Forest Policy and Economics 42 (2014) 42–50

Contents lists available at ScienceDirect

Forest Policy and Economics

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

Confronting problems of method in the study of sustainability

Graham Epstein [a,*], Jessica Vogt [b], Michael Cox [c], Luke Shimek [a]

[a] The Vincent and Elinor Ostrom Workshop in Political Theory and Policy Analysis, Indiana University, Bloomington, IN, USA
[b] Center for the Study of Institutions, Population, and Environmental Change, Indiana University, Bloomington, IN, USA
[c] Environmental Studies Program, Dartmouth College, Hanover, NH, USA

ARTICLE INFO

Article history:
Received 20 August 2013
Received in revised form 15 January 2014
Accepted 19 February 2014
Available online 14 March 2014

Keywords:
Sustainability
Methods
Categorical data analysis
Forest commons
Social–ecological systems

ABSTRACT

Methodological challenges have confounded studies of the commons for much of its recent history. These problems range from a large number of potentially influential variables, difficulties in capturing and measuring these variables, and large sets of interacting and mediating factors. With respect to quantitative methods, a specific dilemma is their use at or beyond their methodologically accepted limits. More specifically, this paper explores the potential implications of applying multinomial logistic regression techniques in small samples. It does so by drawing upon published data from the International Forestry Resources and Institutions (IFRI) program to explore potential problems of method from prominent research findings in the literature on forest commons. A subset of the results is then compared to a similar study of tropical deforestation. The results point to considerable inconsistency in the sign of parameter estimates, and a large number of type II errors. However, they also suggest that type I errors are relatively rare. As a whole, this paper demonstrates the general reliability of multinomial logistic regression in small samples by showing that statistically significant parameters are unlikely to lead policymakers in the wrong direction. Nevertheless it also suggests that this approach is likely to overlook several influential factors, posing potential dilemmas for the development of a theory of sustainability.

1. Introduction

Among the methodologies used by scholars studying the management of common-pool resources (forests, fisheries, irrigation), large-N studies are uniquely capable of producing highly generalizable findings. And yet, large-N studies comprise a relatively small fraction of publications in the commons literature (Poteete et al., 2010). Moreover, many large-N publications rely on a meta-analysis of existing fieldwork (i.e., Cox et al., 2010; Evans et al., 2011; Gutierrez et al., 2011; Ostrom, 1990; Pagdee et al., 2006; Schlager, 1994), the findings of which are potentially confounded by inconsistent coding, missing data, and sampling issues. In recent years, however, the accumulation of data by the collaborative International Forestry Resources and Institutions (IFRI) program has allowed scholars to begin testing hypotheses on larger samples produced by a consistent methodology (Wertime et al., 2007).

Collaborating centers for the IFRI program are now located in Bolivia, Colombia, Guatemala, India, Kenya, Mexico, Nepal, Tanzania, Thailand, Uganda, and the United States, with new centers being established in Ethiopia and China. IFRI is unique among efforts to study forests as it is the only interdisciplinary, long-term monitoring and research program studying forests owned by governments, private organizations, and communities in multiple countries. With a primary focus on community-level characteristics, more than 200 forests and associated user groups have been studied around the world (see http://www.ifriresearch.net/).

The IFRI program is perhaps the single most influential source of information with which commons scholars develop and test hypotheses about the interactions of people, the environment, and institutions. The IFRI database is a great enabler for multiple-methods research, composed of a variety of continuous, ordinal, categorical, and descriptive variables that are collected using a consistent but expandable case-study approach (Wertime et al., 2007). It is thus not surprising that IFRI studies collectively adopt a variety of methodological approaches including comparative case studies (Banana and Gombya-Ssembajjwe, 2000; Fleischman et al., 2010; Karmacharya et al., 2008; Tucker et al., 2007) and quantitative analysis (Andersson and Agrawal, 2011; Chhatre and Agrawal, 2008, 2009; Coleman, 2009; Coleman and Fleischman, 2012; Persha et al., 2011). Collectively, these publications have highlighted the importance of a range of variables in affecting social and biophysical outcomes in community-based forest management systems.

Nevertheless, questions remain about the validity and reliability of IFRI results given unanswered questions concerning the nature of the sample (Coleman, 2009; Ternström et al., 2010), specification of models (Ternström et al., 2010), and the methods that are used to identify relationships between dependent and independent variables. Validity, by which we mean statistical conclusion validity, refers to the extent to which model estimates reflect true relationships. Reliability refers to the consistency of those estimates across applications and contexts.

* Corresponding author. Tel.: +1 812 855 0441.
E-mail address: gepstein@indiana.edu (G. Epstein).

http://dx.doi.org/10.1016/j.forpol.2014.02.006
1389-9341/

This paper responds to these concerns by examining the extent to which multinomial logistic regression (as a representative of categorical modeling techniques) produces consistent estimates of model parameters when applied to IFRI data. The overarching goal is to develop an understanding of the behavior of multinomial logistic regression (MNL) estimates in small samples, and in particular the likelihood that problems of method could (1) lead scholars and policymakers in the wrong direction and/or (2) overlook influential attributes.

This focus is motivated by the fact that recently, quantitative IFRI studies have begun to apply maximum likelihood methods such as binary, ordered, or multinomial logistic regression (i.e., Chhatre and Agrawal, 2008, 2009; Coleman, 2009; Coleman and Fleischman, 2012; Persha et al., 2011), all the while ignoring the "rule of thumb" that such methods are inappropriate in small samples (generally defined as less than 100 cases) (Long, 1997). These methods are, however, particularly appropriate for modeling relationships of interest to commons scholars, such as the correlates of broadly defined success, rule compliance (Kuperan and Sutinen, 1998; Viteri and Chávez, 2007), and provision of public goods (Schündeln, 2013). The question is whether its suitability to these types of questions justifies its application to small samples.

A prominent example of the use of MNL regression is a study of joint outcomes in forest commons that models relationships among four forest outcome categories and five independent variables in a sample of only 80 cases (Chhatre and Agrawal, 2009). While the results are generally supportive of existing institutional theories of the commons, some have questioned their reliability by pointing to potential problems of method (Ternström et al., 2010). These include common concerns such as omitted variables and sampling biases, to which we add additional methodological concerns. Thus, this paper seeks to determine the merits of these critiques using published data from Chhatre and Agrawal (2009), along with other potentially omitted variables drawn from the IFRI database, and a Monte Carlo simulation that approximates sample data.

The results show that the claim that the "analysis is generalizable for the range of values of the independent variables in our data" (Chhatre and Agrawal, 2009, p. SI1) is perhaps slightly overstated, but that the results are unlikely to lead scholars or policymakers in the wrong direction. More specifically, the results show that their analysis is likely to overlook several influential attributes (i.e., type II errors), but are considerably less likely to lead to type I errors. The latter claim is further supported by applying a similar approach to explore the likelihood of type I errors in a small sample study of tropical deforestation (Mahapatr and Kant, 2005). Nonetheless, if the goal of contemporary social–ecological system (SES) theory is to systematically piece together a theory of sustainability based on the effects of a wide range of SES attributes, these models are likely to leave many gaps that demand additional study as part of an interdisciplinary and multimethod research agenda (Poteete et al., 2010).

The remainder of the paper is structured as follows. Section 2 describes some important background information on the original analysis that this paper is reexamining. Section 3 describes the methods used and the results produced from these methods. Section 4 discusses these results, and then explores whether a subset (type I errors) of these results applies to a MNL study of tropical deforestation. Section 5 draws some conclusions from the results and provides general recommendations regarding the use of results from binary, ordered, and categorical data analysis in the study of the commons.

2. Background: forest commons, livelihood benefits, and problems of method

Chhatre and Agrawal (2009) explicitly recognize that forest commons produce multiple social and ecological benefits with a multiplicity of possible synergistic and trade-off relationships. Their analysis focuses on two such benefits – carbon storage and livelihood benefits – that are used to construct four joint outcome categories. Carbon storage is measured as the basal area per hectare of a forest, while the livelihoods index is the product of a principal component analysis based on a forest's contribution to user livelihoods in the form of fodder, fuelwood, timber, and biomass. The outcome categories in the data set are defined by first subtracting the mean values from each observation's carbon storage and livelihood benefit values, recording whether the remainder is positive (1) or negative (0), and then categorizing joint outcomes on the basis of both binary variables in the following way. "Overused forests" have below-average carbon storage and livelihood benefits. "Unsustainable forests" have below-average carbon storage and above-average livelihood benefits. "Deferred-use forests" have above-average carbon storage and below-average livelihood benefits. Finally, "sustainable forest" commons have both above-average carbon storage and livelihood benefits (Chhatre and Agrawal, 2009).

Having constructed the categorical dependent variable, Chhatre and Agrawal (2009) use MNL regression to predict membership in each of the forest categories on the basis of five independent variables. The variables include measures of forest size, ownership, rule-making autonomy, distance from a settlement to the forest, and distance to the nearest administrative center. The results are reported in terms of marginal effects that, in this case, record the partial derivative of continuous or ordinal predictors at their mean while holding all other variables at their mean, or in the case of binary variables, the discrete change in probability. They revealed among a variety of findings that sustainable forests are more likely to occur when users have rule-making autonomy, and that government ownership of forest commons causes users to discount future harvests (i.e., carbon storage) in favor of near-term livelihood benefits.

Problems of method, particularly those associated with quantitative large-N studies, have a long history in the commons literature. In fact, in one of the most influential papers in the field, Agrawal (2003) identifies a range of methodological issues that generally consists of biased sampling, omitted variables, and model specification problems. It is thus not surprising that, in response to Chhatre and Agrawal (2009), Ternström et al. (2010) note that the conclusions of their analysis are debatable on the basis of variable choice, sample selection, and interpretation of the dependent variable, to which we add issues of methodological constraints.

3. Methods and results

3.1. Data collection and analysis

The published data from the original study can be found online,[1] and were used in combination with the IFRI database to extract additional information and produce the data set used for this study. Omitted variables (i.e., variables not present in the initial study) were selected on the basis of the general theoretical prediction that a variable is correlated with the dependent variable and one or more independent variables. Several steps were taken in order to add additional variables to the models. First, IFRI forests were identified on the basis of three variables: (1) the country in which a forest is found, (2) the natural logarithm of its size, and (3) the variable "minimum distance from a settlement to an administrative center." Upon identification of each forest, GPS coordinates of the site and the variable "distance from a settlement to the closest market" were extracted from the database. Distance to market was selected on the basis of the critique that the distance to administrative center is likely correlated with markets (Ternström et al., 2010) and previous studies that have found relationships between markets and forest conditions (Geist and Lambin, 2002; Casse et al., 2004; Agrawal and Chhatre, 2006).

[1] http://sitemaker.umich.edu/ifri/referenced%01datasets.