Consistent Standard Errors in Panel Tobit with Autocorrelation ‡

Meghan R. Busse[1]
Haas School of Business, UC Berkeley

Andrew B. Bernard[2]
Tuck School of Business at Dartmouth
National Bureau of Economic Research

First version: March 2002
This version: August 2003

Abstract

This paper derives consistent standard errors for a panel Tobit model in the presence of correlated errors. The problem is framed in the context of Newey and West (1987), considering the Tobit model as a special case of a GMM estimator.

JEL codes: C23, C24

Keywords: Tobit, panel data, standard error estimation, autocorrelation, censoring

[*] Haas School of Business, 2220 Piedmont Ave., Berkeley, CA 94720-1900; tel (510) 643-1426; fax (510) 643-1420; email: `meghan@haas.berkeley.edu; http://faculty.haas.berkeley.edu/meghan/`

[†] 100 Tuck Hall, Hanover, NH 03755, tel: (603) 646-0302, fax: (603) 646-1308, email: `andrew.b.bernard@dartmouth.edu, http://www.andrew.bernard.org`

[‡] We thank Steve Berry, Ed Kaplan, and Ben Polak for helpful discussions. Elena Krasnokutskaya provided valuable research assistance. Bernard's contribution to the research supported in part by the National Science Foundation (SES-0241474).

1 Introduction

When estimating econometric models with panel data sets, one often has to be concerned with whether the errors in the model are correlated over time. If so, one must account for this autocorrelation in creating estimates of standard errors in order to have correct confidence intervals for coefficient estimates.

In linear models, the method proposed by Newey and West (1987) is the most commonly chosen method for adjusting standard errors to account for serial correlation. However, the method of Newey and West is far more generally applicable than just to linear models. Their article is written in the context of a GMM estimator, which Newey and McFadden (1994) show is a class that includes many kinds of estimators, including maximum likelihood estimators if their first order conditions are thought of as moments. Thus the Newey-West standard error formula can be used for any maximum likelihood estimator. This paper considers one such example, although the derivation presented here could be used as a guide for adjusting standard errors for any maximum likelihood or, less directly, another GMM estimator.

This paper considers the estimation of the covariance matrix for a Tobit model in which the errors may be correlated over time. For simplicity, the notation used in the derivation of the estimator assumes a simple time-series data set. Since the most likely applications of the model are for panel data, the paper also discusses how the estimator would need to be changed to accomodate panel data.

The remainder of the paper proceeds as follows. Section 2 describes the econometric background to this problem. Section 3 derives the formula for Tobit standard errors under autocorrelation. Section 4 expands the formula to accommodate panel data and random effects. Section 5 concludes.

2 Background

Robinson (1982) gives early consideration to the problem of estimating a Tobit model with serially correlated errors. He gives an expression for the limiting covariance matrix, although it does not lend itself readily to estimation.

Newey and West (1987) simplified considerably the problem of estimating covariance matrices in the presence of serial correlation. Their paper presents a simple, positive definite estimator for the covariance matrix of a GMM (Generalized Method of Moments) estimator. GMM estimators are a large class which includes many commonly used estimators, including OLS and instrumental variables, as special cases. In particular, maximum likelihood estimators can be considered as GMM estimators if the solutions to their first order conditions are considered as moments. (See Newey and McFadden (1994) for a very complete treatment of the subject.) Estrella and Rodrigues (1998) apply Newey and West (1987) to the estimation of standard errors for a Probit model by considering Probit as a GMM estimator.

This paper derives the consistent standard errors for a Tobit model. We also apply Newey and West (1987) by considering a Tobit model as a GMM estimator. As a result, we are able to derive a formula that is considerably more useful for empirical implementation than that derived by Robinson (1982).

3 Tobit standard errors under serial correlation

The Tobit model can be considered as a GMM estimator whose moments are the first order conditions of the log likelihood function. Following Robinson's (1982) notation, consider the following model:

$$y_t = \begin{cases} \beta' x_t + \epsilon_t, & \text{if } \beta' x_t + \epsilon_t > 0 \\ 0, & \text{otherwise} \end{cases} \tag{1}$$

for $t = 1, 2, ..., T$, where β and x_t are $k \times 1$ column vectors and ϵ_t is distributed $N(0, \sigma^2)$. We are interested in the case in which the ϵ_t's are not independently distributed. As Robinson (1982) shows, β can be consistently estimated with the Tobit model.

Define an indicator variable $w_t = I(y_t > 0)$. The log likelihood of the Tobit model is given by

$$Q_T(\theta) = \frac{1}{T} \sum_{t=1}^{T} (1 - w_t) \ln(1 - F(\beta' x_t, \sigma)) + w_t \ln f(y_t - \beta' x_t, \sigma), \tag{2}$$

where $\theta = (\beta, \sigma)$ and $f(., \sigma)$ and $F(., \sigma)$ are the probability density function and cumulative density function, respectively, of $N(0, \sigma^2)$. [1]

The first order conditions of this log likelihood function are given by a $(k+1) \times 1$ column vector. The first k elements of this vector are given by the following, where $\psi(\beta' x_t, \sigma) = \frac{f(\beta' x_t, \sigma)}{1 - F(\beta' x_t, \sigma)}$:

$$\frac{\partial Q_T(\theta)}{\partial \beta_i} = \frac{1}{T} \sum_{t=1}^{T} (1 - w_t) \left[-\psi(\beta' x_t, \sigma) x_{ti} \right] + w_t \left[\frac{1}{\sigma^2} (y_t - \beta' x_t) x_{ti} \right], \forall i = 1, ..., k. \tag{3}$$

[1] In the Probit model, it is possible to identify only β/σ, so without loss of generality, one can assume that $\sigma = 1$. In the Tobit model, σ must be estimated.

The $(k+1)$th element in the vector is given by

$$\frac{\partial Q_T(\theta)}{\partial \sigma} = \frac{1}{T}\sum_{t=1}^{T}(1-w_t)\left[\psi(\beta' x_t, \sigma)\frac{\beta' x_t}{\sigma}\right] + w_t\left[\frac{-1}{\sigma} + \frac{(y_t - \beta' x_t)^2}{\sigma^3}\right]. \tag{4}$$

In Newey and West's terminology, this vector is $h_T(\theta) = \frac{1}{T}\sum_{t=1}^{T} h_t(\theta)$.

The general form of Newey and West's (1987) covariance estimator is given by

$$V_T = (H_T' W_T H_T)^{-1} H_T' W_T S_T W_T H_T (H_T' W_T H_T)^{-1}. \tag{5}$$

W_T is a weighting matrix which here, for simplicity, we assume is an identity matrix. H_T is the expected value of the partial derivatives of $h_t(\theta)$, the moment conditions or first order conditions, with respect to θ.

$$H_T = \frac{1}{T}\sum_{t=1}^{T} E(h_{t\theta}(\theta^*)) \tag{6}$$

In a GMM model with p moment condtions and r parameters to be estimated, this would be a $p \times r$ matrix. Since the moments in the Tobit case are the first order condtions of the log likelihood function, the number of moment conditions equals the number of parameters to be estimated, and H_T is square (and symmetric) matrix; in the current example, the dimension of H_T is $(k+1) \times (k+1)$. H_T can be estimated by its sample analog,

$$\hat{H}_T = \frac{1}{T}\sum_{t=1}^{T} h_{t\theta}(\hat{\theta}). \tag{7}$$

Newey and West's (1987) most important contribution is their estimator for the last term in the covariance matrix, S_T.

$$S_T = \frac{1}{T}\sum_{s=1}^{T}\sum_{t=1}^{T} E\left(h_t(\theta^*) h_s(\theta^*)'\right) \tag{8}$$

Newey and West's (1987) estimator for this term is

$$\hat{S}_T = \hat{\Omega}_0 + \sum_{j=1}^{m}(1 - \frac{j}{m+1})(\hat{\Omega}_j + \hat{\Omega}_j'), \tag{9}$$

where $\hat{\Omega}_j = \frac{1}{T}\sum_{t=1}^{T} h_t(\hat{\theta}) h_{t-j}'(\hat{\theta})$ and m is the number of sample autocovariances used to estimate S_T. If the number of nonzero autocorrelations of $h_t(\theta^*)$ is known a priori, then that can be used as m. Other procedures for choosing m are covered elsewhere in the literature, but are beyond the scope of this paper.

For the Tobit model, the $\hat{\Omega}_j$'s, and therefore $\hat{S}_T$, can be calculated using the expression for h_t given in Equations 3 and 4, substituting the estimated $\hat{\beta}$ and $\hat{\sigma}$ for β and σ. $h_{t\theta}$, necessary for estimating H_T, can be found by taking partial derivatives of Equations 3 and 4 with respect to each element of the coefficient vector $\theta = (\beta, \sigma)$. The form of H_T is

$$H_T = \left[\begin{array}{c|c} \frac{\partial Q_T(\theta)}{\partial\beta_i\partial\beta_j} & \frac{\partial Q_T(\theta)}{\partial\beta_i\partial\sigma} \\ \hline \frac{\partial Q_T(\theta)}{\partial\beta_i\partial\sigma} & \frac{\partial Q_T(\theta)}{\partial\sigma\partial\sigma} \end{array}\right] \tag{10}$$

where the top left quadrant is $k \times k$, the bottom right quadrant is 1×1 and the remaining quadrants are conformable. The elements of H_T are:

$$\frac{\partial Q_T(\theta)}{\partial\beta_i\partial\beta_j} = \frac{1}{T}\sum_{t=1}^{T}(1-w_t)\psi(\beta' x_t,\sigma)\left[\frac{\beta' x_t}{\sigma^2} - \psi(\beta' x_t,\sigma)\right]x_{ti}x_{tj} - w_t\left(\frac{1}{\sigma^2}\right)x_{ti}x_{tj}, \forall i,j = 1,...,k, \tag{11}$$

$$\frac{\partial Q_T(\theta)}{\partial\beta_i\partial\sigma} = \frac{1}{T}\sum_{t=1}^{T}(1-w_t)\psi(\beta' x_t,\sigma)\left(\frac{1}{\sigma}\right)\left[\psi(\beta' x_t,\sigma)\beta' x_t - \left(\frac{\beta' x_t}{\sigma}\right)^2 + 1\right]x_{ti} - w_t\left[\frac{2}{\sigma^3}(y_t - \beta' x_t)\right]x_{ti}, \forall i = 1,...,k, \tag{12}$$

$$\frac{\partial Q_T(\theta)}{\partial\sigma\partial\sigma} = \frac{1}{T}\sum_{t=1}^{T}(1-w_t)\psi(\beta' x_t,\sigma)\left(\frac{1}{\sigma^2}\right)\left[\frac{(\beta' x_t)^3}{\sigma^2} - \psi(\beta' x_t,\sigma)(\beta' x_t)^2 - 2\beta' x_t\right] + w_t\left(\frac{1}{\sigma^2}\right)\left[1 - \frac{3}{\sigma^2}(y_t - \beta' x_t)^2\right]. \tag{13}$$

As mentioned in Equation 7, H_T can be estimated by subsituting $(\hat{\beta}, \hat{\sigma})$ for (β, σ) in Equations 11 through 13.

While the expressions for $\hat{S}_T$ and $\hat{H}_T$ may appear daunting at first, they are in fact relatively easily assembled with matrix based software such as MATLAB or Gauss using the sample data and estimates of β and σ produced by another package such as Stata.

4 Panel Data and Random Effects

Tobit models are most often applied in the context of cross sectional or panel data. Thus, autocorrelation in a Tobit model is most likely to arise in a panel, rather than in a univariate time series. With panel data, it is often desirable to have a model that allows for the individuals, firms, countries, or groups that define a cross-sectional unit of the data to differ systematically in the value of the dependent variable for reasons unobserved to the econometrician. In a Tobit model, such individual-specific, time-invariant effects are modelled as random effect, since a fixed effects model is plagued by the incidental parameters problem (Wooldridge, 2002).

In this section, we expand the autocorrelation correction for standard errors to a panel data setting with random effects. First, we expand the model of Equation 1.

$$y_{it} = \begin{cases} \beta' x_{it} + \nu_i + \epsilon_{it}, & \text{if } \beta' x_{it} + \nu_i + \epsilon_{it} > 0 \\ 0, & \text{otherwise} \end{cases} \tag{14}$$

for $t = 1, 2, ..., T$, where β and x_{it} are $k \times 1$ column vectors, ϵ_{it} is distributed $N(0, \sigma_\epsilon^2)$, and ν_i is distributed $N(0, \sigma_\nu^2)$. We assume that $E(\nu_i \nu_j) = 0$, $E(\nu_i \epsilon_{it}) = 0$, and that $E(\epsilon_{it}\epsilon_{jt}) = 0, \forall i \neq j$. We are interested in the case in which ϵ_{is} and ϵ_{it} are not independently distributed.

With the addition of ν_i to the model, the likelihood function becomes somewhat more complicated than that of a simple Tobit model because the distribution of the unobserved component of the model for any one observation is linked through ν_i to the unobserved components of all the other observations in the same cross-sectional unit. The likelihood function, $\Pr(y_i|x_i)$, for cross-sectional unit i is:

$$L_i = \int_{-\infty}^{\infty} \left\{ \prod_{t=1}^{T_i} \left[f(y_{it} - \beta' x_{it} - \nu_i, \sigma_\epsilon) \right]^{w_{it}} \left[\Phi\left(\frac{-\beta' x_{it} - \nu_i}{\sigma_\epsilon} \right) \right]^{(1-w_{it})} \right\} f(\nu_i, \sigma_\nu)\, \mathrm{d}\nu_i \tag{15}$$

where $\Phi(.)$ is the standard normal c.d.f. For ease of notation, we denote the product contained within the curly braces as J_i. Equation 15 can thus be written more compactly as

$$L_i = \int_{-\infty}^{\infty} J_i\, f(\nu_i, \sigma_\nu)\, \mathrm{d}\nu_i. \tag{16}$$

The likelihood function for the whole sample is simply the product of the L_i's over the N cross-sectional units, and the log likelihood is

$$\mathcal{L} = \sum_{i=1}^{N} \ln L_i \tag{17}$$

Note that the log likelihood in Equation 17 does not collapse to a sum, as it would in the case of a purely cross-sectional or time series Tobit. (See Equation 2.) This is because the likelihood function for a given cross-sectional unit (Equation 15) is an *integral* of a product instead of just a product. The log operater cannot be carried through the integral sign, so the natural log of the likelihood function in Equation 17 is a sum of logs of the integral in Equation 15.

We can derive the Newey-West standard error adjustment for the random effects model in much the same way as we derived them for the simple Tobit model in Section 3. The vector of parameters to be estimated is $\theta = (\beta', \sigma_\epsilon, \sigma_\nu)$, and the moment conditions are the first order condtions of $\mathcal{L}$ with respect to θ. Differentiating Equation 16 with respect to the mth element of the vector β yields:

$$\frac{\partial L_i}{\partial \beta_m} = \int_{-\infty}^{\infty} \left\{ \sum_{t=1}^{T_i} \left[w_{it}(y_{it} - \beta' x_{it} - \nu_i) \frac{x_{itm}}{\sigma_\epsilon^2} + (1 - w_{it}) \psi \left(\frac{-\beta' x_{it} - \nu_i}{\sigma_\epsilon} \right) \left(\frac{-x_{itm}}{\sigma_\epsilon} \right) \right] \right\} \cdot J_i f(\nu_i, \sigma_\nu) \, \mathrm{d}\nu_i \tag{18}$$

where $\psi\left(\frac{-\beta' x_{it} - \nu_i}{\sigma_\epsilon}\right) = \phi\left(\frac{-\beta' x_{it} - \nu_i}{\sigma_\epsilon}\right) / \Phi\left(\frac{-\beta' x_{it} - \nu_i}{\sigma_\epsilon}\right)$ and $\phi(.)$ is the standard normal p.d.f. For ease of exposition, we define the summation contained within the curly braces as R_{im}. Thus equation 18 can be rewritten as

$$\frac{\partial L_i}{\partial \beta_m} = \int_{-\infty}^{\infty} R_{im} \, J_i f(\nu_i, \sigma_\nu) \, \mathrm{d}\nu_i. \tag{19}$$

Differentiating equation 17 with respect to β_m and substituting in equation 19 yields the moment condition with respect to β_m:

$$\frac{\partial \mathcal{L}}{\partial \beta_m} = \sum_{i=1}^{N} \frac{1}{L_i} \int_{-\infty}^{\infty} R_{im} \, J_i \, f(\nu_i, \sigma_\nu) \, \mathrm{d}\nu_i. \tag{20}$$

In the random effects Tobit model, there are two variances to be estimated, σ_ϵ and σ_ν. Differentiating equation 16 with respect to σ_ϵ yields

$$\frac{\partial L_i}{\partial \sigma_\epsilon} = \int_{-\infty}^{\infty} S_i \, J_i \, f(\nu_i, \sigma_\nu) \, \mathrm{d}\nu_i, \tag{21}$$

where

$$S_i = \sum_{t=1}^{T_i} w_{it} \left[\frac{1}{\sigma_\epsilon^3} (y_{it} - \beta' x_{it} - \nu_i)^2 - \frac{1}{\sigma_\epsilon} \right] + (1 - w_{it}) \psi \left(\frac{-\beta' x_{it} - \nu_i}{\sigma_\epsilon} \right) \left(\frac{\beta' x_{it} + \nu_i}{\sigma_\epsilon^2} \right). \tag{22}$$

Thus the moment condition corresponding to σ_ϵ can be written as

$$\frac{\partial \mathcal{L}}{\partial \sigma_\epsilon} = \sum_{i=1}^{N} \frac{1}{L_i} \int_{-\infty}^{\infty} S_i \, J_i \, f(\nu_i, \sigma_\nu) \, \mathrm{d}\nu_i. \tag{23}$$

Differentiating equation 16 with respect to σ_ν, the remaining element of the parameter vector θ, yields

$$\frac{\partial L_i}{\partial \sigma_\nu} = \int_{-\infty}^{\infty} \left[\frac{\nu_i^2}{\sigma_\nu^3} - \frac{1}{\sigma_\nu} \right] J_i \, f(\nu_i, \sigma_\nu) \, \mathrm{d}\nu_i \tag{24}$$

Thus the final moment condition, corresponding to σ_ν, is

$$\frac{\partial \mathcal{L}}{\partial \sigma_\nu} = \sum_{i=1}^{N} \frac{1}{L_i} \int_{-\infty}^{\infty} \left[\frac{\nu_i^2}{\sigma_\nu^3} - \frac{1}{\sigma_\nu} \right] J_i\, f(\nu_i, \sigma_\nu)\, \mathrm{d}\nu_i. \tag{25}$$

In order to implement the Newey-West standard error formula, we also need expressions for the Hessian matrix of the log likelihood function. The form of the Hessian matrix is

$$H_T = \left[\begin{array}{c|c|c} \frac{\partial \mathcal{L}}{\partial \beta_m \partial \beta_p} & \frac{\partial \mathcal{L}}{\partial \beta_m \partial \sigma_\epsilon} & \frac{\partial \mathcal{L}}{\partial \beta_m \partial \sigma_\nu} \\ \hline \frac{\partial \mathcal{L}}{\partial \beta_m \partial \sigma_\epsilon} & \frac{\partial \mathcal{L}}{\partial \sigma_\epsilon \partial \sigma_\epsilon} & \frac{\partial \mathcal{L}}{\partial \sigma_\epsilon \partial \sigma_\nu} \\ \hline \frac{\partial \mathcal{L}}{\partial \beta_m \partial \sigma_\nu} & \frac{\partial \mathcal{L}}{\partial \sigma_\epsilon \partial \sigma_\nu} & \frac{\partial \mathcal{L}}{\partial \sigma_\nu \partial \sigma_\nu} \end{array} \right] \tag{26}$$

where the top left submatrix is $k \times k$, the four bottom right submatrices are 1×1 and the remaining submatrices are conformable. The submatrices of H_T are given by the following expressions.

To obtain $\frac{\partial \mathcal{L}}{\partial \beta_m \partial \beta_p}$ we differentiate equation 20 with respect to β_p.

$$\begin{aligned} \frac{\partial \mathcal{L}}{\partial \beta_m \partial \beta_p} = & \sum_{i=1}^{N} -\frac{1}{L_i^2} \frac{\partial L_i}{\partial \beta_p} \int_{-\infty}^{\infty} R_{im}\, J_i\, f(\nu_i, \sigma_\nu)\, \mathrm{d}\nu_i \\ & + \sum_{i=1}^{N} \frac{1}{L_i} \int_{-\infty}^{\infty} \left[\frac{\partial R_{im}}{\partial \beta_p} J_i + R_{im} \frac{\partial J_i}{\partial \beta_p} \right] f(\nu_i, \sigma_\nu)\, \mathrm{d}\nu_i. \end{aligned} \tag{27}$$

Of the elements of this equation, $\frac{\partial L_i}{\partial \beta_p}$ is given by equation 19 and $\frac{\partial J_i}{\partial \beta_p} = R_{ip}\, J_i$. Only $\frac{\partial R_{im}}{\partial \beta_p}$ is yet to be derived.

$$\begin{aligned} \frac{\partial R_{im}}{\partial \beta_p} = \sum_{t=1}^{T_i} & \quad w_{it} \left(-\frac{x_{itp}\, x_{itm}}{\sigma_\epsilon^2} \right) \\ & + (1 - w_{it}) \left[\frac{-\beta' x_{it} - \nu_i}{\sigma_\epsilon} + \psi\left(\frac{-\beta' x_{it} - \nu_i}{\sigma_\epsilon} \right) \right] \psi\left(\frac{-\beta' x_{it} - \nu_i}{\sigma_\epsilon} \right) \left(\frac{-x_{itp}\, x_{itm}}{\sigma_\epsilon^2} \right). \end{aligned} \tag{28}$$

To obtain the Hessian submatrix $\frac{\partial \mathcal{L}}{\partial \beta_m \partial \sigma_\epsilon}$ we differentiate equation 20 with respect to σ_ϵ.

$$\begin{aligned} \frac{\partial \mathcal{L}}{\partial \beta_m \partial \sigma_\epsilon} = & \sum_{i=1}^{N} -\frac{1}{L_i^2} \frac{\partial L_i}{\partial \sigma_\epsilon} \int_{-\infty}^{\infty} R_{im}\, J_i\, f(\nu_i, \sigma_\nu)\, \mathrm{d}\nu_i \\ & + \sum_{i=1}^{N} \frac{1}{L_i} \int_{-\infty}^{\infty} \left[\frac{\partial R_{im}}{\partial \sigma_\epsilon} J_i + R_{im} \frac{\partial J_i}{\partial \sigma_\epsilon} \right] f(\nu_i, \sigma_\nu)\, \mathrm{d}\nu_i. \end{aligned} \tag{29}$$

Of the elements of this equation, $\frac{\partial L_i}{\partial \sigma_\epsilon}$ is given by equation 21 and $\frac{\partial J_i}{\partial \sigma_\epsilon} = S_i\, J_i$. Only $\frac{\partial R_{im}}{\partial \sigma_\epsilon}$ is yet to be derived.

$$\frac{\partial R_{im}}{\partial \beta_p} = \sum_{t=1}^{T_i} \quad w_{it}\left(-\frac{2}{\sigma_\epsilon^3}\right)(y_{it}-\beta' x_{it}-\nu_i)\, x_{itm} \\ +(1-w_{it})\psi\left(\frac{-\beta' x_{it}-\nu_i}{\sigma_\epsilon}\right)\left(\frac{x_{itm}}{\sigma_\epsilon^2}\right)\left[1-\frac{(\beta' x_{it}+\nu_i)^2}{\sigma_\epsilon^2}+\psi\left(\frac{-\beta' x_{it}-\nu_i}{\sigma_\epsilon}\right)\frac{\beta' x_{it}+\nu_i}{\sigma_\epsilon}\right]. \tag{30}$$

To derive the Hessian submatrix $\frac{\partial \mathcal{L}}{\partial \beta_m \partial \sigma_\nu}$ we differentiate equation 20 with respect to σ_ν.

$$\frac{\partial \mathcal{L}}{\partial \beta_m \partial \sigma_\nu} = \sum_{i=1}^{N} -\frac{1}{L_i^2}\frac{\partial L_i}{\partial \sigma_\nu}\int_{-\infty}^{\infty} R_{im}\, J_i\, f(\nu_i,\sigma_\nu)\,\mathrm{d}\nu_i + \sum_{i=1}^{N}\frac{1}{L_i}\int_{-\infty}^{\infty}\left[R_{im}\, J_i \frac{\partial f(\nu_i,\sigma_\nu)}{\partial \sigma_\nu}\right]\mathrm{d}\nu_i. \tag{31}$$

Of the elements of this equation, $\frac{\partial L_i}{\partial \sigma_\nu}$ is given by equation 25, and $\frac{\partial f(\nu_i,\sigma_\nu)}{\partial \sigma_\nu} = \left(\frac{\nu_i^2}{\sigma_\nu^3}-\frac{1}{\sigma_\nu}\right) f(\nu_i,\sigma_\nu)$.

To obtain the Hessian submatrix $\frac{\partial \mathcal{L}}{\partial \sigma_\epsilon \sigma_\epsilon}$ we differentiate equation 23 with respect to σ_ϵ.

$$\frac{\partial \mathcal{L}}{\partial \sigma_\epsilon \partial \sigma_\epsilon} = \sum_{i=1}^{N} -\frac{1}{L_i^2}\frac{\partial L_i}{\partial \sigma_\epsilon}\int_{-\infty}^{\infty} S_i\, J_i\, f(\nu_i,\sigma_\nu)\,\mathrm{d}\nu_i \\ +\sum_{i=1}^{N}\frac{1}{L_i}\int_{-\infty}^{\infty}\left[\frac{\partial S_i}{\partial \sigma_\epsilon} J_i + S_i\frac{\partial J_i}{\partial \sigma_\epsilon}\right] f(\nu_i,\sigma_\nu)\,\mathrm{d}\nu_i. \tag{32}$$

Of the elements of this equation, $\frac{\partial L_i}{\partial \sigma_\epsilon}$ is given by equation 21 and $\frac{\partial J_i}{\partial \sigma_\epsilon} = S_i\, J_i$. Only $\frac{\partial S_i}{\partial \sigma_\epsilon}$ needs to be derived.

$$\frac{\partial S_i}{\partial \sigma_\epsilon} = \sum_{t=1}^{T_i} \quad w_{it}\left[\frac{1}{\sigma_\epsilon^2}-\frac{3}{\sigma_\epsilon^4}(y_{it}-\beta' x_{it}-\nu_i)^2\right] + (1-w_{it})\psi\left(\frac{-\beta' x_{it}-\nu_i}{\sigma_\epsilon}\right)\left(\frac{\beta' x_{it}+\nu_i}{\sigma_\epsilon^3}\right)\cdot \\ \left[\left(\frac{\beta' x_{it}+\nu_i}{\sigma_\epsilon}\right)^2-\psi\left(\frac{-\beta' x_{it}-\nu_i}{\sigma_\epsilon}\right)\left(\frac{\beta' x_{it}+\nu_i}{\sigma_\epsilon}\right)-2\right] \tag{33}$$

To obtain the Hessian sumbatrix $\frac{\partial \mathcal{L}}{\partial \sigma_\epsilon \sigma_\nu}$ we differentiate equation 23 with respect to σ_ν.

$$\frac{\partial \mathcal{L}}{\partial \sigma_\epsilon \partial \sigma_\nu} = \sum_{i=1}^{N} -\frac{1}{L_i^2}\frac{\partial L_i}{\partial \sigma_\nu}\int_{-\infty}^{\infty} S_i\, J_i\, f(\nu_i,\sigma_\nu)\,\mathrm{d}\nu_i + \sum_{i=1}^{N}\frac{1}{L_i}\int_{-\infty}^{\infty} S_i\, J_i \frac{\partial f(\nu_i,\sigma_\nu)}{\partial \sigma_\nu}\,\mathrm{d}\nu_i. \tag{34}$$

Of the elements of this equation, $\frac{\partial L_i}{\partial \sigma_\nu}$ is given by equation 24 and $\frac{\partial f(\nu_i,\sigma_\nu)}{\partial \sigma_\nu} = \left(\frac{\nu_i^2}{\sigma_\nu^3}-\frac{1}{\sigma_\nu}\right) f(\nu_i,\sigma_\nu)$.

The last submatrix of the Hessian is $\frac{\partial \mathcal{L}}{\partial \sigma_\nu \sigma_\nu}$. To obtain this, we differentiate equation 25 with respect to σ_ν.

$$\frac{\partial \mathcal{L}}{\partial \sigma_\nu \partial \sigma_\nu} = \sum_{i=1}^{N}\frac{-1}{L_i^2}\frac{\partial L_i}{\partial \sigma_\nu}\int_{-\infty}^{\infty}\left(\frac{\nu_i^2}{\sigma_\nu^3}-\frac{1}{\sigma_\nu}\right) J_i\, f(\nu_i,\sigma_\nu)\,\mathrm{d}\nu_i + \\ \sum_{i=1}^{N}\frac{1}{L_i}\int_{-\infty}^{\infty}\left(\frac{1}{\sigma_\nu^2}-3\frac{\nu_i^2}{\sigma_\nu^4}\right) J_i\, f(\nu_i,\sigma_\nu) + \left(\frac{\nu_i^2}{\sigma_\nu^3}-\frac{1}{\sigma_\nu}\right) J_i\,\frac{\partial f(\nu_i,\sigma_\nu)}{\partial \sigma \nu}\,\mathrm{d}\nu_i. \tag{35}$$

As above, $\frac{\partial L_i}{\partial \sigma_\nu}$ is given by equation 24 and $\frac{\partial f(\nu_i,\sigma_\nu)}{\partial \sigma_\nu} = \left(\frac{\nu_i^2}{\sigma_\nu^3} - \frac{1}{\sigma_\nu}\right) f(\nu_i, \sigma_\nu)$.

Because the expressions for the moment conditions and Hessian matrix elements contain complicated integrals, they must be evaluated numerically instead of analytically in order to generate estimates of standard errors for a particular case at hand.

5 Conclusion

The advent of large panel data sets with both substantial time and cross-sectional dimensions introduces the need for estimators that are consistent in the presence of autocorrelation. This paper considers the estimation of the covariance matrix for a Tobit model with errors that are correlated over time. The proposed solution is an application of the Newey-West GMM estimator and suggests a general resolution to the problem of autocorrelated errors for estimators that may be characterized as GMM in nature.

References

[1] Estrella, A. and A. Rodrigues (1998) "Consistent Covariance Matrix Estimation in Probit Models with Autocorrelated Errors," mimeo, Federal Reserve Bank of New York.

[2] Newey, W. K. and D. L. McFadden (1994) "Large sample estimation and hypothesis testing." In *Handbook of Econometrics, Volume IV*, Vol. R. F. Engle et. al., Ed. New York: Elsevier.

[3] Newey, W. K. and K. D. West (1987) "A Simple, Positive Semi-definite, Heteroskedasticity and Autocorrelation Consistent Covariance Matrix," *Econometrica*, 55(3): 703-708.

[4] Robinson, P. (1982) "On the Asymptotic Properties of Estimators of Models Containing Limited Dependent Variables," *Econometrica*, 50(1): 27-41.

[5] Wooldrige, J. M. (2002) *Econometric Analysis of Panel Data*, Cambridge, Massachusetts: The MIT Press.