Journal of Environmental Management 211 (2018) 238–246

Contents lists available at ScienceDirect

Journal of Environmental Management

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

Research article

Zero-truncated panel Poisson mixture models: Estimating the impact on tourism benefits in Fukushima Prefecture

Masaki Narukawa [a], Katsuhito Nohara [b, *]

[a] Faculty of Economics, Okayama University, 3-1-1 Tsushima-naka, Kita-ku, Okayama, 700-8530, Japan
[b] School of Economics, Hokusei Gakuen University, 2-3-1 Ohyachi-Nishi, Atsubetsu-ku, Sapporo, 004-8631, Japan

A R T I C L E I N F O

Article history:
Received 16 September 2017
Received in revised form
29 December 2017
Accepted 30 December 2017
Available online 4 February 2018

Keywords:
Contingent behavior
Count data model
Inverse Gaussian distribution
Random-effects model
Travel cost method
Zero truncation

A B S T R A C T

This study proposes an estimation approach to panel count data, truncated at zero, in order to apply a contingent behavior travel cost method to revealed and stated preference data collected via a web-based survey. We develop zero-truncated panel Poisson mixture models by focusing on respondents who visited a site. In addition, we introduce an inverse Gaussian distribution to unobserved individual heterogeneity as an alternative to a popular gamma distribution, making it possible to capture effectively the long tail typically observed in trip data. We apply the proposed method to estimate the impact on tourism benefits in Fukushima Prefecture as a result of the Fukushima Nuclear Power Plant No. 1 accident.

1. Introduction

The travel cost method (TCM) is often used to estimate the recreational or tourism value of a site. There is an inverse relationship between the cost of visiting the recreation site and the number of visits observed. This relationship provides the downward sloping demand curve. However, to estimate the benefits of the site quality, it is necessary to determine the demand for trips to the site for a level of site quality that is not currently observed (Alberini et al., 2007). In recent studies, researchers ask individuals how many trips they would take to a site, under hypothetical conditions. Then, they combine the responses (known as contingent behavior (CB)) with observations on actual trips to the site under the current conditions (e.g., Englin and Cameron, 1996; Whitehead et al., 2000; Grijalva et al., 2002; Hanley et al., 2003; Egan and Herriges, 2006; Christie et al., 2007; Whitehead et al., 2008; Beaumais and Appéré, 2010; Hynes and Greene, 2013, 2016; Voltaire et al., 2017). In addition, when researchers use a conventional TCM approach in recreational demand modeling, it is difficult to estimate changes to the consumer surplus (CS) because the site quality (e.g., degree of air pollution or water pollution) applies uniformly to all individuals who visit there. Even if an environment index is replaced with an indirect effect on environmental quality, such as the catch rate, it is still difficult to separate the differences in individuals' skills. However, combining revealed preference (RP) and stated preference (SP) data with questionnaire responses on trip numbers under a hypothetical level of environmental quality makes it possible to measure the benefits of different levels of environmental quality (Whitehead et al., 2000).

As noted in Landry and Liu (2009), pseudo panels of site-frequency data can be created by collecting information on the number of trips in the current period (RP data) and the number of trips that would be taken under different or similar conditions in future periods (SP data). In other words, when using the contingent behavior TCM, a panel or multivariate count data model is suitable for estimating the demand function. This is because respondents are asked about actual and CB trips. In TCM analyses, a popular way to collect data is an on-site survey, even though there are truncation and endogenous stratification issues (Egan and Herriges, 2006; Martínez-Espiñeira and Amoako-Tuffour, 2008). More specifically, non-users are excluded and those who visit the site more

* Corresponding author.
E-mail address: nohara@hokusei.ac.jp (K. Nohara).

https://doi.org/10.1016/j.jenvman.2017.12.082
0301-4797/

frequently are oversampled, which leads to a biased result (Shaw, 1988; Englin and Shonkwiler, 1995). The effects of these issues in the TCM analysis are examined by Blaine et al. (2015). To deal with the combination of RP and SP data in the contingent behavior TCM, by correcting for the on-site sampling issues, Egan and Herriges (2006) propose a multivariate Poisson log-normal model, and Beaumais and Appéré (2010) extend the random-effects Poisson gamma model of Hausman et al. (1984). Furthermore, as a more general approach, Hynes and Greene (2013, 2016) consider latent class and random parameter approaches, based on a negative binomial (NB) model in the framework of pseudo-panel data collected on-site. However, if an off-site survey is employed to collect data, these approaches cannot avoid providing a problematic analysis.

In off-site surveys, it is expected that many non-users (zero data) will be observed, possibly leading to the issue of excess zeros. In this case, higher costs are incurred in gathering data on enough of the limited sample of site users, and the excess zeros may induce misleading inferences in the TCM analysis (e.g., Hellerstein, 1992). More recently, development of the web environment has seen it become a central means of information transmission in society, and it is now widely utilized for survey data collection. In terms of methodology, web-based surveys can be regarded as being off-site. It is well known that web-based surveys can reduce considerably the time and cost of maintaining a sufficiently large sample size, and it can improve the accuracy of data entry and coding, although they do have disadvantages, such as sample representativeness and a selection bias. However, previous studies, including Berrens et al. (2003), Hudson et al. (2004), Marta-Pedroso et al. (2007), Fleming and Bowden (2009), and Olsen (2009), among others, find no significant difference in CS estimates based on web-based surveys and those based on conventional mail, telephone, or interview surveys. For further details on web-based surveys and their related problems, see the studies referred to here, and the references therein. As a result, web-based surveys are attracting increasing interest. Considering these prior works, we develop zero-truncated panel Poisson mixture (or random-effects) models by focusing on respondents who have visited a site. Here, we assume that a web-based survey holds similar properties as those of conventional surveys. This sampling procedure is similar to a general population survey, which implies that the number of actual trips is truncated at zero. Thus, our approach reduces the costs of a survey and does not need to consider the problem of endogenous stratification, because adjusting for it will incur an overcorrection when estimating the parameters from the data collected off-site. As a study closely related to ours, Hynes and Hanley (2006) consider a zero-truncated NB model, without correcting for endogenous stratification, by combining data collected from web-based and on-site surveys. However, their approach does not fall within the panel data framework and, thus, cannot be applied to the contingent behavior TCM.

In addition, this study introduces an inverse Gaussian (IG) distribution, which can capture the long (heavy) tail of trip data, to unobserved individual heterogeneity in a random-effects model. Because the NB model has limitations in modeling long-tailed data (see, e.g., Cameron and Trivedi, 2013), we propose a random-effects Poisson-inverse Gaussian (PIG) model, with IG-distributed heterogeneity instead of gamma-distributed heterogeneity, within the panel data framework. Dean et al. (1989) use the PIG model to analyze long-tailed count data in a standard regression framework, finding that it renders a more attractive parametric specification than the NB model. As mentioned above, many studies apply the contingent behavior TCM to estimations of recreational or tourism benefits. However, to the best of our knowledge, no studies have used the PIG model to estimate the parameters of the recreational or tourism demand function. Therefore, we construct estimation methods for zero-truncated panel count data, although it is possible to extend these to data collected on-site. Another advantage of our model is its computational simplicity in estimating demand parameters compared with the multivariate Poisson log-normal, latent class, or random parameter models. This is because the probability mass function of the PIG model is given explicitly using a recursive formula.

Then, we apply the proposed zero-truncated random-effects model to estimate the impact on tourism benefits caused by the accident at Tokyo Electric Power Company's Fukushima Nuclear Power Plant No. 1 (hereafter, “NPP No. 1”). The 2011 Great East Japan Earthquake (hereafter, “the Earthquake”) and the accident involving a radiation leakage at NPP No. 1 brought fame to Fukushima Prefecture, but not in a positive way. Here, we conduct a web-based survey of people who had actually traveled to Fukushima Prefecture in the three-year period from the Earthquake to March 2014. This study seems to be the first attempt to estimate the economic impact of the accident at NPP No. 1 on Fukushima Prefecture's tourism benefits for the period 2011–2014.

The remainder of this paper proceeds as follows. In Section 2, we consider panel Poisson mixture models and their extension to zero truncation. Then, we propose an estimation approach based on the PIG model. In Section 3, we describe the setting of our empirical application to the contingent behavior TCM analysis of the impact of the NPP No. 1 accident on tourism benefits caused, and discuss the estimated results. Lastly, Section 4 concludes the paper.

2. The model for estimation

This section first introduces panel Poisson mixture models. These are essential for estimating recreational benefits by combining RP and SP data, because we need to handle multivariate data on each individual's number of trips. As one of the mixture models, we derive a random-effects Poisson-inverse Gaussian model. Then, we develop zero-truncated random effects models by incorporating left truncation at zero in the panel mixture models. This helps to avoid potential bias in the parameter estimation caused by data truncation. Note that in the following, we do not discuss the fixed-effects models that Englin and Cameron (1996) consider. As Hanley et al. (2003) point out, a fixed-effects model cannot include variables that are invariant across scenarios, and, thus, it is not more appropriate than a random-effects model in the context of combined RP and SP data.

2.1. Panel Poisson mixture models

Let y_{ij} denote the number of trips by individual $i = 1, \cdots, N$ in scenario $j = 1, \cdots, J$. Then, let $\mathbf{x}_{ij} = (x_{1ij}, \cdots, x_{kij})'$ denote the k-dimensional independent variable vector, including a constant, in scenario j. Here, J denotes the number of scenarios. In a manner analogous to the exponential mean function in univariate models, the conditional mean of y_{ij} is assumed to be given by

$$\lambda_{ij} = \mathrm{E}\left(y_{ij} \middle| \mathbf{x}_{ij}\right) = \exp\left(\mathbf{x}_{ij}'\beta\right)$$

where β is a vector of parameters.