Two-Step Estimation of a Censored System of Equations
Author(s): J. Scott Shonkwiler and Steven T. Yen
Source: *American Journal of Agricultural Economics*, Vol. 81, No. 4 (Nov., 1999), pp. 972-982
Published by: Oxford University Press on behalf of the Agricultural & Applied Economics Association
Stable URL: http://www.jstor.org/stable/1244339
Accessed: 19-01-2018 15:23 UTC

# Two-Step Estimation of a Censored System of Equations

J. Scott Shonkwiler and Steven T. Yen

A consistent two-step estimation procedure is proposed for a system of equations with limited dependent variables. Monte Carlo simulation results suggest the procedure outperforms an existing two-step method.

*Key words*: Heien and Wessells procedure, limited dependent variables, Monte Carlo simulation.

The increasing availability and use of microdata have made limited dependent variable models indispensable tools in microeconometric modeling. In single-equation applications with a limited dependent variable, maximum-likelihood (ML) estimation of the Tobit model (Tobin) and its parametric generalizations (e.g., Blundell and Meghir, Cragg) is common and straightforward. Whereas theoretical literature exists for systems of equations with limited dependent variables (Amemiya 1974; Lee and Pitt 1986, 1987; Wales and Woodland 1983), direct ML estimation of these models remains difficult when censoring occurs in multiple equations because of the need to evaluate multiple integrals in the likelihood function.

Heien and Wessells (henceforth HW) proposed a two-step estimation procedure for a system of equations (demand system) with limited dependent variables. In the HW procedure, each equation in the system is augmented by a selectivity regressor derived from probit estimates in an earlier step, and the system of equations is estimated with seemingly unrelated regression (SUR) in the second step. Since its inception, this procedure has become increasingly popular in applied demand analysis and the empirical literature has continued to grow.

Although not explicitly stated, the HW procedure is built upon a set of equations which deviate from the unconditional mean expressions for the conventional censored dependent variable specification. In this article, we propose a consistent two-step (henceforth CTS) estimation procedure for systems of equations with limited dependent variables. We conduct Monte Carlo simulations to investigate and compare the performance of the CTS and the HW estimators.

## Two-Step Estimation of a Censored System

We consider the following system of equations with limited dependent variables:

$$(1)\qquad y_{it}^* = f(\mathbf{x}_{it}, \boldsymbol{\beta}_i) + \epsilon_{it}, \qquad d_{it}^* = \mathbf{z}_{it}'\boldsymbol{\alpha}_i + \upsilon_{it},$$

$$d_{it} = \begin{cases} 1 & \text{if } d_{it}^* > 0 \\ 0 & \text{if } d_{it}^* \le 0 \end{cases} \qquad y_{it} = d_{it}y_{it}^*,$$

$$(i = 1, 2, \ldots, m;\ t = 1, 2, \ldots, T)$$

where, for the $i$th equation and $t$th observation, $y_{it}$ and $d_{it}$ are the observed dependent variables, $y_{it}^*$ and $d_{it}^*$ are corresponding latent variables, $\mathbf{x}_{it}$ and $\mathbf{z}_{it}$ are vectors of exogenous variables, $\boldsymbol{\beta}_i$ and $\boldsymbol{\alpha}_i$ are conformable vectors of parameters, and $\epsilon_{it}$ and $\upsilon_{it}$ are random errors.[1] This model is a generalization of Amemiya's (1974) censored system in that the deterministic component $f(\mathbf{x}_{it}, \boldsymbol{\beta}_i)$ can be nonlinear in $\boldsymbol{\beta}_i$ and censoring of each dependent variable is governed by a separate stochastic process; it is a multivariate generalization of Amemiya's (1985) "type 2 Tobit" model, also considered by Cragg and by Heckman.

The authors are, respectively, professor and associate professor of applied economics and statistics at the University of Nevada, Reno.

Helpful comments from two anonymous reviewers are gratefully acknowledged. Research reported herein was supported by the Nevada Agricultural Experiment Station. Senior authorship is shared.

[1] The binary latent equation can be generalized to $d_{it}^* = g(\mathbf{z}_{it}, \boldsymbol{\alpha}_i) + \upsilon_{it}$ but use of the linear form is typical in empirical applications.

*Amer. J. Agr. Econ.* 81 (November 1999): 972–982

The system of equations in (1) implies that for the $i$th equation the single dependent variable $y_{it}$ is observed with nonnegative values. If a nonnegligible proportion of its values are identically zero then it likely cannot be properly represented with a continuous distribution. The selection mechanism represented by the observed dichotomous variable $d_{it}$ is an artifact constructed from knowledge of the censored (zero) observations in $y_{it}$. If, for example, $f(\mathbf{x}_{it}, \boldsymbol{\beta}_i) = \sigma_i \mathbf{z}_{it}'\boldsymbol{\alpha}_i$, where it is assumed $\text{var}(\epsilon_{it}) = \sigma_i$ and $\text{var}(\upsilon_{it}) = 1 \ \forall \ i$, then the $i$th equation follows the Tobit model. However, the selection mechanism may be of a more general form and thus the specification in equation (1) has the attractive property that the censoring mechanism need not vary proportionately with the systematic part of $y_{it}$ when $y_{it}$ is greater than zero.

Direct ML estimation of equation (1) is difficult when error terms are allowed to be contemporaneously correlated, as the likelihood function generally involves multiple integrals. Alternatively, the selection mechanisms can be estimated easily and consistently using $m$ individual ML probit estimators when the $\upsilon_{it}$ are normally distributed. This has led investigators to consider two-step estimators which apply the probit estimator to each selection mechanism and then apply method-of-moments estimators to the observed $y_{it}$ while augmenting its regressors with a correction factor obtained from the first step. For example the HW procedure proceeds as follows: ($i$) obtain ML probit estimates $\hat{\boldsymbol{\alpha}}_i$ for each equation $i$ based on the binary outcomes $d_{it} = 1$ and $d_{it} = 0$; ($ii$) estimate with SUR the system with each equation augmented by an "inverse Mills ratio," defined as $\phi(k_{it}\mathbf{z}_{it}'\hat{\boldsymbol{\alpha}}_i)/\Phi(k_{it}\mathbf{z}_{it}'\hat{\boldsymbol{\alpha}}_i)$ where $k_{it} = 2d_{it} - 1$ and $\phi(\cdot)$ and $\Phi(\cdot)$ are univariate standard normal probability density function and cumulative distribution function, respectively. This two-step procedure, first suggested by HW, has been used extensively in the empirical literature; the applications include, among others, Abdelmagid, Wohlgenant, and Safley; Alderman and Sahn; Gao and Spreen; Gao, Wailes, and Cramer; Han and Wahl; Heien and Durham; Nayga 1995, 1996, 1998; Park et al.; Salvanes and DeVoretz; Wang et al.; and Wellman. In addition, Byrne, Capps, and Saha, and Manrique and Jensen apply the procedure to a single equation with a limited dependent variable.

It is interesting to note that the HW procedure implies $E(y_{it}|\mathbf{x}_{it}, \mathbf{z}_{it}; \upsilon_{it} > -\mathbf{z}_{it}'\boldsymbol{\alpha}_i) = f(\mathbf{x}_{it}, \boldsymbol{\beta}_i) + \delta_i[\phi(\mathbf{z}_{it}'\boldsymbol{\alpha}_i)/\Phi(\mathbf{z}_{it}'\boldsymbol{\alpha}_i)]$ and $E(y_{it}|\mathbf{x}_{it}, \mathbf{z}_{it}; \upsilon_{it} \le -\mathbf{z}_{it}'\boldsymbol{\alpha}_i) = f(\mathbf{x}_{it}, \boldsymbol{\beta}_i) + \delta_i\{\phi(\mathbf{z}_{it}'\boldsymbol{\alpha}_i)/[1 - \Phi(\mathbf{z}_{it}'\boldsymbol{\alpha}_i)]\}$, where the scalar $\delta_i$ is the unknown coefficient of the "correction factor" of the $i$th equation in the second step. These in turn imply the unconditional expectation

$$(2) \qquad E(y_{it}|\mathbf{x}_{it}, \mathbf{z}_{it}) = f(\mathbf{x}_{it}, \boldsymbol{\beta}_i) + 2\delta_i\phi(\mathbf{z}_{it}'\boldsymbol{\alpha}_i).$$

The attraction of the HW procedure stems from the fact that the selection mechanism does not interact with the conditional mean of $y_{it}^*$, that is, $f(\mathbf{x}_{it}, \boldsymbol{\beta}_i)$.[2] However, there is an internal inconsistency in the HW model that can be easily demonstrated by considering what expression (2) implies as $\mathbf{z}_{it}'\boldsymbol{\alpha}_i \to -\infty$, namely, that the unconditional expectation of $y_{it}$ is $f(\mathbf{x}_{it}, \boldsymbol{\beta}_i)$. However the system in equation (1) suggests that as $\mathbf{z}_{it}'\boldsymbol{\alpha}_i \to -\infty$ then $y_{it} \to 0$, as one would expect.

In contrast to two-step procedures based only on nonlimit observations (e.g., Heckman), estimation of a censored system requires a procedure that uses the whole sample since each dependent variable may have a different pattern of censoring (in terms of limit and nonlimit outcomes). To motivate an alternative two-step procedure based on all observations, we draw upon results of Lee (for the strictly Tobit case) and Wales and Woodland (1980) in the single-equation context. Assume for each $i$ the error terms $[\epsilon_{it}, \upsilon_{it}]'$ are distributed as bivariate normal with $\text{cov}(\epsilon_{it}, \upsilon_{it}) = \delta_i$. Then, the conditional mean of $y_{it}$ is (Wales and Woodland 1980)

$$(3) \qquad E(y_{it}|\mathbf{x}_{it}, \mathbf{z}_{it}; \upsilon_{it} > -\mathbf{z}_{it}'\boldsymbol{\alpha}_i) = f(\mathbf{x}_{it}, \boldsymbol{\beta}_i) + \delta_i\frac{\phi(\mathbf{z}_{it}'\boldsymbol{\alpha}_i)}{\Phi(\mathbf{z}_{it}'\boldsymbol{\alpha}_i)}.$$

Because $E(y_{it}|\mathbf{x}_{it}, \mathbf{z}_{it}; \upsilon_{it} \le -\mathbf{z}_{it}'\boldsymbol{\alpha}_i) = 0$, the unconditional mean of $y_i$ is

$$(4) \qquad E(y_{it}|\mathbf{x}_{it}, \mathbf{z}_{it}) = \Phi(\mathbf{z}_{it}'\boldsymbol{\alpha}_i)f(\mathbf{x}_{it}, \boldsymbol{\beta}_i) + \delta_i\phi(\mathbf{z}_{it}'\boldsymbol{\alpha}_i)$$

which clearly differs from equation (2); see also Chung and Goldberger, and Greene.[3]

[2] Note that the presence of the additive term representing the selection mechanism does not permit invariance to equation dropped in a share equation system because each equation will not now have identical regressors.

[3] As one reviewer suggests, errors in the HW procedure may have been caused by a confusion between the observed and latent variables, since it is true that $E(y_{it}^*|\mathbf{x}_{it}, \mathbf{z}_{it}; \upsilon_{it} \le -\mathbf{z}_{it}'\boldsymbol{\alpha}_i) = f(\mathbf{x}_{it}, \boldsymbol{\beta}_i) - \delta_i\phi(\mathbf{z}_{it}'\boldsymbol{\alpha}_i)/[1 - \Phi(\mathbf{z}_{it}'\boldsymbol{\alpha}_i)]$, however, $E(y_{it}|\mathbf{x}_{it}, \mathbf{z}_{it}; \upsilon_{it} \le -\mathbf{z}_{it}'\boldsymbol{\alpha}_i) = 0$.

Based on equation (4) for each $i$, the system of equations (1) can be rewritten as

$$(5)\quad y_{it} = \Phi(\mathbf{z}_{it}'\boldsymbol{\alpha}_i)f(\mathbf{x}_{it}, \boldsymbol{\beta}_i) + \delta_i\phi(\mathbf{z}_{it}'\boldsymbol{\alpha}_i) + \xi_{it} \qquad (i = 1, 2, \ldots, m; t = 1, 2, \ldots, T)$$

where $\xi_{it} = y_{it} - \mathrm{E}(y_{it} | \mathbf{x}_{it}, \mathbf{z}_{it})$. The system (5) can be estimated by a two-step procedure using all observations: ($i$) obtain ML probit estimates $\hat{\boldsymbol{\alpha}}_i$ of $\boldsymbol{\alpha}_i$ using the binary outcome $d_{it} = 1$ and $d_{it} = 0$ for each $i$;[4] ($ii$) calculate $\Phi(\mathbf{z}_{it}'\hat{\boldsymbol{\alpha}}_i)$ and $\phi(\mathbf{z}_{it}'\hat{\boldsymbol{\alpha}}_i)$ and estimate $\boldsymbol{\beta}_1, \boldsymbol{\beta}_2, \ldots, \boldsymbol{\beta}_m, \delta_1, \delta_2, \ldots, \delta_m$ in the system

$$(6)\quad y_{it} = \Phi(\mathbf{z}_{it}'\hat{\boldsymbol{\alpha}}_i)f(\mathbf{x}_{it}, \boldsymbol{\beta}_i) + \delta_i\phi(\mathbf{z}_{it}'\hat{\boldsymbol{\alpha}}_i) + \xi_{it} \qquad (i = 1, 2, \ldots, m; t = 1, 2, \ldots, T)$$

by ML or SUR procedure, where

$$(7)\quad \xi_{it} = \epsilon_{it} + [\Phi(\mathbf{z}_{it}'\boldsymbol{\alpha}_i) - \Phi(\mathbf{z}_{it}'\hat{\boldsymbol{\alpha}}_i)]f(\mathbf{x}_{it}, \boldsymbol{\beta}_i) + \delta_i[\phi(\mathbf{z}_{it}'\boldsymbol{\alpha}_i) - \phi(\mathbf{z}_{it}'\hat{\boldsymbol{\alpha}}_i)]$$

with

$$(8)\quad \mathrm{E}(\xi_{it}) = 0$$

and

$$(9)\quad \begin{aligned}\mathrm{var}(\xi_{it}) = {} & \sigma_i^2\Phi(\mathbf{z}_{it}'\boldsymbol{\alpha}_i) + [1 - \Phi(\mathbf{z}_{it}'\boldsymbol{\alpha}_i)] \\ & \times \{[f(\mathbf{x}_{it}, \boldsymbol{\beta}_i)]^2\Phi(\mathbf{z}_{it}'\boldsymbol{\alpha}_i) \\ & + 2f(\mathbf{x}_{it}, \boldsymbol{\beta}_i)\delta_i\phi(\mathbf{z}_{it}'\boldsymbol{\alpha}_i)\} \\ & - \delta_i^2\{\mathbf{z}_{it}'\boldsymbol{\alpha}_i\phi(\mathbf{z}_{it}'\boldsymbol{\alpha}_i) \\ & + [\phi(\mathbf{z}_{it}'\boldsymbol{\alpha}_i)]^2\}.\end{aligned}$$

Because the ML probit estimators $\hat{\boldsymbol{\alpha}}_i$ are consistent, applying ML or SUR estimation to equation (6) produces consistent estimates in the second step.[5] However, because expression (9) suggests the error terms $\xi_{it}$ in equation (6) are heteroskedastic, the second-step ML or SUR estimator obtained by the usual procedure is inefficient. Efficiency could be achieved by using a weighted system estimator to account for the specific type of heteroskedasticity introduced by equation (9). Another problem caused by the use of the estimated $\hat{\boldsymbol{\alpha}}_i$ in equation (6) is that the covariance matrix of the second-step estimator is incorrect. This covariance matrix can be adjusted by the procedure of Murphy and Topel. Finally, to gain efficiency in estimation and avoid cumbersome correction of covariance matrix using the Murphy-Topel approach, it is possible to combine the first-step (probit) and second-step (system) estimations by jointly estimating each block using the ML procedure within one optimization algorithm.

In light of the large number of studies which employ the inconsistent HW estimator we feel that it is important to determine whether this inconsistency manifests itself in some predictable manner so that results from the previous studies may have some use. Further, we wish to analyze the properties of the CTS estimator in terms of its sampling variability. To accomplish these objectives, we develop a Monte Carlo experiment.

## Monte Carlo Simulations and Comparison with Alternative Estimator

We generate random samples and conduct Monte Carlo simulations to investigate and compare the performance of the CTS and HW estimators.[6] The simulations for a three-equation system proceed as follows:

1. True parameter values are assigned such that

$$\boldsymbol{\beta} = [\boldsymbol{\beta}_1 | \boldsymbol{\beta}_2 | \boldsymbol{\beta}_3] = \begin{bmatrix} 2 & 2 & 2 \\ -0.5 & -0.25 & -0.25 \\ 0.5 & 0.5 & 0.5 \\ 2 & 2 & -2 \end{bmatrix},$$

$$\boldsymbol{\alpha}_1 = \boldsymbol{\alpha}_2 = \boldsymbol{\alpha}_3 = \begin{bmatrix} \alpha^* \\ 0.1 \\ -1 \end{bmatrix}$$

where $\alpha^*$ is allowed to vary with targeted percentages of censoring in the experiment. In particular, $\alpha^* = -1.186$ for 5% censoring, $\alpha^* = -2.79$ for 25% censoring, $\alpha^* = -4.12$

[4] Estimation of the separate probit models implies the restriction $\mathrm{E}(\upsilon_{it}\upsilon_{kt}) = 0$ for $i \neq k$, without which the multivariate probit model would have to be estimated. With some loss in efficiency (relative to multivariate probit) these separate probit estimates are nevertheless consistent.

[5] Estimation of a large system with this procedure should not pose any more computational burden than the Almost Ideal Demand System employed by HW or other nonlinear demand systems.

[6] The programs for data generation and estimation and data generated are available upon request.

for 50% censoring, and $\alpha^* = -5.446$ for 75% censoring.[7]

2. Generate data $x_t$ and $z_t$ for observation $t$ as follows:

$$x_{1t} = z_{1t} = 1, \qquad x_{2t} = 10u_{2t} + 2,$$
$$x_{3t} = z_{2t} = 50u_{3t} + 20,$$
$$x_{4t} = z_{3t} = 1 \quad \text{if } u_{4t} < 0.3;$$
$$= 0 \quad \text{otherwise,}$$

where $u_{2t}$, $u_{3t}$, $u_{4t}$ are random drawings from the uniform distribution with interval [0, 1], mean 0.5, and standard deviation $1/\sqrt{12}$. One thousand draws are used.

3. For observation $t$ a random vector $[\boldsymbol{\epsilon}_t', \boldsymbol{\upsilon}_t']'$ is drawn from the multivariate normal distribution with zero mean and covariance matrix

$$\begin{bmatrix} \boldsymbol{\Sigma}_{\epsilon\epsilon} & \boldsymbol{\Sigma}_{\epsilon\upsilon} \\ \boldsymbol{\Sigma}_{\epsilon\upsilon} & \mathbf{I} \end{bmatrix}$$

where $\mathbf{I}$ is a $3 \times 3$ identity matrix, and

$$\boldsymbol{\Sigma}_{\epsilon\epsilon} = \begin{bmatrix} 4.0 & 1.0 & -1.0 \\ 1.0 & 3.0 & -0.5 \\ -1.0 & -0.5 & 5.0 \end{bmatrix},$$

$$\boldsymbol{\Sigma}_{\epsilon\upsilon} = \begin{bmatrix} 0.5 & 0 & 0 \\ 0 & -0.5 & 0 \\ 0 & 0 & 1.0 \end{bmatrix}.$$

4. Values for the dependent variables are determined using the censoring rule (1). In particular, for $i = 1, 2, 3$,

$$d_{it} = \begin{cases} 1 & \text{if } \mathbf{z}_{it}'\boldsymbol{\alpha}_i + \upsilon_{it} > 0 \\ 0 & \text{otherwise,} \end{cases}$$

$$y_{it} = d_{it}(\mathbf{x}_{it}\boldsymbol{\beta}_i + \epsilon_{it}).$$

5. Repeat steps 3 and 4 $T$ times to obtain a sample of size $T$. We consider $T = 1{,}000$ and $T = 4{,}000$ with different proportions of censored values for the dependent variables. For $T = 4{,}000$, the fixed $x_t$ and $z_t$ are replicated four times.

6. Estimate the system (6) with the generated sample and obtain SUR estimates $\hat{\boldsymbol{\beta}}_i$ and $\hat{\delta}_i$ with the CTS and the HW procedures described above. The cross-equation restrictions $\beta_{13} = \beta_{23} = \beta_{33}$ are imposed to consider the effects of such information on the distribution of the parameter estimates.[8] The first equation is also estimated by itself and compared to the single-equation population following similar procedures; see equation (6) with $i = 1$ for single-equation CTS procedure.

7. Repeat steps 3 through 6 200 times and calculate the simulation-based means and standard deviations of $\hat{\boldsymbol{\beta}}_i$ and $\hat{\delta}_i$ ($\hat{\boldsymbol{\beta}}$ and $\hat{\delta}$ in single-equation case) from results of all replications.

*Results*

Table 1 presents the single-equation simulation results for both the CTS and HW procedures, with repeated samples of 4,000 observations and under different targeted censoring proportions (5%, 25%, 50%, and 75%). The CTS procedure produces estimates fairly close to the true parameters, with all the 95% confidence intervals containing the true parameters. In contrast, the HW procedure produces parameter estimates that have relatively small standard errors but differ notably from the true parameter values. For instance, in the case of 5% censoring, two of the four parameter estimates (excluding the coefficient for the "inverse Mills ratio") deviate so much from the true parameters ($\beta_1$ and $\beta_3$) that the 95% confidence intervals do not contain these parameters. The HW estimator performs particularly poorly as the censoring proportion increases to 25%, 50%, and 75% censoring, in which cases none of the 95% confidence intervals for $\beta$'s contain the true parameters.

The simulation results for a system of equations with 5%, 25%, 50%, and a mixed 75%-50%-25% censoring (for the three equations, respectively) are presented in tables 2, 3, 4, and 5, respectively. Overall, with 5% censoring in all equations (table 2) the CTS procedure performs fairly well, with most parameter estimates significant at the 5% level and with all 95% confidence intervals containing the true parameters. This is also true as the censoring proportion increases to 25% (table 3), 50% (table 4), and mixed 25%-50%-75%

[7] All resulting censoring proportions deviate slightly from targeted proportions as a result of random drawings and are rounded. The use of different $\alpha^*$ values amounts to varying the censoring threshold for the binary mechanism in equation (1). That is, $d_{it} = 1$ if $d_{it}^* > c$; $d_{it} = 0$ if $d_{it}^* \leq c$, where the censoring threshold $c$ is allowed to vary.

[8] Cross-equation restrictions are common in demand systems and highlight the importance of system estimation.

**Table 1. Monte Carlo Simulation Results of Alternative Two-Step SUR Estimates for a Single Equation with Limited Dependent Variable: $n$ = 4,000, 200 Replications, Varied Targeted Censoring Proportions**

| | | | Consistent Two-Step Procedure | | | | Heien-Wessells Two-Step Procedure | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | | | | | 95% Confidence Interval | | | | 95% Confidence Interval | |
| Targeted Censoring % | Parameters[c] | True Value | SUR Estimate | Std. Error | Lower Limit | Upper Limit | SUR Estimate | Std. Error | Lower Limit | Upper Limit |
| 5 | $\beta_1$ | 2.00 | 2.001[a] | 0.216 | 1.537 | 2.421 | 4.242[a,b] | 0.221 | 3.754 | 4.627 |
| | $\beta_2$ | −0.50 | −0.501[a] | 0.018 | −0.543 | −0.470 | −0.474[a] | 0.015 | −0.505 | −0.444 |
| | $\beta_3$ | 0.50 | 0.500[a] | 0.004 | 0.493 | 0.507 | 0.456[a,b] | 0.004 | 0.448 | 0.464 |
| | $\beta_4$ | 2.00 | 2.000[a] | 0.089 | 1.833 | 2.171 | 2.120[a] | 0.095 | 1.921 | 2.296 |
| | $\delta$ | 0.50 | 0.519 | 0.433 | −0.367 | 1.333 | −8.307[a] | 0.178 | −8.642 | −7.954 |
| 25 | $\beta_1$ | 2.00 | 1.934[a] | 0.586 | 0.622 | 3.052 | −4.599[a,b] | 0.725 | −6.042 | −3.429 |
| | $\beta_2$ | −0.50 | −0.496[a] | 0.034 | −0.558 | −0.434 | −0.339[a,b] | 0.031 | −0.400 | −0.267 |
| | $\beta_3$ | 0.50 | 0.501[a] | 0.009 | 0.483 | 0.519 | 0.614[a,b] | 0.012 | 0.592 | 0.6390 |
| | $\beta_4$ | 2.00 | 1.987[a] | 0.166 | 1.688 | 2.317 | −1.488[a,b] | 0.215 | −1.877 | −1.050 |
| | $\delta$ | 0.50 | 0.525 | 0.487 | −0.467 | 1.579 | −5.699[a] | 0.366 | −6.356 | −4.855 |
| 50 | $\beta_1$ | 2.00 | 2.001 | 2.714 | −3.517 | 7.320 | −16.382[a,b] | 0.429 | −17.230 | −15.640 |
| | $\beta_2$ | −0.50 | −0.499[a] | 0.069 | −0.624 | −0.361 | −0.256[a,b] | 0.040 | −0.332 | −0.173 |
| | $\beta_3$ | 0.50 | 0.500[a] | 0.043 | 0.417 | 0.582 | 0.774[a,b] | 0.006 | 0.763 | 0.787 |
| | $\beta_4$ | 2.00 | 2.023[a] | 0.568 | 0.842 | 3.104 | −3.754[a,b] | 0.276 | −4.293 | −3.234 |
| | $\delta$ | 0.50 | 0.543 | 1.319 | −2.344 | 3.150 | −4.576[a] | 0.191 | −4.926 | −4.167 |
| 75 | $\beta_1$ | 2.00 | −3.548 | 26.973 | −59.009 | 44.715 | −14.880[a,b] | 0.619 | −16.077 | −13.734 |
| | $\beta_2$ | −0.50 | −0.490[a] | 0.181 | −0.822 | −0.099 | −0.191[a,b] | 0.049 | −0.294 | −0.097 |
| | $\beta_3$ | 0.50 | 0.578 | 0.389 | −0.107 | 1.364 | 0.554[a,b] | 0.020 | 0.511 | 0.594 |
| | $\beta_4$ | 2.00 | 1.185 | 3.900 | −7.087 | 7.985 | −3.872[a,b] | 0.302 | −4.422 | −3.257 |
| | $\delta$ | 0.50 | 2.053 | 6.916 | −10.795 | 15.754 | 0.353 | 0.568 | −0.783 | 1.572 |

[a] Significant at the 5% level.
[b] Ninety-five percent confidence interval does not include the true parameter.
[c] The Heien-Wessells estimates for $\delta$ are not comparable to the true parameter and the consistent estimates.

**Table 2. Monte Carlo Simulation Results of Alternative Two-Step SUR Estimates for System of Equations with Limited Dependent Variables: $n$ = 4,000, 200 Replications, 5% Targeted Censoring**

| | | | Consistent Two-Step Procedure | | | | Heien-Wessells Two-Step Procedure | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| Targeted Censoring | | | | | 95% Confidence Interval | | | | 95% Confidence Interval | |
| % | Parameter[c] | True Value | SUR Estimate | Std. Error | Lower Limit | Upper Limit | SUR Estimate | Std. Error | Lower Limit | Upper Limit |
| | $\beta_{11}$ | 2.00 | 1.987[a] | 0.155 | 1.678 | 2.284 | 4.297[a,b] | 0.167 | 3.880 | 4.618 |
| | $\beta_{12}$ | −0.50 | −0.500[a] | 0.015 | −0.528 | −0.469 | −0.474[a,b] | 0.013 | −0.499 | −0.447 |
| 5 | $\beta_{13}$ | 0.50 | 0.500[a] | 0.002 | 0.496 | 0.504 | 0.455[a,b] | 0.003 | 0.450 | 0.460 |
| | $\beta_{14}$ | 2.00 | 1.984[a] | 0.088 | 1.806 | 2.131 | 2.128[a] | 0.094 | 1.939 | 2.293 |
| | $\delta_1$ | 0.50 | 0.550 | 0.330 | −0.140 | 1.181 | −8.336[a] | 0.154 | −8.642 | −8.059 |
| | $\beta_{21}$ | 2.00 | 1.979[a] | 0.153 | 1.674 | 2.277 | 4,387[a,b] | 0.163 | 4.045 | 4.736 |
| | $\beta_{22}$ | −0.25 | −0.249[a] | 0.017 | −0.280 | −0.220 | −0.235[a] | 0.012 | −0.260 | −0.215 |
| 5 | $\beta_{23}$ | 0.50 | 0.500[a] | 0.002 | 0.496 | 0.504 | 0.455[a,b] | 0.003 | 0.450 | 0.460 |
| | $\beta_{24}$ | 2.00 | 2.003[a] | 0.075 | 1.846 | 2.157 | 2.095[a] | 0.088 | 1.931 | 2.262 |
| | $\delta_2$ | −0.50 | −0.481 | 0.307 | −1.095 | 0.128 | −9.496[a] | 0.133 | −9.809 | −9.260 |
| | $\beta_{31}$ | 2.00 | 1.998[a] | 0.172 | 1.658 | 2.332 | 4.370[a,b] | 0.160 | 4.053 | 4.672 |
| | $\beta_{32}$ | −0.25 | −0.252[a] | 0.018 | −0.285 | −0.217 | −0.241[a] | 0.013 | −0.268 | −0.217 |
| 5 | $\beta_{33}$ | 0.50 | 0.500[a] | 0.002 | 0.496 | 0.504 | 0.455[a,b] | 0.003 | 0.450 | 0.460 |
| | $\beta_{34}$ | −2.00 | −1.998[a] | 0.100 | −2.208 | −1.776 | −1.578[a,b] | 0.100 | −1.769 | −1.394 |
| | $\delta_3$ | 1.00 | 1.043[a] | 0.378 | 0.311 | 1.761 | −8.068[a] | 0.145 | −8.371 | −7.817 |

[a] Significant at the 5% level.
[b] Ninety-five percent confidence interval does not include the true parameter.
[c] The Heien-Wessells estimates for $\delta$ are not comparable to the true parameters and the consistent estimates.

**Table 3. Monte Carlo Simulation Results of Alternative Two-Step SUR Estimates for System of Equations with Limited Dependent Variables: $n$ = 4,000, 200 Replications, 25% Targeted Censoring**

| | | | Consistent Two-Step Procedure | | | | Heien-Wessells Two-Step Procedure | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | | | | | 95% Confidence Interval | | | | 95% Confidence Interval | |
| Targeted Censoring % | Parameter[c] | True Value | SUR Estimate | Std. Error | Lower Limit | Upper Limit | SUR Estimate | Std. Error | Lower Limit | Upper Limit |
| | $\beta_{11}$ | 2.00 | 1.965[a] | 0.455 | 1.177 | 2.855 | −4.711[a,b] | 0.474 | −5.702 | −3.780 |
| | $\beta_{12}$ | −0.50 | −0.496[a] | 0.034 | −0.562 | −0.437 | −0.340[a,b] | 0.028 | −0.397 | −0.288 |
| 25 | $\beta_{13}$ | 0.50 | 0.500[a] | 0.006 | 0.489 | 0.512 | 0.616[a,b] | 0.007 | 0.602 | 0.630 |
| | $\beta_{14}$ | 2.00 | 2.015[a] | 0.154 | 1.750 | 2.319 | −1.471[a,b] | 0.203 | −1.958 | −1.125 |
| | $\delta_1$ | 0.50 | 0.514 | 0.380 | −0.178 | 1.322 | −5.728[a] | 0.281 | −6.282 | −5.249 |
| | $\beta_{21}$ | 2.00 | 2.019[a] | 0.444 | 1.159 | 2.770 | −4.350[a,b] | 0.479 | −5.344 | −3.385 |
| | $\beta_{22}$ | −0.25 | −0.252[a] | 0.033 | −0.305 | −0.191 | −0.156[a,b] | 0.028 | −0.218 | −0.106 |
| 25 | $\beta_{23}$ | 0.50 | 0.500[a] | 0.006 | 0.489 | 0.512 | 0.616[a,b] | 0.007 | 0.602 | 0.630 |
| | $\beta_{24}$ | 2.00 | 1.994[a] | 0.164 | 1.712 | 2.310 | −1.753[a,b] | 0.227 | −2.211 | −1.311 |
| | $\delta_2$ | −0.50 | −0.505 | 0.379 | −1.206 | 0.301 | −6.940[a] | 0.295 | −7.527 | −6.398 |
| | $\beta_{31}$ | 2.00 | 2.006[a] | 0.435 | 1.080 | 2.784 | −4.362[a,b] | 0.449 | −5.255 | −3.513 |
| | $\beta_{32}$ | −0.25 | −0.250[a] | 0.035 | −0.321 | −0.185 | −0.164[a,b] | 0.030 | −0.225 | −0.113 |
| 25 | $\beta_{33}$ | 0.50 | 0.500[a] | 0.006 | 0.489 | 0.512 | 0.616[a,b] | 0.007 | 0.602 | 0.630 |
| | $\beta_{34}$ | −2.00 | −2.022[a] | 0.158 | −2.323 | −1.708 | −4.181[a,b] | 0.207 | −4.591 | −3.771 |
| | $\delta_3$ | 1.00 | 0.993[a] | 0.367 | 0.342 | 1.757 | −6.027[a] | 0.259 | −6.571 | −5.517 |

[a] Significant at the 5% level.
[b] Ninety-five percent confidence interval does not include the true parameter.
[c] The Heien-Wessells estimates for $\delta$ are not comparable to the true parameters and the consistent estimates.

**Table 4. Monte Carlo Simulation Results of Alternative Two-Step SUR Estimates for System of Equations with Limited Dependent Variables: $n$ = 4,000, 200 Replications, 50% Targeted Censoring**

| Targeted Censoring % | Parameter[c] | True Value | Consistent Two-Step Procedure: SUR Estimate | Consistent Two-Step Procedure: Std. Error | Consistent Two-Step Procedure: 95% Confidence Interval: Lower Limit | Consistent Two-Step Procedure: 95% Confidence Interval: Upper Limit | Heien-Wessells Two-Step Procedure: SUR Estimate | Heien-Wessells Two-Step Procedure: Std. Error | Heien-Wessells Two-Step Procedure: 95% Confidence Interval: Lower Limit | Heien-Wessells Two-Step Procedure: 95% Confidence Interval: Upper Limit |
|---|---|---|---|---|---|---|---|---|---|---|
| | $\beta_{11}$ | 2.00 | 2.005 | 1.670 | −1.0850 | 5.086 | −17.204[a,b] | 0.412 | −18.040 | −16.418 |
| | $\beta_{12}$ | −0.50 | −0.498[a] | 0.072 | −0.6249 | −0.350 | −0.253[a,b] | 0.044 | −0.337 | −0.163 |
| 50 | $\beta_{13}$ | 0.50 | 0.499[a] | 0.026 | 0.4480 | 0.547 | 0.792[a,b] | 0.004 | 0.785 | 0.798 |
| | $\beta_{14}$ | 2.00 | 2.019[a] | 0.433 | 1.2521 | 2.876 | −3.757[a,b] | 0.291 | −4.294 | −3.202 |
| | $\delta_1$ | 0.50 | 0.599 | 0.951 | −1.1452 | 2.505 | −4.529[a] | 0.224 | −5.002 | −4.151 |
| | $\beta_{21}$ | 2.00 | 2.085 | 1.696 | −1.3384 | 5.341 | −17.029[a,b] | 0.379 | −17.826 | −16.343 |
| | $\beta_{22}$ | −0.25 | −0.254[a] | 0.071 | −0.3879 | −0.114 | −0.146[a,b] | 0.041 | −0.223 | −0.075 |
| 50 | $\beta_{23}$ | 0.50 | 0.499[a] | 0.026 | 0.4480 | 0.547 | 0.792[a,b] | 0.004 | 0.785 | 0.798 |
| | $\beta_{24}$ | 2.00 | 1.965[a] | 0.478 | 0.9377 | 2.823 | −4.112[a,b] | 0.282 | −4.657 | −3.578 |
| | $\delta_2$ | −0.50 | −0.551 | 0.849 | −2.2968 | 0.757 | −4.929[a] | 0.196 | −5.284 | −4.558 |
| | $\beta_{31}$ | 2.00 | 2.021 | 1.726 | −0.9088 | 5.541 | −16.842[a,b] | 0.379 | −17.585 | −16.143 |
| | $\beta_{32}$ | −0.25 | −0.249[a] | 0.076 | −0.3918 | −0.087 | −0.148[a,b] | 0.046 | −0.234 | −0.060 |
| 50 | $\beta_{33}$ | 0.50 | 0.499[a] | 0.026 | 0.4480 | 0.547 | 0.792[a,b] | 0.004 | 0.785 | 0.798 |
| | $\beta_{34}$ | −2.00 | −1.931[a] | 0.406 | −2.7279 | −1.063 | −5.499[a,b] | 0.286 | −6.123 | −4.977 |
| | $\delta_3$ | 1.00 | 1.018 | 0.934 | −0.7794 | 2.632 | −4.616[a] | 0.220 | −4.983 | −4.152 |

[a] Significant at the 5% level.
[b] Ninety-five percent confidence interval does not include the true parameter.
[c] The Heien-Wessells estimates for $\delta$ are not comparable to the true parameters and the consistent estimates.

**Table 5. Monte Carlo Simulation Results of Alternative Two-Step SUR Estimates for System of Equations with Limited Dependent Variables: $n$ = 4,000, 200 Replications, Varied Targeted Censoring Proportions**

| | | | Consistent Two-Step Procedure | | | | Heien-Wessells Two-Step Procedure | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | | | | | 95% Confidence Interval | | | | 95% Confidence Interval | |
| Targeted Censoring % | Parameter[c] | True Value | SUR Estimate | Std. Error | Lower Limit | Upper Limit | SUR Estimate | Std. Error | Lower Limit | Upper Limit |
| 75 | $\beta_{11}$ | 2.00 | 2.066 | 1.631 | −1.001 | 5.122 | −19.141[a,b] | 0.485 | −20.075 | −18.224 |
| | $\beta_{12}$ | −0.50 | −0.509[a] | 0.190 | −0.925 | −0.168 | −0.172[a,b] | 0.051 | −0.288 | −0.079 |
| | $\beta_{13}$ | 0.50 | 0.500[a] | 0.011 | 0.479 | 0.519 | 0.667[a,b] | 0.007 | 0.652 | 0.679 |
| | $\beta_{14}$ | 2.00 | 1.955 | 1.442 | −0.835 | 4.513 | −4.070[a,b] | 0.325 | −4.666 | −3.492 |
| | $\delta_1$ | 0.50 | 0.516 | 1.338 | −1.710 | 3.337 | −2.021[a] | 0.312 | −2.556 | −1.407 |
| 50 | $\beta_{21}$ | 2.00 | 1.992[a] | 0.850 | 0.401 | 3.500 | −11.258[a,b] | 0.477 | −12.226 | −10.346 |
| | $\beta_{22}$ | −0.25 | −0.249[a] | 0.067 | −0.399 | −0.116 | −0.173[a,b] | 0.037 | −0.249 | −0.102 |
| | $\beta_{23}$ | 0.50 | 0.500[a] | 0.011 | 0.479 | 0.519 | 0.667[a,b] | 0.007 | 0.652 | 0.679 |
| | $\beta_{24}$ | 2.00 | 2.020[a] | 0.405 | 1.261 | 2.842 | −4.206[a,b] | 0.284 | −4.763 | −3.676 |
| | $\delta_2$ | −0.50 | −0.436 | 0.613 | −1.671 | 0.833 | −4.958[a] | 0.208 | −5.389 | −4.620 |
| 25 | $\beta_{31}$ | 2.00 | 1.988[a] | 0.655 | 0.499 | 3.216 | −6.983[a,b] | 0.458 | −7.850 | −6.012 |
| | $\beta_{32}$ | −0.25 | −0.247[a] | 0.034 | −0.308 | −0.181 | −0.153[a,b] | 0.031 | −0.218 | −0.093 |
| | $\beta_{33}$ | 0.50 | 0.500[a] | 0.011 | 0.479 | 0.519 | 0.667[a,b] | 0.007 | 0.652 | 0.679 |
| | $\beta_{34}$ | −2.00 | −2.004[a] | 0.173 | −2.340 | −1.651 | −4.257[a,b] | 0.210 | −4.664 | −3.803 |
| | $\delta_3$ | 1.00 | 0.989[a] | 0.492 | −0.164 | 1.882 | −5.025[a] | 0.314 | −5.699 | −4.498 |

[a] Significant at the 5% level.
[b] Ninety-five percent confidence interval does not include the true parameter.
[c] The Heien-Wessells estimates for $\delta$ are not comparable to the true parameters and the consistent estimates.

(table 5), with some compromise in statistical significance (notably so for δ's). As the censoring proportion increases, the HW procedure produces significant parameter estimates in most cases but performs very poorly in that few 95% confidence intervals contain the true parameters. Drawing on comparisons with results from a preliminary analysis using 1,000 observations, we find that precision of the CTS estimates can be accomplished by increasing the sample size. We also find that imposition of the cross-equation restrictions tightens the parameter estimates (as reflected in smaller standard errors) as these restrictions improve efficiency.

With regard to finding any systematic tendencies to the bias of the HW estimator such as parameter attenuation or inflation, the simulation results are particularly troubling. We see that sets of parameters can have instances of both attenuation and inflation. Further, the occurrences of sign reversals provide evidence that studies which used the HW approach may suggest relationships that are actually opposite those of the underlying data-generating mechanism.

## Concluding Remarks

The issue of zero observations is "one of the most pressing in applied demand analysis" (Deaton, p. 1809) and other microeconometric applications. Two-step estimation procedures for single-equation limited dependent variable models, common in the 1970s and 1980s, will become obsolete as direct ML estimation of these models becomes feasible. In the case of a censored system of equations, although recent development in simulation estimators resolves some of the practical problems in evaluating multiple integrals (e.g., Hajivassiliou, McFadden, and Ruud), direct ML estimation remains complicated for most empirical practitioners, and a consistent two-step estimation procedure can be invaluable.

The HW estimator has been a favorite choice for empirical analysts for nearly a decade. We point out that the HW estimator is inconsistent and propose an alternative estimation procedure for a system of equations with limited dependent variables. Monte Carlo simulation results suggest that our proposed procedure performs well while the HW procedure performs poorly.

Although we consider only a three-equation linear system in the simulation, application of the methodology to the case of multiple and/or nonlinear equations (e.g., "theoretically plausible" demand systems) is equally straightforward. The two-step procedure proposed here resolves a very practical problem in estimating systems of equations with limited dependent variables and will remain a powerful tool in microeconometric analysis.

*[Received August 1998; accepted March 1999.]*

## References

Abdelmagid, B.D., M.K. Wohlgenant, and C.D. Safley. "Demand for Plants Sold in North Carolina Garden Centers." *Agr. and Resour. Econ. Rev.* 25(April 1996):28–37.

Alderman, H., and D.E. Sahn. "Substitution between Goods and Leisure in a Developing Country." *Amer. J. Agr. Econ.* 75(November 1993):875–83.

Amemiya, T. *Advanced Econometrics.* Cambridge: Harvard University Press, 1985.

———. "Multivariate Regression and Simultaneous Equation Models When the Dependent Variables Are Truncated Normal." *Econometrica* 42(November 1974):999–1012.

Blundell, R.W., and C. Meghir. "Bivariate Alternatives to the Univariate Tobit Model." *J. Econometrics* 34(January/February 1987): 179–200.

Byrne, P.J., O. Capps, Jr., and A. Saha. "Analysis of Food-Away-From-Home Expenditure Patterns for U.S. Households, 1982–89." *Amer. J. Agr. Econ.* 78(August 1996):614–27.

Chung, C.F., and A.S. Goldberger. "Proportional Projections in Limited Dependent Variable Models." *Econometrica* 52(March 1984): 531–34.

Cragg, J.G. "Some Statistical Models for Limited Dependent Variables with Applications to the Demand for Durable Goods." *Econometrica* 39(September 1971):829–44.

Deaton, A. "Demand Analysis." *Handbook of Econometrics,* vol. 3. Z. Griliches and M.D. Intriligator, eds., chap. 30, pp. 1769–1839. Amsterdam: North-Holland, 1986.

Gao, X.M., E.J. Wailes, and G.L. Cramer. "A Microeconometric Analysis of Consumer Taste Determination and Taste Change for Beef." *Amer. J. Agr. Econ.* 79(May 1997): 573–82.

Gao, X.M., and T. Spreen. "A Microeconometric Analysis of the U.S. Meat Demand." *Can. J. Agr. Econ.* 42(November 1994):397–412.

Greene, W.H. "On the Asymptotic Bias of the Ordinary Least Squares Estimator of the Tobit Model." *Econometrica* 49(March 1981): 505–13.

Hajivassiliou, V.A., D. McFadden, and P. Ruud. "Simulation of Multivariate Normal Rectangle Probabilities and Their Derivatives: Theoretical and Computational Results." *J. Econometrics* 72(May/June 1996):85–134.

Han, T., and T. Wahl. "China's Rural Household Demand for Fruit and Vegetables." *J. Agr. and Appl. Econ.* 30(July 1998):141–50.

Heckman, J.J. "The Common Structure of Statistical Models of Truncation, Sample Selection, and Limited Dependent Variables and A Simple Estimator for Such Models." *Annals Econ. and Soc. Meas.* 5(Fall 1976):475–92.

Heien, D., and C. Durham. "A Test of the Habit Formation Hypothesis Using Household Data." *Rev. Econ. and Statist.* 73(May 1991):189–99.

Heien, D., and C.R. Wessells. "Demand Systems Estimation with Microdata: A Censored Regression Approach." *J. Bus. and Econ. Statist.* 8(July 1990):365–71.

Lee, L.F. "Estimation of Limited Dependent Variable Models by Two-Stage Methods." PhD dissertation, University of Rochester, 1976.

Lee, L.F., and M.M. Pitt. "Microeconometric Demand Systems with Binding Nonnegativity Constraints: The Dual Approach." *Econometrica* 54(September 1986):1237–42.

———. "Microeconometric Models of Rationing, Imperfect Markets, and Non-negativity Constraint." *J. Econometrics* 36(September/October 1987):89–110.

Manrique, J., and H.H. Jensen. "Spanish Household Demand for Convenience Meat Products." *Agribus.* 13(December 1997):579–86.

Murphy, K.M., and R.H. Topel. "Estimation and Inference in Two-Step Econometric Models." *J. Bus. and Econ. Statist.* 3(October 1985):370–79.

Nayga, R. "Microdata Expenditure Analysis of Disaggregate Meat Products." *Rev. Agr. Econ.* 17(September 1995):275–85.

———. "A Sample Selection Model for Prepared Food Expenditures." *Appl. Econ.* 29(March 1998):345–52.

———. "Wife's Labor Force Participation and Family Expenditures for Prepared Food, Food Prepared at Home, and Food away from Home." *Agr. and Resour. Econ. Rev.* 25(October 1996):179–86.

Park, J.L., R.B. Holcomb, K.C. Raper, and O. Capps, Jr. "A Demand Systems Analysis of Food Commodities by U.S. Households Segmented by Income." *Amer. J. Agr. Econ.* 78(May 1996):290–300.

Salvanes, K.G., and D.J. DeVoretz. "Household Demand for Fish and Meat Products: Separability and Demographic Effects." *Marine Resour. Econ.* 12(Spring 1997):37–55.

Tobin, J. "Estimation of Relationships for Limited Dependent Variables." *Econometrica* 26(January 1958):24–36.

Wales, T.J., and A.D. Woodland. "Estimation of Consumer Demand Systems with Binding Nonnegativity Constraints." *J. Econometrics* 21(April 1983):263–85.

———. "Sample Selectivity and the Estimation of Labor Supply Functions." *Int. Econ. Rev.* 21(June 1980):437–68.

Wang, J., X.M. Gao, E.J. Wailes, and G.L. Cramer. "U.S. Consumer Demand for Alcoholic Beverages: Cross-Section Estimation of Demographic and Economic Effects." *Rev. Agr. Econ.* 18(September 1996):477–88.

Wellman, K.F. "The U.S. Retailed Demand for Fish Products: An Application of the Almost Ideal Demand System." *Appl. Econ.* 24(April 1992):445–57.