Economics Letters 18 (1985) 361–365
North-Holland

A TOBIT-TYPE ESTIMATOR FOR THE CENSORED POISSON REGRESSION MODEL

Joseph V. TERZA *

University of Georgia, Athens, GA 30602, USA

Received 2 January 1985

The Poisson counterpart to the Tobit model is presented. Formulae for the gradient and Hessian of the relevant log-likelihood function are given and incorporated into a Newton–Raphson optimization algorithm. The asymptotic covariance matrix of the estimator is detailed. As an illustration, the NR algorithm is applied to a model of individual shopping behavior.

1. Introduction

Considerable attention has been given to Poisson regression models in the econometric and statistical literatures. Advances in the theory of estimation can be found in the works of Jorgenson (1961), Frome et al. (1973) and Gourieroux et al. (1984). The technique has been successfully applied by Flowerdew and Aitkin (1982), Hausman et al. (1984) and Marlow et al. (1984). The present note augments the existing literature by detailing a Newton–Raphson algorithm for a Poisson regression model whose dependent variable is censored above a known threshold. This is the Poisson counterpart to the now famous Tobit method proposed by Tobin (1958) for which Amemiya (1973) later supplied workable gradient and Hessian formulae.

2. Poisson regression

Consider the Poisson random variable y_t where $t = 1,\dots,T$. The pdf of y_t is

$$f_t(y_t) = m_t^{y_t} \exp\{-m_t\}/y_t!,$$

where $m_t = \exp\{X_t\beta\}$, X_t is a $1 \times K$ vector of exogenous variables and β is an unknown $K \times 1$ vector of parameters to be estimated. The parameter vector β can be estimated by one of two popular methods. The first is iterative maximum likelihood estimation via a Newton–Raphson (NR) algorithm. The NR update of β at the $(i+1)$th iteration is

$$\beta_{i+1} = \beta_i - H^{-1}G, \tag{1}$$

where $G(K \times 1)$ is the gradient and $H(K \times K)$ is the Hessian with

$$G = \sum_{t=1}^{T} (m_t - y_t) X_t' \quad \text{and} \quad H = -\sum_{t=1}^{T} m_t X_t' X_t.$$

* The excellent research assistance of A. Ason Okoruwa is gratefully acknowledged.

0165-1765/85/$3.30 © 1985, Elsevier Science Publishers B.V. (North-Holland)

Alternatively, the maximum likelihood estimator of β can be obtained by the Iterative Weighted Least Squares (IWLS) method developed by Frome et al. (1973). The IWLS update of β at the $(i+1)$th iteration is

$$\beta_{i+1} = \beta_i + \left[\sum_{t=1}^{T} (1/m_t) X_t' X_t \right]^{-1} \left[\sum_{t=1}^{T} ((m_t - y_t)/m_t) X_t' \right]. \tag{2}$$

3. The censored model

Suppose that the value of y_t is observable for the tth member of the sample iff $y_t < C$ where C is a known positive integer. Let d_t be a dummy variable defined such that

$$\begin{aligned} d_t &= 1 \quad \text{iff} \quad y_t < C, \\ &= 0 \quad \text{otherwise.} \end{aligned}$$

Censoring of this kind may be imposed on the data by survey design, or it may reflect relevant theoretical or institutional constraints. In this case both IWLS and the NR algorithm defined in (1) will yield biased and inconsistent estimates of β. A consistent estimate can, however, be obtained by maximizing the true likelihood function.

The likelihood of the observed outcome for the tth member of the sample is $f_t(y_t)$ iff $d_t = 1$ and $F_t(C) = 1 - \sum_{i=0}^{C-1} f_t(i)$ otherwise. The log-likelihood of the sample can therefore be written

$$L(\beta) = \sum_{t=1}^{T} \{ d_t \log f_t(y_t) + (1 - d_t) \log F_t(C) \}. \tag{3}$$

Table 1
Independent variable definitions.

EMPL	=	1 if employed, 0 otherwise,
MART	=	1 if married, 0 otherwise,
AGE	=	respondent's age,
SEX	=	1 if male,
SCHC	=	last grade of schooling completed,
RACE	=	1 if white, 0 otherwise,
LADD	=	length of time at current address (years),
NIND	=	total number of individuals in the household,
TINC	=	total household income (dollars),
DRTM	=	driving time from respondent's residence to retail center,
DPR1	=	1 if retail center contains respondent's favorite department store with respect to variety and selection of merchandise, 0 otherwise,
DPR2	=	1 if center has favorite department store with respect to the quality of merchandise, 0 otherwise,
DPR3	=	1 if center has favorite department store with respect to quality for price, 0 otherwise,
DPR4	=	1 if center has department store with most helpful and friendly sales people, 0 otherwise,
SPR1–*SPR4*	=	variables analogous to *DPR1*–*DPR4* for specialty stores.

Table 2
Estimation results. [a]

Variable	Frome's method	Censored NR
CONSTANT	−1.199	−1.036
EMPL	−0.602	−0.668
	(4.308)	(−4.773)
MART	0.134	0.174
	(0.815)	(1.054)
AGE	−0.038	−0.039
	(−1.360)	(−1.367)
AGE^2	0.2E−04	1E−04
	(0.502)	(0.470)
SEX	−0.057	−0.083
	(−0.417)	(−0.616)
SCHC	0.083	0.082
	(3.631)	(3.600)
RACE	0.702	0.744
	(3.896)	(4.147)
LADD	0.007	0.006
	(0.507)	(0.392)
NIND	−0.284	−0.31
	(−5.297)	(−5.453)
TINC	3E−05	3E−05
	(4.739)	(4.837)
DRTM	−0.014	−0.015
	(3.468)	(−3.166)
DPR1	0.578	0.634
	(3.463)	(3.825)
DPR2	0.145	0.119
	(0.831)	(0.693)
DPR3	0.377	0.412
	(2.538)	(2.774)
DPR4	−0.010	−0.119
	(−0.723)	(−0.859)
SPR1	−0.074	−0.04164
	(−0.402)	(−0.217)
SPR2	0.645	0.710
	(3.449)	(3.711)
SPR3	0.101	0.080
	(0.471)	(0.361)
SPR4	−0.192	−0.264
	(−0.922)	(−1.232)
Likelihood ratio [b]	–	318.92
$T = 828$		

[a] Asymptotic t-statistics in parentheses.
[b] Null hypothesis: all coefficients in β except the constant are equal to zero.

The NR algorithm defined in (1) can be used to maximize (3) after respectively replacing G and H with

$$\sum_{t=1}^{T} \{ d_t(y_t - m_t) - (1-d_t)(\Phi_{1t}/F_t(C)) \} X_t', \quad \text{and}$$

$$\sum_{t=1}^{T} \left\{ d_t m_t + (1-d_t)\left[(\Phi_{2t}/F_t(C)) + (\Phi_{1t}/F_t(C))^2 \right] \right\} X_t' X_t, \quad \text{where}$$

$$\Phi_{1t} = \sum_{i=0}^{C-1} f_t(i)(i-m_t) \quad \text{and} \quad \Phi_{2t} = \sum_{i=0}^{C-1} f_t(i)\left[(i-m_t)^2 - m_t \right].$$

The asymptotic covariance matrix of the maximum likelihood estimator of β is

$$\left[\sum_{t=1}^{T} \{ (1-F_t(C)) m_t + (\Phi_{2t} + (\Phi_{1t}^2/F_t(C))) \} X_t' X_t \right]^{-1}.$$

To obtain an initial value of β, assign a dependent variable value of C to all observations for which $d_t = 0$, then apply either IWLS or the uncensored NR algorithm.

4. An application

For the purpose of illustration, consider the censored Poisson regression whose dependent variable is obtained from the following survey question: 'How many times have you been to shopping area Q in the past thirty days?', with the following possible responses: (a) zero, (b) one, (c) two, and (d) three or more. The data for this application were drawn from a 1977 survey of the Atlanta SMSA conducted for the Atlanta Journal/Constitution newspaper. Information was collected on a wide range of economic and demographic variables. Many of these were used as exogenous variables in the specification of the Poisson regression equation. Table 1 gives the definitions of all variables included in the analysis. The response frequencies for the dependent variable were: (a) 682, (b) 47, (c) 34, and (d) 65.

Initially, censoring was ignored and responses above the censoring threshold ($C = 3$) were assigned a dependent variable value of 3. The Poisson regression program developed by Frome (1981) was then applied to the doctored data and the results were used as initial values for the NR algorithm described in section 3. The initial values along with the NR results are shown in table 2. As a means of investigating the potential bias when censoring is ignored, predicted patronization rates for the average respondent were computed using the initial values and the NR results. These predicted values are 0.231 and 0.111, respectively, representing a bias in excess of 100%. This result is particularly interesting in view of the fact that only a small percentage of the sample (7.9%) was censored.

References

Amemiya, T., 1973, Regression analysis when the dependent variable is truncated normal, Econometrica 41, 997–1016.

Flowerdew, R. and M. Aitkin, 1982, A method of fitting the gravity model based on the Poisson distribution, Journal of Regional Science 22, 191–202.

Frome, E.L., 1981, PREG: A computer program for Poisson regression analysis (Biomedical Computing Technology Information Center, Vanderbilt University, Nashville, TN).

Frome, E.L., M.H. Kutner and J.J. Beauchamp, 1973, Regression analysis of Poisson-distributed data, Journal of the American Statistical Association 68, 935–940.

Gourieroux, C., A. Monfort and A. Trognon, 1984, Pseudo maximum likelihood methods: Applications to Poisson models, Econometrica 52, 701–720.

Hausman, J., B.H. Hall and Z. Griliches, 1984, Econometric models for count data with an application to the patents-R&D relationship, Econometrica 52, 909–938.

Jorgenson, D.W., 1961, Multiple regression analysis of a Poisson process, Journal of the American Statistical Association 56, 235–245.

Marlow, M.L., J.P. Link and R.P. Trost, 1984, Market structure and rivalry in financial markets: Some new evidence with a nonlinear model, Review of Economics and Statistics 66, forthcoming.

Tobin, J., 1958, Estimation of relationships for limited dependent variables, Econometrica 26, 24–36.