	82
Small cars	*50*	
Medium cars	*135*	
Large cars	*58*	
By foot	23	7.8
Bicycle	24	8.2
By horse	1	0.3
Motorbike	2	0.7
Public transport	2	0.7

	Fuel cost per kilometer (in 2018 EUR)
Car size category	
Small cars	0.07
Medium cars	0.08
Large cars	0.11

[a] We used self-reported travel times. Traffic congestion can be significant during peak hours in Belgium, and *ex post* travel time calculations may be inaccurate without additional information about the relevant traffic conditions. We compared self-reported travel times with the calculated travel distances and, based on this, removed one respondent who reported an implausibly large travel time.
[b] One respondent indicated a travel time of zero minutes as he stayed in the adjacent campsite.
[c] Income statistics are reported in Table 4.

[15] We explore whether the inclusion of these observations affects our results in the robustness analysis.
[16] A Kruskal-Wallis test indicates that income differs significantly across three sub-samples stratified by employment status (employee, business owner or manager, unemployed) ($\chi^2 = 33.27$, $p < 0.0001$).

Table 4
Summary statistics of final sample (n = 293).

	Mean	Std deviation	Minimum	Maximum	Median
Variable					
Current travel cost (EUR per person per trip)	4.9	5.9	0.1	45.9	3.1
After-tax household income[a] (EUR per month) (n = 200)	2719	1346	500	6500	2500
Age (in years)	48.3	16.5	18	84	50
Member of environmental organization (1 = yes, 0 = no)	0.26	0.44	0	1	0
Gender (1 = male, 0 = female)	0.46	0.50	0	1	0
Higher education (1 = diploma of higher professional education or higher, 0 = otherwise)	0.56	0.50	0	1	1
RP trips (t = 1 – current situation)	29.9	78.9	1	182	8
SP trips (t = 2 – 50% grass cover in year+2)	32.5	39.5	0	156	12
SP trips (t = 3 – 100% grass cover in year+2)	25.6	42.0	0	182	12
SP trips (t = 4 – 50% burned area in year+1)	27.1	38.8	0	182	6
SP trips (t = 5 – 100% burned area in year+1)	19.4	32.5	0	182	10
SP trips (t = 6 – current situation, but with access fee in year+1)	28.5	41.6	0	156	3

[a] Based on the mid-points of the income brackets.

Table 5
Correlations between revealed and stated trips (r).

	RP trips (t = 1)	SP trips (t = 2)	SP trips (t = 3)	SP trips (t = 4)	SP trips (t = 5)	SP trips (t = 6)
RP trips (t = 1 – current situation)	1					
SP trips (t = 2 – 50% grass)	0.93	1				
SP trips (t = 3 – 100% grass)	0.84	0.94	1			
SP trips (t = 4 – 50% burned)	0.89	0.87	0.87	1		
SP trips (t = 5 – 100% burned)	0.79	0.97	0.80	0.85	1	
SP trips (t = 6 – current situation plus access fee)	0.92	0.91	0.84	0.87	0.81	1

$$p_i^x = \frac{\textit{Stated importance}}{\textit{of heathland}_i} \cdot \left(\frac{2 \cdot \textit{Travel distance}_i \cdot (\textit{Petrol cost per km} | \textit{Car size}_i)}{\textit{Group size}_i} + \frac{1}{3} \frac{\textit{Net monthly wage}_i}{165\ \textit{Hours/month}} 2 \cdot \textit{Travel time}_i \right) + \textit{Access fee}_i. \tag{19}$$

We used individual data for this calculation, except for fuel cost per kilometer, which we based on national data. The fuel cost per kilometer was calculated by multiplying the weighted average price per liter based on the share of diesel and gasoline cars in Belgium by the average fuel consumption of the car size category (small, medium, large) to which the car of the individual belongs. Travel distances were obtained by means of Google Maps. We account for multiple purpose day trips by weighting the on-the-road cost with the stated, relative importance of the heathland for the decision to make the current day trip (Martínez-Espiñeira and Amoako-Tuffour, 2008). This measure was obtained directly from the respondents. Respondents were asked to indicate the relative importance of the heathland on a seven-point scale, where one is totally unimportant and seven is most important. Subsequently, the relative importance of the heathland is obtained by dividing the response by seven. The fee to be paid to enter the heathland area is hypothetical since entrance is currently free of charge.

Summary statistics of the overall dataset are provided in Table 4. Furthermore, the correlations between RP and SP trips are presented in Table 5. The correlations, which are all 0.79 or higher, suggest that on-site sampling bias may affect both observed RP trips and observed SP trips (see Eq. (16)).

5. Results and discussion

5.1. Regression results

The recreation demand model specification includes travel cost and demand-shifting dummies for the hypothetical quality change treatments as the core variables, and several demographic and travel behavior variables that previous studies have found to influence recreation demand. Initially, these variables include income, gender, age, educational attainment, membership of an environmental organization, and whether the respondent visits as part of a family trip with children (Englin and Shonkwiler, 1995; Richardson and Loomis, 2004). The recreation demand model is estimated with and without constraining the intercept, travel cost, and income parameters across RP and SP equations. In the recreation demand model without constraints, the intercept, travel cost and income parameters in the RP equation are estimated with actual visitation

Table 6
Recreation demand models.[a]

	Corrected for on-site sampling bias		Not corrected for on-site sampling bias	
	Coeff.	Standard error	Coeff.	Standard error
Variable				
Travel Cost (RP)	−0.110	0.029***	−0.077	0.010***
Travel Cost (SP)	−0.070	0.011***	−0.067	0.007***
Gender	0.190	0.137	0.188	0.103*
Family visit	−0.157	0.695	0.581	0.202***
Missing income	0.303	0.166*	0.201	0.113*
50% Grass cover in year+2 (dummy)	0.011	0.161	−0.168	0.081**
100% Grass cover in year+2 (dummy)	−0.185	0.106*	−0.259	0.076***
50% Burned cover in year+1 (dummy)	−0.216	0.083***	−0.209	0.071***
100% Burned cover in year+1 (dummy)	−0.414	0.091***	−0.443	0.096***
Constant (RP)	−0.170	0.392	0.731	0.159***
Constant (SP)	0.776	0.169***	0.933	0.142***
a_{rp}	0.383	0.457	0.016	0.024
a_{sp}	0.181	0.142	0.023	0.034
p	2.196	0.288***	2.796	0.399***
γ_{rp}	4.409	0.527***	3.799	0.187***
γ_{sp}	3.858	0.197***	3.605	0.155***
λ_2[b]	0.544	0.120***	0.534	0.114***
λ_3[b]	−1.046	0.235***	−1.113	0.146***
θ_1[b]	0.419	0.292	0.387	0.150***
θ_2[b]	0.360	0.379	0.000	0.204
θ_3[b]	0.271	0.370	−0.173	0.327
Log Likelihood	−4076.0		−4103.1	
AIC	8193.9		8248.2	
BIC	8300.3		8354.6	

[a] Robust standard errors clustered by individual are reported. Significance levels are indicated by * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.
[b] The estimated parameter values of λ and θ before transformation are presented.

data only, and the same parameters in the SP equation are estimated with contingent behavior data only. Furthermore, these models are estimated with and without on-site sampling bias correction, resulting in a total of four initial models. From these initial models we obtained a set of final models, which we present in this section.[17] We selected the final models based on three grounds. First, we tested the parameter restrictions across the RP and SP equations. We found that the unconstrained models outperform the constrained models ($\beta_{rp} = \beta_{sp}$) based on an LRT[18] at a significance level of 1%. Given the rejection of constrained parameters, we selected the unconstrained models for further consideration. Second, following von Haefen and Domanski (2018), we selected the optimal number of discrete classes based on the corrected Akaike Information Criterion (cAIC). This criterion imposes a higher penalty on a larger number of parameters than the uncorrected AIC and Bayesian Information Criterion (BIC). The cAIC supports four discrete classes over three discrete classes for all initial models.[19] Third, the covariates income, membership of environmental organization, age and higher education are not significant in any of the initial models. Hence, we dropped these covariates and re-estimated the unconstrained models with four discrete classes. The conditional mean of these final models is specified as:

$$\mu_{itk} = \begin{cases} exp\left(\beta_{0,RP} + \beta_{p,SP} p_i^x + \beta_D D_i + \gamma_{RP}\lambda_k\right), t = 1 \\ exp\left(\beta_{0,SP} + \beta_{p,SP} p_i^x + \beta_D D_i \ + \beta_{Q,SP} Q_t + \gamma_{SP}\lambda_k\right), t = 2,3,4,5,6 \end{cases}, \quad \text{for } \ k = 1,2,3,4 \ , \tag{20}$$

[17] The initial model specifications and regression results are included in Online Appendix 2.

[18] For the models corrected for on-site sampling bias, $\chi^2_{df=3} = 104.2$ ($p < 0.0001$). For the models uncorrected for on-site sampling bias, $\chi^2_{df=3} = 14.4$ ($p = 0.006$). We also explored parameter consistency by pooling the RP and SP data in a single demand equation in which we interacted an RP dummy variable with the constant, the travel cost and income parameters, and the factor loading (see Online Appendix 3). Even though all interaction terms (apart from *Constant*RP* in the corrected model) are insignificant, LRTs favor the models in which these interaction terms are not constrained to zero ($\chi^2_{df=4} = 138.2$ ($p < 0.0001$) for the corrected model, and $\chi^2_{df=4} = 11.8$ ($p = 0.018$) for the uncorrected model). Since we found that the DFM approach is more appropriate based on an AIC-adjusted Vuong non-nested test ($p = 0.0001$ and $p = 0.018$ for the corrected and uncorrected models, respectively), we focus on the DFM model in the remainder of this paper.

[19] The cAIC scores of the initial models with three classes are included in Online Appendix 2. We also attempted to estimate the initial models with five classes. These models consistently converged to more negative log likelihood values, with several parameters converging to values that are theoretically inconsistent with utility maximizing behavior.

where p_i^x is travel cost, D_i is a vector of the demographic variables: gender, family visit, missing income, and Q_t is a vector of dummy variables for: 50% grass cover, 100% grass cover, 50% burned cover and 100% burned cover. The regression results of the final models are presented in Table 6. The regression results shown in this paper are obtained by selecting the set of parameter estimates with the least negative log-likelihood value after running the BADS algorithm from 20 different starting value sets. Following the same procedure for the EM algorithm, we found that both algorithms produced the same convergence results in terms of log likelihood value and parameter values. We also found that the BADS algorithm did not converge to local optima, whereas the EM algorithm sometimes converged to local optima when the initial parameter values were far off the final parameter values. Finally, in cases where the initial parameter values were set such that both algorithms converged to the maximum likelihood values in Table 6, the BADS algorithm converged faster than the EM algorithm.[20]

Both travel cost parameters are negative and significant. Being a male has a significant positive effect on recreation demand, but this effect is only significant in the uncorrected model. Recreation demand is significantly higher when a visitor comes as part of a family in the uncorrected model. This effect is negative and not significant in the corrected model. Furthermore, respondents who did not report their income bracket have a significantly higher visitation rate than other respondents, in both models.

The estimates for the treatment parameters are negative and statistically significant in both models, except for the 50 percent grass cover treatment, which is also negative, but significant in the uncorrected model only. Based on the size of the treatment parameters, we expect the biggest decrease in visitation rates directly after a 100 percent wildfire. Visitation increases but does not fully recover when the burned area becomes encroached by grasses in the years after a 100 percent wildfire. After a 50 percent wildfire we also expect a decrease in visitation, but as soon as the burned area is covered with grasses, we expect no significant difference with pre-burn visitation rates.[21]

The dispersion parameter a is not statistically significant, whereas dispersion parameter p is statistically significant in both models. The unobserved heterogeneity parameters λ_2 and λ_3 are statistically significant in both models. However, the parameters θ, which enter the multinomial logit used to calculate class probabilities, are all insignificant except θ_1 in the uncorrected model.[22] This implies that individuals are equally likely to belong to each of the classes with insignificant θ parameters.

A comparison of the RP and SP parameters for the travel cost parameter, intercept, and factor loadings reveals that these parameters diverge to a larger extent in the on-site sampling bias corrected model compared to the uncorrected model. Specifically, when moving from the uncorrected model to the corrected model, the RP factor loading becomes larger, whereas the travel cost parameter and intercept become smaller. This implies that, in the corrected model, some of the RP variation that can be explained in the uncorrected model is shifted to unobserved heterogeneity. The SP parameters do not show this shift. A possible explanation is that RP parameter estimates are directly affected by the on-site sampling bias correction because a lower expected RP demand directly improves the sample likelihood function (see Eq. (18)).

Based on the AIC and BIC, we find support for using the model corrected for on-site sampling bias. A Vuong non-nested likelihood ratio test also lends support to using the corrected model ($p = 0.07$).

The travel cost and hypothetical quality change parameter estimates may be sensitive to the selection of covariates. We use the definition of robustness introduced by Neumayer and Plümper (2017) to assess the extent to which these two parameter estimates are sensitive to covariate selection. Specifically, we quantify the percentage of the probability density distribution of the estimate from the initial model that lies within the 95% confidence bounds of the estimate from the final model. The results of this test indicate that the effects of the travel cost and quality changes on recreation demand do not depend much on covariate selection, because the robustness of the parameter estimates varies between 0.90 and 1.00. Furthermore, the statistical significance of these parameters is not sensitive to the model specification.[23]

We also estimated the final models without removing highly frequent visitors (see Online Appendix 6). The parameter estimates for the quality change dummy variables remain similar in magnitude and statistical significance, except that the "50% burned" dummy variable becomes insignificant in the uncorrected model. Furthermore, the travel cost parameter estimate is smaller in both the corrected and uncorrected models and remains significant in the corrected model only. These results suggest that, on average, predicted recreation demand is less sensitive to travel cost if highly frequent visitors are

[20] For example, when estimating the final on-site corrected model with zero starting values for all parameters, the BADS algorithm converged to the log likelihood value in Table 4 in approximately 6 min, while the EM algorithm stopped after reaching a local optimum of the objective function (−4312.44). When the starting values of all parameters were set to an arbitrary value between −0.05 and +0.05 relative to the maximum likelihood parameter values, we found that the estimation time of the BADS and EM algorithms was 4 and 11 min, respectively.

[21] We find that the "First scenario" dummy variable is an insignificant predictor of SP trips (Online Appendix 4), suggesting that there are no ordering effects.

[22] The regression results for the latent class models with class membership covariates are included in Online Appendix 5. Using an AIC-adjusted Vuong non-nested test statistic (Vuong, 1989) (that is, adjusted for the number of parameters), we find that our DFM approach with three classes is more appropriate than the latent class model with three classes (on-site corrected model: $p = 0.003$; uncorrected model: $p < 0.0001$). Furthermore, our DFM approach with four classes is more appropriate than the latent classes model with four classes (on-site corrected model: $p < 0.0001$; uncorrected model: $p = 0.038$). We have not included the latent class model with four classes in the appendix because we are not completely confident that all 27 parameters were recovered without identification issues.

[23] The Vuong statistic also supports the on-site correction based on the models with all initial covariates included ($p = 0.05$).

Table 7
Predicted trips and consumer surplus.[a]

	Corrected for on-site sampling bias		Not corrected for on-site sampling bias	
	Mean	95% CI	Mean	95% CI
Predicted trips per person per year				
Current situation	28.90	(18.02–39.77)	31.80	(24.92–38.67)
50% grass cover in year+2	29.22	(18.06–40.37)	26.89	(20.57–33.20)
100% grass cover in year+2	24.02	(12.80–35.23)	24.55	(18.65–30.44)
50% burned cover in year+1	23.28	(14.77–31.78)	25.80	(19.95–31.64)
100% burned cover in year+1	19.09	(11.32–26.85)	20.41	(15.21–25.60)
Current situation, but with access fee in year+1	26.30	(16.18–36.41)	29.07	(22.75–35.38)
Predicted consumer surplus per trip	14.28	(12.24–16.31)	14.96	(13.00–16.92)
Predicted consumer surplus per person per year				
Current situation	412.59	(208.37–616.80)	475.85	(350.76–600.93)
50% grass cover in year+2	417.15	(224.07–610.22)	402.35	(287.55–517.14)
100% grass cover in year+2	342.95	(142.26–543.63)	367.31	(258.78–475.83)
50% burned cover in year+1	332.44	(175.60–489.27)	386.09	(277.32–494.85)
100% burned cover in year+1	272.62	(136.06–409.17)	305.42	(215.63–395.20)
Current situation, but with access fee in year+1	375.55	(184.97–566.12)	434.95	(316.93–552.96)

[a] Standard errors are obtained by means of the delta method.

included in the analysis. Again, based on a Vuong test, we find support for the on-site correction using these models ($p = 0.09$).

5.2. Welfare change estimation

The expected consumer surplus of individual i under treatment t belonging to class k is calculated by (Hellerstein and Mendelsohn, 1993):

$$E(CS_{itk}) = \int_{p_0^x}^{p_c^x} \int_E f(\varepsilon)\mu_{itk} d\varepsilon dp, \tag{21}$$

where ε represents unobserved heterogeneity, of which the probability density function is f over a range E. We integrate over the probability density function by computing the probability weighted sum over K discrete classes. Taking into account the semi-log functional form of recreation demand (Bockstael et al., 1989), the expected consumer surplus can be calculated as:

$$E(CS_{it}) = \sum_{k=1}^{K} \Pr(\lambda_k) \frac{\mu_{itk}}{-\beta_p^t}, \tag{22}$$

where β_p^t is the travel cost coefficient of treatment t. The inconsistency between RP and SP parameters makes the use of both demand equations somewhat problematic for welfare estimation, because the difference between recreation demand and consumer surplus in the current situation and hypothetical treatments is not only caused by quality or price changes, but also by the difference between the RP and SP parameter estimates for the travel cost, intercept and factor loading. Whereas the difference in factor loadings indicates inconsistency between the degree of preference dispersion of the RP and SP data sets, the difference of the intercept and travel cost terms implies a difference between actual and intended behavior under the same quality and price. This means that quality-induced changes in recreation demand are confounded by the change caused by inconsistency between actual and intended behavior if both the RP and SP equations are used for prediction. This implies that we have to use either the RP or SP equation to examine the effect of quality changes on recreation demand only.[24] Since only the SP equation was informed with recreation demand after quality changes, we use the SP equation to forecast future recreation demand. In particular, we use the SP equation to calculate trips and consumer surplus under baseline conditions by setting the treatment variables to zero, and for each hypothetical treatment by setting the associated treatment variable to one and the others to zero. We obtain predicted mean trips and consumer surplus under relevant demand conditions by computing the mean for each individual in the sample and subsequently computing the sample average. The predicted trips and consumer surplus are provided in Table 7.

[24] Another commonly adopted approach is to estimate an SP dummy covariate and set it to zero in welfare estimation. However, this requires including an extra treatment in which anticipated recreation demand under the status quo is elicited.

The corrected and uncorrected model predict similar impacts of the hypothetical scenarios on trips and consumer surplus. Expected trips show the largest decrease directly after a wildfire in which 100 percent of the heathland area was burned in both models. Expected trips are also lower in the subsequent years compared to the pre-burn situation but are higher than the expected trips in the year immediately following a 100 percent wildfire. The same pattern is observed for consumer surplus. Furthermore, expected trips and consumer surplus are not significantly affected by a hypothetical access fee.

5.3. Discussion

The consumer surplus calculated with the SP travel cost coefficient (EUR 14.28 (95% CI: 12.24–16.32)) is significantly larger than its RP counterpart (EUR 9.08 (95% CI: 6.93–11.23)). However, it is common for actual and hypothetical stated values to differ by a factor of three (List and Gallet, 2001). In general, our consumer surplus estimates are higher than comparable estimates in the literature. For example, a recent meta-analysis of recreational valuation studies concerning European National Parks found a mean consumer surplus of EUR 7.55 per trip (Schägner et al., 2016). A potential explanation for this difference might lie in the fact that the study area is the only National Park in Belgium, potentially attracting more visitors from relatively far away due to the lack of substitutes. A second potential explanation for this difference might lie in the relatively high mean household income of the sample (EUR 2719 per household per month), which was 64 percent higher than the mean household income across European Union Member States in the study year (2018) (Eurostat, 2019).

As noted, we found that the dummy variables for all hypothetical quality change treatments, except "50 percent grass cover," have significant coefficients in the regression models, but the predicted decrease of expected consumer surplus is significant only for the "100 percent burned" scenario. This finding could be the result of DFM–GNB producing relatively large standard errors in the final welfare estimates relative to parametric estimators (Landry and Liu, 2009). Another explanation is that visitors are indifferent about whether there is grass or heath vegetation during recovery years. Indeed, some respondents who did not report a change of visitation rates after a hypothetical wildfire indicated that the vegetation is not important for their visitation rates, because they come to enjoy nature in general. Another possible explanation is that visitors anticipate that the heathland area will recover in the future, and do not change their visitation behavior during the process of restoration. With regard to the hypothetical increase of travel cost, we also find that an access fee of EUR 3 per visit or an annual subscription pass of EUR 5 appears not to significantly decrease visitation rates to the study area. This implies that the park management may consider asking for a contribution from visitors without significantly decreasing current visitation rates.

Our finding that recreation demand decreases after an extreme wildfire is consistent with decreases found in early studies on forest recreation in the United States (Flowers et al., 1985; Hesseln et al., 2003). Some recent studies have reported an increase of annual visitation rates immediately or several years after forest fires (Hilger and Englin, 2009; Sánchez et al., 2016), which can be (partially) attributed to the attractiveness of observing the sudden change in the ecosystem induced by the fire or of observing a slowly recovering ecosystem. Our study showed no corresponding increases in recreation demand immediately or several years after the wildfire, except that the on-site corrected model predicts a small but insignificant increase of recreation demand immediately after an intermediate wildfire. This finding suggests that the negative impact caused by the loss of heath vegetation weights heavier than the positive experience of observing wildfire effects. We also observed that visitation rates partly recover during the years following an initial wildfire year. This contrasts with the findings of Englin et al. (2001) who found that recreation demand declines during post-wildfire recovery years compared to immediately after the wildfire.

Comparing the corrected model with the uncorrected model, the predicted trips in the former are, on average, 5 percent lower. For example, without correcting for on-site sampling bias, the predicted number of trips per person per year in the current situation is 31.8 (95% confidence interval (CI): 24.9–38.7). With correction of on-site sampling bias, the number of predicted trips decreases to 28.9 (95% CI: 18.0–39.8). The opposite pattern is observed for consumer surplus, because the uncorrected model has a smaller travel cost parameter than the corrected model, which reduces the consumer surplus per trip. The estimated consumer surplus in the current situation is EUR 476 (95% CI: 351–601) per person per year without on-site sampling bias correction, and with on-site sampling bias correction this amount decreases to EUR 413[25] (95% CI: 208–617). On average, the correction for on-site sampling bias via the cross-equation correlation imposed by the discrete factor method decreases RP and SP welfare estimates by 13 percent and 8 percent, respectively. The impact of the correction is at the lower end of the range of impacts found in previous studies in which recreation demand models are corrected for on-site sampling bias. For example, regarding the number of expected trips under current conditions, we found a 5 percent decrease due to the correction, whereas Beaumais and Appéré (2010), Egan and Herriges (2006), and Moeltner and Shonkwiler (2005) reported decreases of 3 percent, 71 percent, and 94 percent, respectively.[26] The relatively small impact of on-site correction on welfare estimates may be attributed to the fact that RP-SP parameter inconsistency and a lack of RP observations regarding wildfire impacts on recreation necessitate the use of SP coefficients for the computation of welfare estimates. However, the SP coefficients are affected by the on-site correction less than the RP coefficients.

[25] This is equal to 1.3% of mean annual household income.

[26] Other studies that estimated corrected recreation demand models with empirical data (e.g. Hynes and Greene, 2013) did not report the impact of the correction on welfare estimation.

We note the following limitations of this study: First, the hypothetical treatments presented to the respondents imply that a wildfire occurs with 100 percent probability. Within a risk analysis framework, the welfare loss after a wildfire can then be multiplied with the expected increase in probability of a wildfire to obtain the expected value of increased wildfire risk in the future. It is possible that different welfare estimates would be obtained if visitors were asked directly about their *ex-ante* willingness to avoid an increase of wildfire risk. Hence, a possible avenue for future research is to investigate whether such an alternative approach would lead to different welfare estimates. Second, we did not take into account the travel cost to substitute sites. Ignoring the travel cost to substitute sites may have introduced some omitted variable bias. However, the assumption that there are no close substitutes for our study area seems reasonable because it is located in the only national park in Belgium and the availability of other accessible heathland areas of similar size is low. Third, the rejection of consistency between revealed and stated preference parameters indicates that the data sources should not be combined. The difference of the RP and SP parameter estimates implies that the provided welfare estimates are upward biased under the assumption that revealed preference data are more valid than stated preference data. However, the rejection of parameter consistency does not tell us which data source is more valid, or whether neither is valid. Further research should address the sources of bias in the revealed and stated preference data, particularly when the effect of climate hazards on visitation behavior is concerned. Fourth, the estimates may be less efficient than estimates obtained with competing parametric models, such as the Multivariate Poisson Log-Normal model. However, the latter model may not be able to deal with the relatively large number of treatments. This implies that a tradeoff exists between the number of treatments that researchers want to include and the need for efficiency in the estimates. This tradeoff is particularly relevant for valuing the impact of climate-induced events on recreation, because the magnitude of climate-induced events is unknown and different magnitudes can have contrasting effects on recreation. In our case, without including treatments with wildfires of different magnitude, the non-linearity property of the effect of wildfires on recreational value would have gone unnoticed. This means that DFM−GNB may be able to provide useful information in studies in which the quality change is uncertain, and the aim is to investigate a range of outcomes. On the other hand, parametric models may be preferred if the magnitude of the quality change is less uncertain. Finally, we used Marshallian consumer surplus to compute welfare changes, in the absence of sufficient data to reliably estimate the income coefficients. In many cases, Marshallian welfare is a reasonable approximation of Hicksian welfare, although the approximation error depends on the income elasticity and the amount of income spent on the good (Hausman, 1981).

6. Conclusions and recommendations

In this study we combined a finite mixture model − the Discrete Factor Method−Generalized Negative Binomial (DFM−GNB) model − with on-site survey data after extending the model to control for on-site sampling bias. The survey data consisted of revealed and stated preference data collected among heathland visitors in the Hoge Kempen National Park (Belgium).

Consumer surplus estimates obtained with the on-site sampling bias corrected and uncorrected models suggest that if a heathland area is affected by an intermediate or extreme wildfire, recreation demand will significantly decrease. Consumer surplus significantly decreases in the case of an extreme wildfire only. In the years after an intermediate or extreme wildfire, there is no significant effect on recreation demand or consumer surplus.

We find that our model specification provides robust estimates for the travel cost and hypothetical quality change parameters. We reject consistency between revealed and stated preference parameters based on a likelihood ratio test. However, based on a Vuong test, we conclude that the models corrected for on-site sampling have a better model fit relative to the uncorrected models. This implies that our analysis still benefits from joint RP−SP estimation, because without the estimation of an RP equation we are not able to properly correct for on-site sampling bias correction. However, we do note that this correction does not have a significant impact on welfare measures in our specific application.

The welfare estimates obtained in this study can be compared with the costs of wildfire prevention measures, which include a combination of mowing, grazing, controlled burning, and maintaining compartments that prevent the spread of wildfires. Verburg et al. (2017) recently estimated the costs of these measures for a number of European, terrestrial ecosystems. Based on this, the conservation costs of our study area are estimated to amount to EUR 240,000 per year. Depending on whether revealed or stated preference data are used, we estimate that the recreational value exceeds these costs by a factor 6 or 10 (EUR 1,590,750 per year or 2,500,750 per year, respectively, based on 175,000 trips to the *Mechelse Heide* per year (Pollaris, 2018)). Our results also suggest that recreational value alone may not fully justify investment in restoration management after wildfires because of the relatively small difference between visitation rates in a grass-encroached area and the current situation. However, heathland areas also provide a variety of other benefits that are not accounted for in our study and that would need to be included in a comprehensive cost-benefit assessment of wildfire prevention measures.

Financial disclosure statement

We certify that no party having a direct interest in the results of the research supporting this article has or will confer a benefit on us or on any organization with which we are associated and we certify that all financial and material support for this research and work are clearly identified in the title page of the manuscript.

Declaration of competing interest

We certify that there are no conflicts of interest to disclose.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.jeem.2020.102317.

References

Acerbi, L., Ma, W.J., 2017. Practical bayesian optimization for model fitting with bayesian adaptive direct search. Adv. Neural Inf. Process. Syst. 30, 1834–1844.
Adamowicz, W., Louviere, J., Williams, M., 1994. Combining revealed and stated preference methods for valuing environmental amenities. J. Environ. Econ. Manag. 26 (3), 271–292.
Agentschap Natuur en Bos, 2016. *Tevredenheidsonderzoek Mechelse Heide*. Technical Report. Agentschap Natuur en Bos, Brussels.
Akaike, H., 1973. Information theory and an extension of the maximum likelihood principle. In *selected Papers of hirotugu Akaike*. In: Kotz, S., Johnson, N.L. (Eds.), (1992), Breakthroughs in Statistics, I. Springer-Verlag, pp. 610–624.
Allman, E.S., Ba, H., Evans, R., Ho, S., 2019. Maximum likelihood estimation of the latent class model through model boundary decomposition. J. Algebraic Stat. 10 (1), 51–84.
Audet, C., Dennis, J.E.J., 2006. Mesh adaptive direct search algorithms for constrained optimization. SIAM J. Optim. 17 (1), 188–217.
Bateman, I.J., Carson, R.T., Day, B., Hanemann, M., Hanley, N., Hett, T., et al., 2002. Economic Valuation with Stated Preference Techniques. Edward Elgar Publishing, Cheltenham, UK.
Beaumais, O., Appéré, G., 2010. Recreational shellfish harvesting and health risks: a pseudo-panel approach combining revealed and stated preference data with correction for on-site sampling. Ecol. Econ. 69 (12), 2315–2322.
Ben-Akiva, M., Bradley, M., Morikawa, T., Benjamin, J., Novak, T., Oppewal, H., Rao, V., 1994. Combining revealed and stated preferences data. Market. Lett. 5 (4), 335–349.
Bockstael, N.E., McConnell, K.E., 1993. Public goods as characteristics of non-market commodities. Econ. J. 103, 111–132.
Bockstael, N.E., Hanemann, W.M., Strand, I.E., 1989. Measuring the Benefits of Water Quality Improvements Using Recreation Demand Models. United States Environmental Protection Agency, Washington. EPA-230-10-89-069.
Cherchi, E., Ortúzar, J. de D., 2011. On the use of mixed RP/SP models in prediction: accounting for systematic and random taste heterogeneity. Transport. Sci. 45 (1), 98–108.
Cordingley, J.E., Newton, A.C., Rose, R.J., Clarke, R.T., Bullock, J.M., 2015. Habitat fragmentation intensifies trade-offs between biodiversity and ecosystem services in a heathland ecosystem in southern England. PloS One 10 (6), 1–16.
Depicker, A., Baets, B. De, Baetens, J.M., 2020. Wildfire ignition probability in Belgium. Nat. Hazards Earth Syst. Sci. 20 (2), 1–32.
Egan, K., Herriges, J., 2006. Multivariate count data regression models with individual panel data from an on-site sample. J. Environ. Econ. Manag. 52 (2), 567–581.
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. Rev. Econ. Stat. 77 (1), 104–112.
Englin, J., Loomis, J., Gonzalez-Caban, A., 2001. The dynamic path of recreational values following a forest fire. Can. J. For. Res. 10, 18–37.
European Commission, 1992. Council Directive 92/43/EEC of 21 May 1992 on the Conservation of Natural Habitats and of Wild Fauna and Flora. Retrieved from ec europa.eu/environment/nature/legislation/habitatsdirective/index_en.htm.
Eurostat, 2019. Mean and Median Income by Household Type - EU-SILC and ECHP Surveys. Retrieved from. https://ec.europa.eu/eurostat/web/products-datasets/product?code=ilc_di04.
Fagúndez, J., 2013. Heathlands confronting global change: drivers of biodiversity loss from past to future scenarios. Ann. Bot. 111 (2), 151–172.
Flannigan, M.D., Wotton, B.M., Marshall, G.A., de Groot, W.J., Johnston, J., Jurko, N., Cantin, A.S., 2016. Fuel moisture sensitivity to temperature and precipitation: climate change implications. Climatic Change 134 (1–2), 59–71.
Flowers, P.J., Vaux, H., Gardner, P.D., Mills, T.J., 1985. Changes in Recreation Values after Fire in the Northern Rocky Mountains. United States Department of Agriculture, Berkeley.
Gurmu, S., Elder, J., 2012. Flexible bivariate count data regression. J. Bus. Econ. Stat. 30 (2), 265–274.
Hausman, J.A., 1981. Exact consumer's surplus and deadweight loss. Am. Econ. Rev. 71 (4), 662–676.
Heckman, J., Singer, B., 1984. A method for minimizing the impact of distributional assumptions in econometric models for duration data. Econometrica 52 (2), 271.
Hellerstein, D., Mendelsohn, R., 1993. A theoretical foundation for count data models. Am. J. Agric. Econ. 75, 604–611.
Herriges, J.A., Phaneuf, D.J., Tobias, J.L., 2008. Estimating demand systems when outcomes are correlated counts. J. Econom. 147 (2), 282–298.
Hesseln, H., Loomis, J.B., González-Cabán, A., Alexander, S., 2003. Wildfire effects on hiking and biking demand in New Mexico: a travel cost study. J. Environ. Manag. 69 (4), 359–368.
Hilger, J., Englin, J., 2009. Utility theoretic semi-logarithmic incomplete demand systems in a natural experiment : forest fire impacts on recreational values and use, 31, 287–298.
Hobbs, R.J., Gimingham, C.H., 1984. Studies on fire in scottish heathland communities II . Post-fire vegetation development. J. Ecol. 72 (2), 585–610.
Huang, P., 2017. Asymptotics of AIC, BIC and RMSEA for model selection in structural equation modeling. Psychometrika 82 (2), 407–426.
Hynes, S., Greene, W., 2013. A panel travel cost model accounting for endogenous stratification and truncation: a latent class Approach. Land Econ. 89 (1), 177–192.
Jollois, F.X., Nadif, M., 2007. Speed-up for the expectation-maximization algorithm for clustering categorical data. J. Global Optim. 37, 513–525.
Landry, C.E., Liu, H., 2009. A semi-parametric estimator for revealed and stated preference data-An application to recreational beach visitation. J. Environ. Econ. Manag. 57 (2), 205–218.
Lewis, F., Butler, A., Gilbert, L., 2011. A unified approach to model selection using the likelihood ratio test. Methods Ecol. Evol. 2, 155–162.
List, J.A., Gallet, C.A., 2001. What experimental protocol influence disparities between actual and hypothetical stated values? Environ. Resour. Econ. 20 (3), 241–254.
Loomis, J., Gonzalez-Cabán, A., Champ, J., Downing, J., 2011. The International Handbook on Non-market Environmental Valuation. Edward Elgar Publishing, Cheltenham, UK.
Martínez-Espiñeira, R., Amoako-Tuffour, J., 2008. Recreation demand analysis under truncation, overdispersion, and endogenous stratification: an application to Gros Morne National Park. J. Environ. Manag. 88 (4), 1320–1332.
Meyerhoff, J., Liebe, U., 2008. Do protest responses to a contingent valuation question and a choice experiment differ? Environ. Resour. Econ. 39 (4), 433–446.
Moeltner, K., Shonkwiler, J.S., 2005. Correcting for on-site sampling in random utility models. Am. J. Agric. Econ. 87 (2), 327–339.

Molina, J.R., González-Cabán, Silva, F.R., 2019. Wildfires impact on the economic susceptibility of recreation activities: application in a Mediterranean protected area. J. Environ. Manag. 245, 454–463.
Neumayer, E., Plümper, T., 2017. Robustness Tests for Quantitative Research. Methodological Tools in the Social Sciences. Cambridge University Press, Cambridge.
Patil, G.P., Rao, C.R., 1978. Weighted distributions and size-biased sampling with applications to wildlife populations and human families. Biometrics 34 (2), 179–189.
Pollaris, S., 2018. De waarde van het Nationaal Park Hoge Kempen voor bezoekers, specifiek voor de Mechelse Heide. Unpublished Master's thesis. Hasselt University, Hasselt, Belgium.
Pudney, S., 1991. Modelling Individual Choice: the Econometrics of Corners, Kinks and Holes. Blackwell Publishing, New York.
Richardson, R.B., Loomis, J.B., 2004. Adaptive recreation planning and climate change: a contingent visitation approach. Ecol. Econ. 50 (1–2), 83–99.
Sánchez, J.J., Baerenklau, K., González-cabán, A., 2016. Forest Policy and Economics Valuing hypothetical wild fi re impacts with a Kuhn – tucker model of recreation demand. For. Pol. Econ. 71, 63–70.
Schägner, J.P., Brander, L., Maes, J., Paracchini, M.L., Hartje, V., 2016. Mapping recreational visits and values of European National Parks by combining statistical modelling and unit value transfer. J. Nat. Conserv. 31, 71–84.
Schepers, L., Haest, B., Veraverbeke, S., Spanhove, T., Borre, J. Vanden, Goossens, R., 2014. Burned area detection and burn severity assessment of a heathland fire in Belgium using airborne imaging spectroscopy (APEX). Rem. Sens. 6 (3), 1803–1826.
Shaw, D., 1988. On-site samples' regression: problems of non-negative integers, truncation, and endogenous stratification. J. Econom. 37 (2), 211–223.
Thiene, M., Scarpa, R., Tempesta, T., 2007. Latent class count models of total visitation demand: days out hiking in the eastern Alps. Environ. Resour. Econ. 38 (4), 447–460.
Van Marwijk, R.B.M., Elands, B.H.M., Kampen, J.K., Terlouw, S., Pitt, D.G., Opdam, P., 2012. Public perceptions of the attractiveness of restored nature. Restor. Ecol. 20 (6), 773–780.
Vaux, H.J., Gardner, P.D., Mills, T.J., 1984. *Methods For Assessing the Impact of Fire on Forest Hecreation*. PSW-079. U.S. Department of Agriculture, Berkeley.
Verburg, R.W., Puister, L.F., Michels, R., Duijvendijk, K. Van, 2017. Estimating Costs of Nature Management in the European Union. Wageningen University and Research, Wageningen. Research Report.
von Haefen, R.H., 2002. A complete characterization of the incomplete demand system models. J. Agric. Resour. Econ. 27 (2), 281–319.
von Haefen, R.H., Domanski, A., 2018. Estimation and welfare analysis from mixed logit models with large choice sets. J. Environ. Econ. Manag. 90, 101–118.
Vuong, Q.H., 1989. Likelihood ratio tests for model selection and non-nested hypotheses. Econometrica 57 (2), 307–333.
Whitehead, J.C., Pattanayak, S.K., Van Houtven, G.L., Gelso, B.R., 2008. Combining revealed and stated preference data to estimate the nonmarket value of ecological services: an assessment of the state of the science. J. Econ. Surv. 22 (5), 872–908.
Whitehead, J., Haab, T., Huang, J.C., 2011. Preference Data for Environmental Valuation: Combining Revealed and Stated Approaches. Routledge, New York.
Willig, R.D., 1978. Incremental consumer's surplus and hedonic price adjustment. J. Econ. Theor. 17 (2), 227–253.