Available online at www.sciencedirect.com
ScienceDirect
Energy Economics 30 (2008) 1672–1692
Energy Economics
www.elsevier.com/locate/eneco

Applying models for ordinal logistic regression to the analysis of household electricity consumption classes in Rio de Janeiro, Brazil ☆

Mauricio Fuks [a,*], Esther Salazar [b,1]

[a] *Programa de Planejamento Energético (PPE/COPPE) Universidade Federal do Rio de Janeiro, Brazil*
[b] *Department of Statistical Methods of the Universidade Federal do Rio de Janeiro, Brazil*

Received 2 October 2006; received in revised form 19 July 2007; accepted 17 September 2007
Available online 22 September 2007

Abstract

This study applies the proportional odds and partial proportional odds models for ordinal logistic regression to analyze household electricity consumption classes. Micro-data from households situated in the state of Rio de Janeiro during 2004 was used to measure the performance of the models in correctly classifying household electricity consumption classes via sociodemographic, electricity usage and dwelling characteristics. The strategy of using binary logistic regressions to test the main hypothesis of the proportional odds model, suggested by Bender and Grouven, was successful in identifying which of the independent variables could be estimated via the proportional odds assumption. Results indicate that the partial proportional odds models is slightly superior to the more restrictive approach. The study includes probabilistic examples to describe how changes in the independent variables affect the probability of a household belonging to specific classes of electricity consumption. Projections using the final model indicated that the approach may be useful for estimating aggregate household electricity consumption.

JEL classification: C21

Keywords: Probability model; Electricity demand; Logistic regression; Partial proportional odds model; Proportional odds model

☆ The authors would like to thank Fundação de Amparo à Pesquisa do Estado do Rio de Janeiro (FAPERJ) for the postdoctoral fellowship which financed this project and for the support of the Programa de Planejamento Energético (COPPE/UFRJ). The authors are indebted to: Prof. Luiz Fernando Legey (COPPE/UFRJ); an anonymous reviewer; Prof. Richard Williams (Univ. Notre Dame) for guidance with gologit2; Prof. Keith Head (UBC); and Prof. Reinaldo Castro Souza (PUC-RJ) for allowing access to the data. Any errors or omissions are the sole responsibility of the authors.
* Corresponding author. Rua Bulhões de Carvalho, 50/502. Rio de Janeiro, RJ, CEP 22081-000, Brazil. Tel.: +55 21 2287 3261.
E-mail addresses: mfuks@terra.com.br (M. Fuks), esalazar@dme.ufrj.br (E. Salazar).
[1] Corresponding Address: Rua Tonelero 265/801 Rio de Janeiro, RJ, CEP 22030-000, Brazil. Tel.: +55 21 2236 7331.

0140-9883/$ - see front matter © 2007 Elsevier B.V. All rights reserved.
doi:10.1016/j.eneco.2007.09.006

1. Introduction

Micro-data analysis on household electricity demand has generally focused on KWh level consumption (Andersson and Damsgaard, 1999; Halvorsen and Larsen, 1999; Fung et al., 1999; Westley 1992), mainly for the estimation of short-run and long-run elasticities. However, when information on electricity consumption is only available in collapsed form (i.e., consumption classes) traditional OLS approach cannot be used due to the discrete nature and limited range of the dependent variable. One option for researchers with restricted information on electricity consumption is to apply ordinal regression techniques. In this case, although elasticity estimation is not possible via ordinal models, probabilistic examples can be used to illustrate the influence of the independent variables on the likelihood of a household belonging to one of the consumptions classes. Moreover, as in Jung (1993), ordinal regression techniques can also be used to estimate aggregate annual household electricity demand. Thus, we argue that researchers with limited information on household electricity consumption can still obtain informative results via this approach.

A review of the literature on cross-sectional modeling of household electricity demand indicates only one article which has applied the logistic model in this field (Jung, 1993). This study applies the proportional odds model (POM) and the partial proportional odds model (PPOM) for ordinal logistic regression in order to analyze household electricity consumption classes in the state of Rio de Janeiro, Brazil. The data was obtained from face-to-face interviews with approximately 2000 households during the year of 2004. The questionnaire included modules related to sociodemographic, dwelling characteristics and electricity usage information.[2]

The objectives of this research are: (i) to find the best models for the three approaches: proportional odds (POM), partial proportional odds (PPOM) and generalized ordered logit; (ii) compare the final models via the likelihood ratio (LR) test, Akaike Information Criteria (AIC), Schwarz's Bayesian Information Criterion (BIC); (iii) test the strategy suggested by Bender and Grouven (1998) to identify which of the independent variables (IVs) can be summarized by a single coefficient (i.e., by means of the proportional odds); (iv) evaluate and compare the performance of the proportional odds and the partial proportional odds models in correctly classifying the households' electricity consumption classes based on their sociodemographic; electricity usage and dwelling related information; (v) describe, via probabilistic examples, how changing the values of the independent variables affect the probabilities of a household belonging to specific electricity consumption classes and (vi) evaluate the performance of the final model in predicting actual annual aggregate household electricity demand.

Regarding objective (iv), most statistical softwares provide output which include observed and predicted classification for each observation, thus allowing us to calculate not only the accuracy rate of the entire sample, but also the rate for each consumption class. Lastly, split-sample validation will be implemented to evaluate the stability of the final model.

[2] The data pertains only to households covered by a company which does not supply electricity to the entire state of Rio de Janeiro. Thus, conclusions from this study are not applicable to the total number of households in the state. Additionally, illegal household connections, locally known as *gatos*, are abundant. Since such households are not listed as clients, they could not participate in the sample selection process. Values of the dependent variable consumption class (CLASS) were obtained directly from the company's database. The authors, however, did not have access to KWh level consumption. Given the confidential nature of the data, the authors have agreed not to specify the company nor the specific municipalities of the households. See Souza (2004) for a complete description of the project.