UNIVERSITY OF COPENHAGEN

Estimation of Tobit Type Censored Demand Systems

A Comparison of Estimators

Barslund, Mikkel Christoffer

Publication date:
2007

Document version
Publisher's PDF, also known as Version of record

Citation for published version (APA):
Barslund, M. C. (2007). *Estimation of Tobit Type Censored Demand Systems: A Comparison of Estimators*. Department of Economics, University of Copenhagen.

Discussion Papers
Department of Economics
University of Copenhagen

No. 07-16

Estimation of Tobit Type Censored Demand Systems: A Comparison of Estimators

Mikkel Barslund

Studiestræde 6, DK-1455 Copenhagen K., Denmark
Tel. +45 35 32 30 82 - Fax +45 35 32 30 00
http://www.econ.ku.dk

ISSN: 0902-6452 (print) ISSN: 1601-2461 (online)

Estimation of Tobit type censored demand systems: A comparison of estimators *

Mikkel Barslund
Department of Economics, University of Copenhagen
Development Economics Research Group (DERG)

August 2007

Abstract

Recently a number of authors have suggested to estimate censored demand systems as a system of Tobit multivariate equations employing a Quasi Maximum Likelihood (QML) estimator based on bivariate Tobit models. In this paper I study the efficiency of this QML estimator relative to the asymptotically more efficient Simulated ML (SML) estimator in the context of a censored Almost Ideal demand system. Further, a simpler QML estimator based on the sum of univariate Tobit models is introduced. A Monte Carlo simulation comparing the three estimators is performed on three different sample sizes. The QML estimators perform well in the presence of moderate sized error correlation coefficients often found in empirical studies. With absolute larger correlation coefficients, the SML estimator is found to be superior. The paper lends support to the general use of the QML estimators and points towards the use of simple etimators for more general censored systems of equations.

Keywords: Censored demand system, Monte Carlo, Quasi maximum likelihood, Simulated maximum likelihood.

JEL Classifications: D12, C15, C34

*Address for correspondence: Mikkel Barslund, Department of Economics, Studiestraede 6, DK-1455 Copenhagen K, Denmark. Email: Mikkel.Barslund@econ.ku.dk. I would like to thank Katleen Van den Broek, Carol Newman, Heino Bohn Nielsen and Mette Ejrnæs for useful comments and suggestions. The usual disclaimer applies.

1 Introduction

Analysis of individuals and households consumption patterns and their response to relative price changes has a long tradition in economics and goes back at least to Engels seminal work on expenditure shares. Systems of flexible functional forms such as the translog and almost ideal demand systems (Jorgensen et al. 1979, Deaton and Muellbauer 1980) and the advance of fast computers have made estimation of price response coefficients in large demand systems with many goods based on household survey data feasible. Hence, a large literature has grown. However, until recently the problem of censoring of the expenditure shares (i.e. the minimum consumption share is zero) was largely ignored or only addressed in systems with a small number of goods (see Wales & Woodland 1983, Lee & Pitt 1986).

To account for censoring a model which allows for a positive probability of observing zero consumption must be estimated. Thus, whether implicit or explicit, the model should accommodate a market participation decision and a consumption decision. Further, the estimation procedure must be capable of accommodating cross-equation restrictions, making joint estimation of all equations necessary. If errors are normal and assumed to covary between the decisions to consume each good, then - with multiple goods not consumed for some households - the contribution to the likelihood function will require evaluation of multiple integrals over a multivariate normal density function. As an example, consider the case of a five good demand system where a non-neglible number of households only consume two of the five goods, thus equations for the three non-consumed goods are censored. For these households part of the likelihood contribution will be the probability that the three error terms fall within a range consistent with observed censoring of these three goods. Difficulties associated with evaluating multiple integrals over the multivariate normal density function explain why accounting for censoring in applications of large demand system is rare.

One way to account for censoring which has been used in the literature is to model the consumption shares as a multivariate Tobit model (see Yen, Lin and Smallwood 2003), such that implicitly the participation decision and the consumption decision are determined by the same process. In the context of demand system estimation two maximum likelihood based estimators have recently been used to estimate multivariate Tobit systems. Harris and Shonkwiler (1997) proposed a Quasi Maximum Likelihood (QML) estimator based on linking bivariate Tobit models to avoid evaluating high dimensional integrals. More recently Yen, Lin and Smallwood (2003) have used a Simulated Maximum Likelihood (SML) estimator of a similar system.[1] While both estimators are consistent, the SML estimator is asymptotically more efficient, since it uses more sample information than the QML estimator. However, the relative performance of the estimators in applications with empirically relevant sample sizes is unknown.

The contribution of this paper is twofold. First - inspired by the idea of linking bivariate Tobit models - I introduce a simpler QML estimator based on the maximization of the sum of univariate Tobit models over all equations. Although the proposed estimator does not identify the error correlation across equations this is of secondary importance in a demand system context since error correlations are not used to calculate elasticities or other quantities of interest. Second, I compare the three estimators using Monte Carlo simulations in a setup with four simultaneous equations subject to a large degree of censoring. Their performance is assessed in three different sample sizes with respectively 200, 1,000 and 3,000 observations. The sample sizes are chosen to resemble a 'small' sample of households (200 observations), a larger sample (1,000 observations) and a typical (sub)-sample from the World Banks LSMS surveys (3,000 observations).

There are a number of reasons contributing to the relevance of this exercise. First, it is not evident which estimator is preferable for relatively small sample sizes. Second, even if the SML estimator is superior, the cost of implementation and the computational burden associated with simulating the likelihood function might warrant the use of a sufficiently good second best estimator. Third, the SML estimator has difficulties converging from arbitrary starting values and computation time is reduced substantially by using good starting values possibly obtained from less efficient estimators. Further, the type of QML estimators used in this paper can be applied to more general systems of censored systems, i.e. the system suggested by Yen and Lin (2006).

There exist few application specific comparisons of the SML estimator and the bivariate Tobit QML estimator considered here. Yen, Lin and Smallwood (2003) estimate a large demand system

[1] Shonkwiler & Yen (1999) propose a two step estimator where the participation decision is modelled as a univarite probit on each equation (or alternatively, a multivariate probit over all equations). In the second step, the equations determining the expenditure shares are augmented to take account of the censoring and errors are assumed multivariate normal. The estimator is consistent but less efficient relative to the SML estimator due to the two-step nature. It is not considered in the present work, since it is not suitable for estimation of Tobit type models.

with both the SML and the bivariate Tobit QML and conclude that the QML and SML estimator deliver very similar results. In a similar application Yen and Lin (2002) find QML and SML estimates to be close and similar. Clearly, since the true data generating mechanism and parameters are unknown these studies cannot shed light on the relative performance of the estimators in question.

In the following section the model is outlined together with the three estimators. Section 3 describes the Monte Carlo setup, while section 4 presents results.[2] Some brief concluding remarks are offered at the end.

2 Estimation of a multivariate system of Tobit equations

The point of departure is a multivariate generalization of the Tobit model. Denote the dependent variable by y_i, the matrix of explanatory variables by X and the full set of parameters to be estimated by θ, then the system of equations $(i = 1, .., M)$ can be written (suppressing observation indices)

$$y_i = \max\left(f_i\left(X;\theta\right) + \varepsilon_i, 0\right), \quad i = 1, .., M \tag{1}$$

where ε_i is an equation specific error term. Define the vector of errors $\varepsilon = [\varepsilon_1, .., \varepsilon_M]$ and allow parametric estimation by assuming multivariate normal errors with zero mean and covariance Σ. In the context of demand system estimation, y_i is the expenditure share on good i and the functions f_i are of some flexible form. Note that in applications where there is no need for cross equation restrictions (1) can be estimated consistently using a univariate Tobit model equation by equation. However, even in this case efficiency is gained by estimating all equations jointly as a system.

Simulated Maximum Likelihood estimation

To construct the likelihood function for the system given by (1), let a censoring regime z_c be a $1 \times M-$vector with entries equal to zero for the censored equations and one for the non-censored equations. Each observation belongs to a particular censoring regime. Thus, an observation with the first k equations non-censored and the remaining censored would have ones in the first k entries

[2] While the simulations have been done in the context of estimating an almost ideal demand system with five goods (the last good being determined residually as suggested by Pudney (1989), i.e. four goods/equations estimated), this is not emphasized in the discussion of the results.

and zeros for the rest. Call this regime z_c and note that all censoring regimes can be written like this with k equal to the number of non-censored equations and a suitable reorganization of the equations. That is, no generality is lost. To develop the likelihood function for the observations belonging to the censoring regime, z_c, partition the error vector and the covariance matrix such that

$$\begin{aligned} \boldsymbol{\varepsilon} &\equiv [\boldsymbol{\varepsilon}_1, \boldsymbol{\varepsilon}_2] \equiv [\varepsilon_1, \varepsilon_2, .., \varepsilon_k : \varepsilon_{k+1}, \varepsilon_{k+2}, .., \varepsilon_M] \\ \Sigma &\equiv \begin{bmatrix} \Sigma_{11} & \\ \Sigma_{21} & \Sigma_{22} \end{bmatrix} \end{aligned}$$

where Σ_{11} is a $k \times k$ matrix, Σ_{21} is a $(M-k) \times k$ matrix and Σ_{22} is a $(M-k) \times (M-k)$ matrix. Let $g(\boldsymbol{\varepsilon}_1)$ be the joint marginal probability density function (pdf) for the first k errors. The pdf function for the M errors can be written in terms of $g(\boldsymbol{\varepsilon}_1)$ and the joint marginal pdf of the remaining $(M-k)$ error terms conditional on observing $\boldsymbol{\varepsilon}_1$, $h(\boldsymbol{\varepsilon}_2 \mid \boldsymbol{\varepsilon}_1)$.Thus, the joint marginal pdf, $f(\boldsymbol{\varepsilon}_1, \boldsymbol{\varepsilon}_2)$, can be written as

$$f(\boldsymbol{\varepsilon}_1, \boldsymbol{\varepsilon}_2) \equiv g(\boldsymbol{\varepsilon}_1) \cdot h(\boldsymbol{\varepsilon}_2 \mid \boldsymbol{\varepsilon}_1)$$

It can be shown that $h(\boldsymbol{\varepsilon}_2 \mid \boldsymbol{\varepsilon}_1)$ is distributed multivariate normal with mean and covariance matrix given by (Greene 2000)

$$\begin{aligned} \mu_{2.1} &= \Sigma_{21}\Sigma_{11}^{-1}\boldsymbol{\varepsilon}_1 \\ \Sigma_{22.1} &= \Sigma_{22} - \Sigma_{21}\Sigma_{11}^{-1}\Sigma_{21}^{'} \end{aligned}$$

The contribution to the likelihood function for an observation belonging to censoring regime z_c is then given by

$$L^{Z_c} = g(\mathbf{e}_1) \int\limits_{-\infty}^{-f_{k+1}(X;\theta)} \int\limits_{-\infty}^{-f_{k+2}(X;\theta)} .. \int\limits_{-\infty}^{-f_M(X;\theta)} \phi_{(M-k)}\left(u_{k+1}, u_{k+2}, .., u_M\right) \partial u_M .. \partial u_{k+2} \partial u_{k+1} \tag{2}$$

where $g(\mathbf{e}_1)$ is the k-variate normal density with zero mean and covariance matrix Σ_{11} evaluated at $\mathbf{e}_1 = \left[y_1 - f_1\left(X;\theta\right), y_2 - f_2\left(X;\theta\right), .., y_k - f_k\left(X;\theta\right)\right]$. The integration is with respect to the $M-k$-variate normal density with mean and covariance given above. To write the likelihood function for the sample define the indicator function $I_h^{Z_c}$ being one if observation h is in censoring regime z_c and zero otherwise. Since each observation belongs to only one censoring regime the sample likelihood can be written as

$$L = \prod_h \prod_{z_c} [L_h^{z_c}]^{I_h^{z_c}}$$

The set of censoring regimes includes the two special cases where respectively none and all equations are censored. The contributions to the likelihood function are $L^{Z_c} = g(\mathbf{e}_1)$ and $L^{Z_c} = \int\limits_{-\infty}^{-f_1(X;\theta)} \int\limits_{-\infty}^{-f_2(X;\theta)} .. \int\limits_{-\infty}^{-f_M(X;\theta)} \phi_{(M)}(u_1, u_2, .., u_M) \partial u_M .. \partial u_2 \partial u_1$, both M-variate normal density functions having zero mean and covariance Σ.

If for just one observation the number of censored equations exceeds two, a simulation method has to be relied upon to evaluate the integral in (2). I rely on the GHK (Geweke, Hajivassiliou and Keane) simulator to evaluate the integrals[3].

Quasi Maximum Likelihood Estimation

Although implementation of the SML estimator is feasible in most statistical packages (such as Stata and Gauss) it is likely to be computationally intensive. In addition without good starting values obtaining convergence can be difficult. Thus, it is of interest to explore simpler estimators which allow for cross-equation restrictions and do not require simulation techniques. Analogous to the literature on multivariate probit models (Avery et al. 1983) a simple alternative is a QML estimator which maximizes the sum of individual equation Tobit models (QML_{T1}). Formally, the QML_{T1} estimator maximizes

$$\ln QML_{T1} = \sum_h \sum_i \ln L_{T1,i} \tag{3}$$

where the subscript $T1$ indicates that the QML is with respect to the univariate Tobit model. $L_{T1,i}$ is the Tobit likelihood function for the i'th equation and as before h indexes observations. The estimator is based on maximizing the sum of marginal densities of the system in (1) and is therefore consistent (Cameron and Trivedi, 2005). However, because the likelihood function in (3) is mis-specified relative to the true likelihood function for the model in (1), Whites robust standard errors should be used for statistical inference (White, 1982). The estimator incorporates

[3] The methodology behind the GHK simulator is explained elsewhere and is beyond the scope of the present paper (see Börsch-Supan and Hajivassiliou, 1993, and Cappellari and Jenkins, 2006). In practice the GHK simulator has been shown to work well (Greene 2000).

all equations simultaneously in the procedure so cross equation restrictions can be imposed. The entries in the covariance matrix outside the diagonal in the system in (1) are not identified. Hence, the vector of estimated coefficients has fewer elements than the vector of coefficients from the SML estimator. For demand system applications where cross equation correlations are not of particular value this is of minor importance. On the other hand, if the purpose of the QML estimation is to get starting values for the SML estimator cross equation correlations are valuable in their own right.

An extension to the approach in (3) which yields estimates of cross equation correlation coefficients is to estimate a sequence of pair-wise bivariate Tobit models (QML_{T2}). The QML_{T2} estimator maximizes

$$\ln QML_{T2} = \sum_{h} \sum_{i}^{M-1} \sum_{j=i+1}^{M} \ln L_{T2;(i,j)}$$

$T2$ indicates that the sequence of likelihood functions are bivariate Tobits. The coefficient of correlation between the error terms in the i'th and j'th equations is identified from the contribution of $L_{T2;(i,j)}$ to the sample quasi likelihood. Thus, this approach yields as many estimated coefficients as the SML estimator. This last method has recently been used in a number of studies (see Yen, Lin and Smallwood 2003, Barslund 2006, Lin and Yen 2002, Harris and Shonkwiler 1997).

3 Monte Carlo simulations

The relative performance of each estimator is explored along two dimensions. First, a system of four equations is estimated on three different sample sizes to investigate how closely their performance is related to sample size. The sample sizes of 1,000 and 3,000 observations are chosen to resemble empirically relevant samples from typical cross sectional data sets. The third sample with 200 observations is employed to look at performance in a 'small' sample. Second, the effect of the absolute size of the error correlation coefficients is examined. In particular, since the QML_{T1} estimator ignores cross equation error correlations its performance should deteriorate as the absolute size of the correlation coefficients increases. The QML_{T2} estimator identifies the correlation coefficient via the bivariate Tobit formulation, but unlike the SML estimator it does not take into account the complete correlation structure when estimating the pair-wise correlation coefficients.

Overall, the SML estimator should improve relative to the other two estimators when the correlation between equations increases. I compare the estimators for each sample size and for two error correlation structures; namely an empirically relevant correlation matrix ('base' correlations) and a matrix where the base correlations are doubled ('large' correlations). For comparison, and using the 'base' correlation matrix, the system of latent shares is estimated ignoring the issue of censoring and the errors are assumed multivariate normally distributed. Each scenario consists of 500 simulations.

Monte Carlo setup

The system of equations is based on a censored almost ideal demand system. In the context of an empirical application this corresponds to a five good system where the last good is residually determined as suggested by Pudney (1989). Although adding up (expenditure shares sum to one) is accommodated in this way, parameter restrictions designed to facilitate adding up in the latent system of expenditure shares are still imposed. In addition, in order to see how the estimators perform in the presence of cross equation restrictions, slutsky symmetry is imposed on latent shares even if the theoretical justification for this is blurred in censored systems.[4] The latent almost ideal demand system has the form (observation indices suppressed)

$$\begin{aligned} w_i^* &= \alpha_i + \sum_{j=1}^{M+1} \gamma_{ij} \log p_j + \beta_i \log(x/a(p)) \qquad (4) \\ \text{with } \log a(p) &= \alpha_0 + \sum_{j=1}^{M+1} \alpha_j \log p_j + 1/2 \sum_{i=1}^{M+1} \sum_{j=1}^{M+1} \gamma_{ij} \log p_i \log p_j \end{aligned}$$

Where w_i^* is the latent expenditure share on commodity i. α_i, β_i and γ_{ij} are parameters to be estimated. Exogenous variables are prices, p_i and income/expenditure x. As is often done in empirical applications, the unidentified parameter α_0 is set equal to zero (Moschini, 1998). The indices i, j denote commodities, thus $i, j \in \{1, .., M+1\}$. Adding-up, slutsky symmetry and homogeneity of the latent shares are ensured by the parameter restrictions: $\sum_{i=1}^{M+1} \alpha_i = 1$, $\gamma_{ij} = \gamma_{ji}, \forall i, j$ and $\sum_{i=1}^{M+1} \beta_i = 0$.Denoting the full parameter vector by θ the observed shares are

[4] In any case, imposing slutsky symmetry in censored demand systems is standard practice in empirical applications.

given by (equivalent to the system in (1))

$$w_i = \max\left(w_i^*(\theta) + \varepsilon_i, 0\right), \quad i = 1, .., M$$

For each scenario the simulations are done in the following steps:

1) Exogenous variables (logarithmic prices and incomes) are drawn from a standard normal distribution.

2) Errors are drawn from the specified multivariate normal distribution (cf. below).

3) Latent shares are calculated, errors added, and the observed share is determined from the censoring rule.

4) Each estimator is estimated using the observed shares and exogenous variables.

5) Estimates are saved.

Step 2 to 5 are carried out 500 times for each scenario with fixed exogenous variables. Parameters are chosen such that they are within a range often found in empirical applications.

	Alpha	Beta	Gamma:	Eq. 1	Eq. 2	Eq. 3	Eq. 4	Eq. 5
Equation 1	0.3	-0.025		-0.06	-0.03	0.05	0.20	0.02
Equation 2	0.25	0.03		-0.03	-0.01	0.02	0.01	0.01
Equation 3	0.05	-0.01		0.05	0.02	-0.03	-0.02	-0.02
Equation 4	0.1	0.02		0.20	0.01	-0.02	0.01	-0.02
Equation 5	0.3	-0.015		0.02	0.01	-0.02	-0.02	0.01

The values ensure both adding up and slutsky symmetry of the latent shares. The errors are drawn from a multivariate normal distribution with the base error correlation structure given by:

	Standard	Probability	Correlation matrix (base corr. coef.):			
	deviation (σ).	censored (%)	Eq. 1	Eq. 2	Eq. 3	Eq. 4
Equation 1	0.6	30.9	1.00	-0.20	-0.15	-0.08
Equation 2	0.5	30.9	-0.20	1.00	-0.15	-0.07
Equation 3	0.4	45.0	-0.15	-0.15	1.00	-0.10
Equation 4	0.3	36.9	-0.08	-0.07	-0.10	1.00

The probability of an observation being censored is calculated using that for any given observation the expected latent expenditure share w_i^* is equal to α_i since expected logarithmic prices and income are zero (drawn from a standard normal distribution). The base correlation matrix is chosen to resemble the range of values found in empirical studies. The average absolute value over the error correlation coefficients is 0.125 with a maximum absolute value of 0.20. This compares well with the average of 0.083 over absolute correlation coefficients found in Yen, Lin and

Smallwood (2003) with only one coefficient out of 66 being significantly larger than 0.20. Yen, Fang and Su (2004) report slightly larger coefficients. The absolute average is 0.118 and 6 out of 45 coefficients are significantly larger than 0.20 with a maximum of 0.288. Similarly, Barslund (2006) finds an absolute average of 0.116 with 5 out of 55 correlation coefficients being significantly larger than 0.20. The maximum value reported is 0.327. Lastly, Yen and Lin (2002) estimate a three equation system with the largest correlation coefficient not significantly larger than 0.20 and with an average value of 0.136 in absolute terms. Although the absolute size of the correlation coefficients is application specific, the scenarios with 'large' correlations should provide an upper bound for differences in the estimators likely to be found in empirical applications.

A final issue relates to the evaluation of the SML log-likelihood using the GHK simulator. The GHK simulator relies on a specific number, R, of random draws from the unit interval. Because the accuracy of the GHK simulator relies on the size of R, formally, the efficiency of the SML estimator hinges on $\sqrt{N}/R \longrightarrow 0$, where N is the number of observations (see Train 2003). The pitfall to avoid in relation to the Monte Carlo simulation is that the relative performance of the SML versus the QML_{T1} and QML_{T2} is not confounded with poor accuracy of the SML estimator due to an inadequate number of draws when using the GHK simulator. The random draws were generated by Statas `mdraws` command (Cappellari and Jenkins, 2006). In practice, the number of draws were determined following a suggestion by Haan and Uhlendorff (2006). They propose to start by maximizing a simulated log-likelihood function using R equal to $N^{0.55}$ random draws and then increase the number of draws until the maximized log-likelihood function stabilizes on a value. For all three sample sizes Halton sequences with $R = 84$ and antithetic draws were used (Cappellari and Jenkins, 2006). For the sample of 3,000 observations the change in the log-likelihood value at $R = 84$ was below 1/100 of a percentage point. The change in parameter values was on average less than 1/25 of the difference between the SML and the QML_{T2} estimator.[5]

[5] All estimations were done in Stata with 'seeding' of the random generator used for drawing errors so as to facilitate replicability. Files are available from the author.

4 Results

To manage the amount of output the discussion of the results will concentrate on differences in the mean squared error (MSE) between the estimators. The performance of the QML estimators, QML_{T1} and QML_{T2}, is measured relative to the asymptotically more efficient SML estimator. Table 1 shows the results for the base correlation specification with a sample of 200 observations.

Columns numbered 2 through 10 show the percentage deviation of the mean over the 500 simulations from the true value and the MSE for each of the estimated parameters for respectively, the SML, QML_{T1}, QML_{T2}, and the non-censored estimator. The deviation from the mean is reported in order to gorge the biasness in finite samples. As expected - given the degree of censoring - the non-censored estimator shows a large bias for all coefficients (column 8). Turning to the three estimators of primary interest, the most interesting thing coming out of Table 1 is how similar the results are. Looking across the rows it is clear that the differences between the estimators for both measures are small. When one estimator performs particularly well with respect to a point estimate of a coefficient the other two also do well. And similar when coefficients are less precisely estimated. For an illustration look at the estimated standard deviations for the error term of equation 1 (σ_1) and 4 (σ_4), respectively. In terms of their MSE, σ_4 performs well over all three estimators whereas the opposite is true for σ_1. Column 10, 11 and 12 summarize the differences between the coefficient MSEs over the three estimators. Column 10 shows a comparison of the SML and QML_{T2} estimator, where a plus indicates that the SML has the lowest MSE. Similar for column 11 where the SML is compared to the QML_{T1} estimator. Lastly, the QML_{T2} and QML_{T1} estimators are compared in column 12. Thus, for all three columns a plus signifies that the estimator using the most sample information performed better. Reflecting the resemblance of column 2 to 7 none of the estimators perform better than the two others for all coefficients. However, the QML_{T2} seems to have on average slightly lower MSEs than both the SML (minus in column 10) and the QML_{T1} (plus in column 12).

It is of interest to test if the small differences in performance between the estimators are statistically significant. For this purpose I perform two-tailed t-tests of equality of MSEs based on the sample of 500 replications. Significance levels are indicated in the three last columns by one, two or three asterisks equivalent to significance at 10, 5 and 1 percent, respectively. For only one coefficient (β_1) does the SML estimator perform significantly better than the two others, while the QML_{T2} does significantly better than the SML for five coefficients and better than the QML_{T1} for seven coefficients. In sum, for small samples with error correlations of empirical relevant size both the QML_{T1} and QML_{T2} perform very well.

Table 2 is similar to Table 1, but the sample size is increased to 1,000 observations. First, note that for the three estimators of primary interest the deviations of the mean for most estimated coefficients are smaller than in Table 1. This is to be expected from consistent estimators. Contrast that with the biased non-censored estimator, where deviations from the mean are more or less unchanged between Table 1 and 2. Also the MSEs are reduced substantially. Regarding the comparison in the last three columns, the SML estimator has lower MSEs for a majority of coefficients than both the QML_{T1} and QML_{T2} estimators. However, this better performance is not statistically significant (except for one coefficient in the comparison between the SML and QML_{T1} estimators), again reflecting that the coefficient estimates coming from the three estimators are very close for each simulation. The QML_{T2} estimator has lower MSEs for all but two parameters (five are significantly lower) compared with the QML_{T1} estimator. In Table 3 the number of observations is further increased to 3,000 while keeping the same base error correlation structure. Except for the non-censored estimator the effect on the deviation of the mean and the MSEs for all estimators are as expected. The great majority of parameters have means within one percent of the true value and the MSEs have decreased compared to Table 2. However, the SML estimator is now superior to both QML estimators. Not only does it perform better for the great majority of parameters (lower MSEs) as indicated in column 10 and 11 in Table 3, but it is also significantly better for a small number of parameters. The QML_{T2} does a better job than the QML_{T1} estimator (column 12).

The main message from Table 1 to 3 where simulations are done with the base error correlation structure is that it takes a relatively large sample size before the theoretical better performance of the SML estimator shows up. Even then the gains from employing the SML estimator are small. In particular, it is clear that both QML estimators provide very accurate approximations of the SML estimator for the sample sizes examined here, although only the QML_{T2} estimator yields error correlation estimates. To illustrate the last point consider the difference in individual point estimates between the SML and QML_{T1} estimators for the 500 simulations with 3,000 observations. The two estimators have 22 parameters in common since the QML_{T1} estimator does not identify error correlations. For 13 of the 22 parameters the QML_{T1} estimator is never more than 5 percent worse than the SML estimator. For the remaining parameters, more than 85 of the 500 point estimates are not more than 5 percent further from the true value than the SML estimator. The only exception being γ_{22} where only 71 percent lies within this criterion.

Table 4 to 6 are analogous to table 1, 2 and 3, but with the correlation matrix multiplied by two ('large' correlations). The non-censored estimator is not included since the results above showed it to be clearly biased. Although all three tables are presented for completeness, table 4 with 200 observations provides a clear case of how the results differ between the two sets of tables. With the absolute larger correlation coefficients the SML estimator is superior to both the QML_{T2} and the QML_{T1} estimators with very few (five) MSEs larger than those for the two QML estimators. Further, significant differences show up for a substantial number of coefficients. Similarly, the QML_{T2} estimator, which takes error correlations into account, performs much better than the QML_{T1} estimator that does not. In table 5 with 1,000 observations this is even more evident. The majority of parameters show the SML estimator to be significantly better performing than the two other estimators, whereas the same is the case for the QML_{T2} versus the QML_{T1} estimator. The picture is the same in table 6 where the simulations are done on 3,000 observations.

Table 4, 5 and 6 show, that with larger correlation coefficients there are significant gains from using more sophisticated estimators and that the gains are apparent at all sample sizes analyzed here. Since it is often difficult to have a prior opinion on the size of the correlation coefficients for a given application, one recommendation would be to first apply the QML_{T2} estimator to assess the size of the correlation coefficients before considering to go on with the SML estimator. In that case, the QML_{T2} provides some very accurate starting values.

5 Concluding remarks

The results in this paper indicate that there is very little to gain from using a SML estimator compared to the two simpler QML estimators investigated here, if the absolute size of the error correlation coefficients is of the same magnitude as usually found in empirical studies. However, the error correlation structure can not be known prior to an application, and if these are large in absolute value there will be gains from using the asymptotically better SML estimator. In this case both QML estimators provide good starting values for the SML estimator; something which is useful in the cause of achieving convergence of the maximum likelihood routine.

Even if Monte Carlo simulations are subject to the problem of specificity which makes broad generalizations of the results difficult, this study has shown that for moderate sample sizes most commonly found in empirical applications simple QML estimators perform surprisingly well. The results herein also suggest that QML estimators of a similar type to those presented here might be useful in more general systems of censored equations.

References

[1] Avery, R.B., Hansen, L.P. & Hotz, V.J., 1983, 'Multi-period probit models and orthogonality condition estimation', *International Economic Review 24*, pp. 21-35.

[2] Barslund, M., 2007, 'Food consumption in urban Mozambique: A censored demand system approach', Discussion Paper, Department of Economics, University of Copenhagen.

[3] Börsch-Supan, A. & Hajivassiliou, V.A., 1993, 'Smooth Unbiased Multivariate Probability Simulatorfor Maximum Likelihood Estimation of Limited Dependent Variable Models', *Journal of Econometrics 58*, pp. 347-368.

[4] Cameron, A.C. & Trivedi, P.K., 2005, '*Microeconometrics - methods and applications*', Cambridge University Press, US.

[5] Cappellari, L & Jenkins, S.P., 2006, 'Calculation of multivariate normal probabilities by simulation, with applications to maximum simulated likelihood estimation', *The Stata Journal 6*, pp. 156-189.

[6] Christensen L.R., Jorgenson, D.W. & Lau, L.J., 1975, 'Transcendental Logarithmic Utility Functions', *American Economic Review 65*, pp. 367-383.

[7] Deaton, A. & Muellbauer, J., 1980, 'An almost ideal demand system', *American Economic Review 70*, pp. 312-326.

[8] Greene, W.H., 2000, *Econometric analysis*, 4. ed., Prentice Hall Int., US.

[9] Haan, P., and Uhlendorff, A., 2006, 'Estimation of multinomial logit models with unobserved heterogeneity using maximum simulated likelihood', *Stata Journal 6*, pp. 229-245.

[10] Harris, T.R. & Shonkwiler, J.S., 1997, 'Interdependence of Retail Businesses', *Growth and Change 28*, pp. 520-533.

[11] Lee, F. & Pitt, M.M., 1986, 'Microeconometric demand systems with binding non-negativity constraints: the dual approach', *Econometrica 54*, No. 5, pp. 1237-42.

[12] Moschini, G., 1998, 'The semiflexible almost ideal demand system', *European Economic Review 42*, pp. 349-364.

[13] Pudney, S., 1989, '*Modelling Individual Choice: Econometrics of Corners, Kinks and Holes*', Cambridge: Blackwell Publishers.

[14] Shonkwiler, J.S. & Yen, S.T., 1999, 'Two-step etsimation of a censored system of equations', *American Journal of Agricultural Economics 81*, pp. 972-82.

[15] Train, K., 2003, 'Discrete Choice Methods with Simulation', Cambridge, Cambridge University Press.

[16] Yen, S.T., Fang, C. & Su, S-J., 2004, 'Household food demand in urban China: a censored system approach', *Journal of Comparative Economics 32*, pp. 564-585.

[17] Yen, S.T. & Lin, B-H., 2002, 'Beverage consumption among US children and adolesscents: full-information and quasi maximum-likelihood estimation of a censored system', *European Review of Agricultural Economics 29*, pp. 85-103.

[18] Yen, S.T. & Lin, B-H., 2006, 'A sample selection approach to censored demand systems', *American Journal of Agricultural Economics 88*, pp. 742-749.

[19] Yen, S.T., Lin, B-H. & Smallwood, D.M., 2003, 'Quasi- and simulated-likelihood approaches to censored demand systems: food consumption by food stamp recipients in the United States', *American Journal of Agricultural Economics 85*(2), pp. 458-478.

[20] Wales, T.J. & Woodland A.D., 1983, 'Estimation of consumer demand systems with binding non-negativity constraints', *Journal of Econometrics 21*, pp. 262-285.

[21] White, H., 1982, 'Maximum likelihood estimation of misspecified models', *Econometrica 50*, pp.1-25.

Table 1: Simulation results. Base correlations, 200 observations (500 simulations).

Parameter		SML		QML-T2		QML-T1		Non-censored		Comparison		
	True value	Deviation mean (%)	MSE (x10000)	Deviation mean (%)	MSE (x10000)	Deviation mean (%)	MSE (x10000)	Deviation mean (%)	MSE (x10000)	(5) − (3)	(7) − (3)	(7) − (5)
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
$\gamma 11$	-0.06	-2.77	15.70	-2.75	15.87	-2.75	16.01	-31.86	11.20	$+$	$+$	$+^{**}$
$\gamma 12$	-0.03	5.72	7.42	5.58	7.43	5.47	7.46	-35.01	4.71	$+$	$+$	$+^{*}$
$\gamma 13$	0.05	-0.22	5.57	-0.09	5.61	-0.02	5.64	-45.79	7.24	$+$	$+$	$+^{**}$
$\gamma 14$	0.02	2.43	3.38	2.84	3.32	3.08	3.31	-26.90	1.78	$-^{**}$	$-^{*}$	$-$
$\gamma 22$	-0.01	-10.19	10.80	-9.07	10.89	-8.63	10.98	-78.08	6.48	$+$	$+$	$+^{*}$
$\gamma 23$	0.02	4.82	4.74	4.38	4.77	4.24	4.81	-34.82	2.24	$+$	$+$	$+^{**}$
$\gamma 24$	0.01	-9.88	2.91	-9.61	2.90	-9.53	2.91	-22.01	1.40	$-$	$+$	$+$
$\gamma 33$	-0.03	3.59	7.44	3.09	7.43	2.83	7.48	-37.77	3.80	$-$	$+$	$+^{*}$
$\gamma 34$	-0.02	0.60	2.93	0.53	2.89	0.61	2.88	-43.33	1.89	$-^{*}$	$-^{*}$	$-$
$\gamma 44$	0.01	1.35	3.26	1.40	3.24	1.50	3.24	-24.16	1.45	$-$	$-$	$+$
$\sigma 1$	0.6	-0.89	16.75	-0.93	16.65	-0.94	16.62	-26.34	255.86	$-$	$-$	$-$
$\sigma 2$	0.5	-1.32	11.10	-1.34	11.10	-1.34	11.11	-26.30	177.26	$-$	$+$	$+$
$\sigma 3$	0.4	-1.64	8.97	-1.65	9.00	-1.65	9.02	-38.66	241.84	$+$	$+$	$+$
$\sigma 4$	0.3	-1.45	4.11	-1.45	4.08	-1.44	4.07	-31.59	91.33	$-^{*}$	$-^{*}$	$-^{*}$
$\rho 12$	-0.2	-0.93	56.10	-1.50	55.90	.	.	-17.03	54.22	$-$	.	.
$\rho 13$	-0.15	-2.28	66.39	-2.95	65.21	.	.	-25.61	58.59	$-^{**}$	.	.
$\rho 14$	-0.08	2.92	69.92	2.37	68.95	.	.	-15.14	53.43	$-^{**}$	.	.
$\rho 23$	-0.15	0.58	64.63	-0.23	63.93	.	.	-21.61	50.37	$-$	.	.
$\rho 24$	-0.07	-1.99	58.73	-2.26	58.37	.	.	-17.42	46.81	$-$	.	.
$\rho 34$	-0.1	-3.81	69.18	-4.45	68.52	.	.	-25.92	56.72	$-$	.	.
$\alpha 01$	0.3	-0.82	21.60	-0.80	21.54	-0.79	21.51	40.07	154.52	$-$	$-$	$-$
$\alpha 02$	0.25	0.81	14.47	0.80	14.46	0.80	14.47	40.75	110.84	$-$	$+$	$+$
$\alpha 03$	0.05	-1.64	10.67	-1.59	10.73	-1.58	10.76	274.73	191.48	$+$	$+$	$+$
$\alpha 04$	0.1	1.45	5.89	1.46	5.87	1.46	5.87	78.76	64.36	$-$	$-$	$-$
$\beta 1$	-0.025	-9.35	17.37	-9.02	17.52	-8.87	17.57	-37.63	9.20	$+^{**}$	$+^{**}$	$+$
$\beta 2$	0.03	0.85	13.00	0.62	13.01	0.56	13.01	-31.85	7.59	$+$	$+$	$+$
$\beta 3$	-0.01	9.63	8.80	10.50	8.80	10.86	8.81	-47.66	2.98	$+$	$+$	$+$
$\beta 4$	0.02	6.68	4.41	6.47	4.41	6.40	4.42	-31.14	2.24	$+$	$+$	$+$

Notes: Parameters refer to the main text (eq. 4). σ and ρ are respectively the standard error and correlation of the error term. Deviation mean (column 2, 4, 6 and 8) shows the percentage deviation of the mean over the 500 simualtions from the true value. Column 10, 11 and 12 show the sign of difference between column 5 and 3, 7 and 3, and 7 and 5, respectively. *, ** and *** indicate that the compared columns are significantly different at 10, 5 and 1 percent based on two sided t-tests.

Table 2: Simulation results. Base correlations, 1000 observations (500 simulations).

Parameter	True value	SML Deviation mean (%)	SML MSE (x10000)	QML-T2 Deviation mean (%)	QML-T2 MSE (x10000)	QML-T1 Deviation mean (%)	QML-T1 MSE (x10000)	Non-censored Deviation mean (%)	Non-censored MSE (x10000)	Comparison (5) − (3)	Comparison (7) − (3)	Comparison (7) − (5)
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
$\gamma 11$	-0.06	-2.480	2.737	-2.483	2.718	-2.467	2.715	-29.71	4.53	$-$	$-$	$-$
$\gamma 12$	-0.03	1.124	1.284	1.095	1.293	1.093	1.300	-37.29	1.89	$+$	$+^{*}$	$+^{***}$
$\gamma 13$	0.05	-1.251	0.979	-1.251	0.982	-1.252	0.985	-45.74	5.61	$+$	$+$	$+$
$\gamma 14$	0.02	0.703	0.591	0.801	0.591	0.843	0.593	-30.16	0.64	$+$	$+$	$+$
$\gamma 22$	-0.01	-8.771	2.171	-8.451	2.186	-8.383	2.205	-72.66	1.66	$+$	$+$	$+^{***}$
$\gamma 23$	0.02	-1.747	0.898	-1.782	0.906	-1.793	0.912	-39.12	0.98	$+$	$+$	$+^{**}$
$\gamma 24$	0.01	-0.674	0.573	-0.875	0.572	-0.973	0.574	-18.70	0.29	$-$	$+$	$+$
$\gamma 33$	-0.03	-4.146	1.392	-4.184	1.395	-4.181	1.401	-42.23	2.08	$+$	$+$	$+$
$\gamma 34$	-0.02	2.443	0.506	2.472	0.508	2.477	0.510	-41.47	0.89	$+$	$+$	$+^{*}$
$\gamma 44$	0.01	3.529	0.779	3.238	0.779	3.157	0.780	-25.82	0.39	$+$	$+$	$+$
$\sigma 1$	0.6	-0.164	3.016	-0.163	3.022	-0.161	3.027	-25.95	243.61	$+$	$+$	$+$
$\sigma 2$	0.5	-0.190	2.182	-0.202	2.187	-0.206	2.191	-25.80	167.25	$+$	$+$	$+$
$\sigma 3$	0.4	-0.363	1.660	-0.359	1.667	-0.358	1.673	-37.16	221.60	$+$	$+$	$+^{**}$
$\sigma 4$	0.3	-0.164	0.825	-0.159	0.822	-0.156	0.823	-30.78	85.55	$-$	$-$	$+$
$\rho 12$	-0.2	0.532	11.451	0.421	11.494	.	.	-15.55	18.28	$+$	.	.
$\rho 13$	-0.15	-0.685	11.999	-0.770	12.013	.	.	-22.55	18.82	$+$	.	.
$\rho 14$	-0.08	0.701	14.008	0.592	13.913	.	.	-17.17	12.03	$-$	.	.
$\rho 23$	-0.15	0.866	12.489	0.818	12.442	.	.	-20.93	17.86	$-$	.	.
$\rho 24$	-0.07	-0.573	11.692	-0.757	11.769	.	.	-17.59	10.39	$+$	.	.
$\rho 34$	-0.1	-2.027	13.230	-2.049	13.162	.	.	-22.59	14.37	$-$	.	.
$\alpha 01$	0.3	-0.369	4.104	-0.376	4.104	-0.380	4.109	39.91	145.28	$-$	$+$	$+$
$\alpha 02$	0.25	0.212	2.894	0.228	2.874	0.233	2.870	40.13	101.93	$-^{***}$	$-^{**}$	$-^{*}$
$\alpha 03$	0.05	3.536	2.152	3.511	2.145	3.508	2.146	278.07	193.97	$-$	$-$	$+$
$\alpha 04$	0.1	-0.610	1.111	-0.615	1.110	-0.617	1.110	77.51	60.51	$-$	$-$	$+$
$\beta 1$	-0.025	-3.961	3.113	-3.985	3.115	-4.007	3.117	-36.44	2.37	$+$	$+$	$+$
$\beta 2$	0.03	0.829	1.939	0.998	1.945	1.066	1.947	-33.14	2.03	$+$	$+$	$+$
$\beta 3$	-0.01	8.575	1.536	8.675	1.538	8.681	1.539	-49.14	0.73	$+$	$+$	$+$
$\beta 4$	0.02	0.411	0.994	0.383	0.994	0.383	0.995	-37.86	0.97	$+$	$+$	$+$

Notes: Parameters refer to the main text (eq. 4). σ and ρ are respectively the standard error and correlation of the error term. Deviation mean (column 2, 4, 6 and 8) shows the percentage deviation of the mean over the 500 simualtions from the true value. Column 10, 11 and 12 show the sign of difference between column 5 and 3, 7 and 3, and 7 and 5, respectively. *, ** and *** indicate that the compared columns are significantly different at 10, 5 and 1 percent based on two sided t-tests.

Table 3: Simulation results. Base correlations, 3000 observations (500 simulations).

Parameter		SML		QML-T2		QML-T1		Non-censored		Comparison		
	True value	Deviation mean (%)	MSE (x10000)	Deviation mean (%)	MSE (x10000)	Deviation mean (%)	MSE (x10000)	Deviation mean (%)	MSE (x10000)	(5) − (3)	(7) − (3)	(7) − (5)
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
$\gamma11$	-0.06	-0.775	0.916	-0.838	0.924	-0.865	0.930	-28.76	3.45	+	+*	+**
$\gamma12$	-0.03	-1.078	0.428	-1.064	0.426	-1.059	0.425	-38.01	1.52	−	−	−
$\gamma13$	0.05	-0.163	0.292	-0.154	0.294	-0.150	0.295	-45.59	5.31	+	+	+**
$\gamma14$	0.02	-0.138	0.199	-0.103	0.201	-0.096	0.201	-30.31	0.45	+	+	+
$\gamma22$	-0.01	-2.218	0.688	-1.680	0.692	-1.471	0.695	-71.25	0.88	+	+	+**
$\gamma23$	0.02	-2.307	0.324	-2.391	0.324	-2.410	0.325	-39.38	0.75	−	+	+
$\gamma24$	0.01	-0.099	0.201	-0.278	0.202	-0.333	0.202	-17.71	0.12	+	+	+
$\gamma33$	-0.03	-1.212	0.493	-1.286	0.490	-1.297	0.490	-41.52	1.72	−	−	+
$\gamma34$	-0.02	-0.810	0.187	-0.833	0.188	-0.838	0.188	-43.20	0.82	+	+	+
$\gamma44$	0.01	-0.374	0.272	-0.343	0.273	-0.330	0.273	-27.92	0.19	+	+	+
$\sigma1$	0.6	-0.077	0.986	-0.075	0.990	-0.074	0.993	-25.57	235.79	+	+	+**
$\sigma2$	0.5	-0.107	0.684	-0.107	0.681	-0.106	0.680	-25.58	163.88	−	−	−
$\sigma3$	0.4	-0.067	0.522	-0.067	0.527	-0.067	0.529	-37.51	225.31	+**	+**	+***
$\sigma4$	0.3	-0.065	0.269	-0.064	0.270	-0.064	0.270	-30.73	85.10	+	+	+
$\rho12$	-0.2	-0.173	3.596	-0.209	3.592	.	.	-15.98	12.93	−	.	.
$\rho13$	-0.15	-0.067	4.164	-0.131	4.228	.	.	-22.20	13.91	+**	.	.
$\rho14$	-0.08	2.393	4.109	2.330	4.106	.	.	-15.06	4.74	−	.	.
$\rho23$	-0.15	0.756	4.211	0.666	4.215	.	.	-20.86	12.56	+	.	.
$\rho24$	-0.07	-1.203	4.073	-1.242	4.084	.	.	-17.35	4.72	+	.	.
$\rho34$	-0.1	-0.183	4.531	-0.227	4.592	.	.	-22.78	8.44	+**	.	.
$\alpha01$	0.3	0.189	1.330	0.186	1.331	0.186	1.332	40.36	147.24	+	+	+
$\alpha02$	0.25	-0.069	0.959	-0.069	0.959	-0.069	0.959	40.07	100.82	−	−	+
$\alpha03$	0.05	-1.180	0.776	-1.171	0.781	-1.170	0.783	275.76	190.32	+**	+**	+***
$\alpha04$	0.1	0.299	0.393	0.301	0.393	0.301	0.393	77.79	60.67	−	−	+
$\beta1$	-0.025	0.445	1.007	0.564	1.006	0.615	1.006	-33.30	1.20	−	−	+
$\beta2$	0.03	0.560	0.770	0.560	0.772	0.565	0.774	-31.76	1.31	+	+	+
$\beta3$	-0.01	3.056	0.530	3.139	0.527	3.171	0.526	-55.25	0.48	−	−	−
$\beta4$	0.02	-0.024	0.273	-0.019	0.273	-0.017	0.273	-36.71	0.66	+	+	+

Notes: Parameters refer to the main text (eq. 4). σ and ρ are respectively the standard error and correlation of the error term. Deviation mean (column 2, 4, 6 and 8) shows the percentage deviation of the mean over the 500 simualtions from the true value. Column 10, 11 and 12 show the sign of difference between column 5 and 3, 7 and 3, and 7 and 5, respectively. *, ** and *** indicate that the compared columns are significantly different at 10, 5 and 1 percent based on two sided t-tests.

Table 4: Simulation results. Large correlations, 200 observations (500 simulations).

Parameter	True value	SML Deviation mean (%)	SML MSE (x10000)	QML-T2 Deviation mean (%)	QML-T2 MSE (x10000)	QML-T1 Deviation mean (%)	QML-T1 MSE (x10000)	Comparison (5) − (3)	Comparison (7) − (3)	Comparison (7) − (5)
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
$\gamma 11$	-0.06	-2.626	16.185	-2.902	16.923	-3.068	17.348	$+^{**}$	$+^{***}$	$+^{***}$
$\gamma 12$	-0.03	5.202	7.938	5.446	8.029	5.342	8.119	$+$	$+$	$+^{***}$
$\gamma 13$	0.05	-0.847	5.784	-0.753	5.964	-0.688	6.036	$+^{**}$	$+^{**}$	$+^{***}$
$\gamma 14$	0.02	2.229	3.384	2.305	3.365	2.538	3.386	$-$	$+$	$+$
$\gamma 22$	-0.01	-16.906	11.324	-13.258	11.826	-11.972	12.203	$+^{**}$	$+^{***}$	$+^{***}$
$\gamma 23$	0.02	5.599	5.004	5.381	5.119	5.273	5.222	$+$	$+^{*}$	$+^{***}$
$\gamma 24$	0.01	-10.918	3.022	-10.530	3.096	-10.348	3.127	$+$	$+^{*}$	$+^{**}$
$\gamma 33$	-0.03	3.444	7.858	2.715	8.069	2.477	8.227	$+$	$+^{**}$	$+^{***}$
$\gamma 34$	-0.02	-1.043	2.961	-1.160	2.983	-0.929	2.988	$+$	$+$	$+$
$\gamma 44$	0.01	3.549	3.414	4.721	3.403	5.032	3.408	$-$	$-$	$+$
$\sigma 1$	0.6	-0.798	16.560	-0.980	16.587	-0.999	16.647	$+$	$+$	$+$
$\sigma 2$	0.5	-1.522	11.383	-1.546	11.412	-1.535	11.456	$+$	$+$	$+$
$\sigma 3$	0.4	-1.787	8.933	-1.841	9.297	-1.825	9.398	$+^{**}$	$+^{**}$	$+^{***}$
$\sigma 4$	0.3	-1.726	4.044	-1.722	4.071	-1.715	4.079	$+$	$+$	$+$
$\rho 12$	-0.4	0.018	42.495	-0.545	43.286	.	.	$+$	.	.
$\rho 13$	-0.3	-0.485	54.258	-1.272	55.339	.	.	$+$	.	.
$\rho 14$	-0.16	3.140	66.555	2.427	66.150	.	.	$-$	.	.
$\rho 23$	-0.3	0.420	56.633	-0.538	57.726	.	.	$+$	.	.
$\rho 24$	-0.14	-1.745	55.820	-1.287	57.044	.	.	$+$	.	.
$\rho 34$	-0.2	-1.683	62.616	-2.445	63.271	.	.	$+$	.	.
$\alpha 01$	0.3	-0.859	21.747	-0.755	21.653	-0.743	21.654	$-$	$-$	$+$
$\alpha 02$	0.25	1.280	14.350	1.181	14.481	1.157	14.513	$+$	$+$	$+^{*}$
$\alpha 03$	0.05	-1.166	10.672	-1.220	10.889	-1.307	10.962	$+$	$+^{*}$	$+$
$\alpha 04$	0.1	1.599	5.453	1.584	5.489	1.580	5.499	$+$	$+$	$+$
$\beta 1$	-0.025	-8.928	17.212	-8.708	17.594	-8.450	17.727	$+^{**}$	$+^{**}$	$+$
$\beta 2$	0.03	1.123	12.583	0.290	12.782	0.208	12.873	$+$	$+^{**}$	$+$
$\beta 3$	-0.01	12.967	9.272	15.359	9.332	16.089	9.391	$+$	$+$	$+^{**}$
$\beta 4$	0.02	6.409	4.480	5.921	4.522	5.905	4.538	$+$	$+$	$+$

Notes: Parameters refer to the main text (eq. 4). σ and ρ are respectively the standard error and correlation of the error term. Deviation mean (column 2, 4, 6 and 8) shows the percentage deviation of the mean over the 500 simualtions from the true value. Column 10, 11 and 12 show the sign of difference between column 5 and 3, 7 and 3, and 7 and 5, respectively. *, ** and *** indicate that the compared columns are significantly different at 10, 5 and 1 percent based on two sided t-tests.

Table 5: Simulation results. Large correlations, 1,000 observations (500 simulations).

Parameter	True value	SML Deviation mean (%)	SML MSE (x10000)	QML-T2 Deviation mean (%)	QML-T2 MSE (x10000)	QML-T1 Deviation mean (%)	QML-T1 MSE (x10000)	Comparison (5) − (3)	Comparison (7) − (3)	Comparison (7) − (5)
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
$\gamma 11$	-0.06	-2.253	2.965	-2.459	2.975	-2.397	2.995	$+$	$+$	$+^{*}$
$\gamma 12$	-0.03	1.594	1.376	1.407	1.411	1.185	1.450	$+^{*}$	$+^{**}$	$+^{***}$
$\gamma 13$	0.05	-1.285	1.016	-1.201	1.057	-1.182	1.073	$+^{**}$	$+^{***}$	$+^{***}$
$\gamma 14$	0.02	1.096	0.583	1.304	0.593	1.249	0.598	$+$	$+$	$+^{**}$
$\gamma 22$	-0.01	-9.747	2.259	-8.951	2.350	-8.396	2.435	$+^{*}$	$+^{**}$	$+^{***}$
$\gamma 23$	0.02	-2.811	0.930	-2.730	0.986	-2.741	1.008	$+^{***}$	$+^{***}$	$+^{***}$
$\gamma 24$	0.01	-1.001	0.605	-1.395	0.613	-1.589	0.618	$+$	$+$	$+^{**}$
$\gamma 33$	-0.03	-4.361	1.354	-4.240	1.456	-4.236	1.492	$+^{***}$	$+^{***}$	$+^{***}$
$\gamma 34$	-0.02	2.006	0.522	2.410	0.534	2.436	0.539	$+$	$+^{*}$	$+^{***}$
$\gamma 44$	0.01	4.491	0.761	3.452	0.773	3.678	0.780	$+$	$+$	$+^{*}$
$\sigma 1$	0.6	-0.166	2.920	-0.174	3.009	-0.171	3.030	$+^{*}$	$+^{**}$	$+^{***}$
$\sigma 2$	0.5	-0.130	2.061	-0.153	2.129	-0.162	2.148	$+^{***}$	$+^{***}$	$+^{***}$
$\sigma 3$	0.4	-0.307	1.579	-0.283	1.663	-0.279	1.681	$+^{***}$	$+^{***}$	$+^{***}$
$\sigma 4$	0.3	-0.295	0.822	-0.283	0.819	-0.283	0.820	$-$	$-$	$+$
$\rho 12$	-0.4	0.170	8.536	0.029	8.922	.	.	$+^{***}$	.	.
$\rho 13$	-0.3	-0.632	9.654	-0.786	10.316	.	.	$+^{***}$	.	.
$\rho 14$	-0.16	0.232	13.439	0.219	13.460	.	.	$+$	.	.
$\rho 23$	-0.3	0.447	10.124	0.439	10.440	.	.	$+$	.	.
$\rho 24$	-0.14	0.166	11.079	-0.042	11.511	.	.	$+^{***}$	.	.
$\rho 34$	-0.2	-0.514	12.119	-0.512	12.586	.	.	$+^{**}$	.	.
$\alpha 01$	0.3	-0.352	4.113	-0.365	4.113	-0.376	4.124	$+$	$+$	$+$
$\alpha 02$	0.25	0.160	2.856	0.183	2.821	0.198	2.823	$-$	$-$	$+^{*}$
$\alpha 03$	0.05	2.892	2.160	2.628	2.198	2.590	2.210	$+$	$+^{**}$	$+$
$\alpha 04$	0.1	-0.723	1.094	-0.759	1.095	-0.788	1.097	$+$	$+$	$+$
$\beta 1$	-0.025	-3.892	3.094	-3.931	3.119	-3.915	3.127	$+$	$+$	$+$
$\beta 2$	0.03	-0.887	1.826	-0.544	1.862	-0.435	1.865	$+^{**}$	$+^{**}$	$+^{**}$
$\beta 3$	-0.01	9.496	1.465	9.814	1.519	9.259	1.542	$+^{***}$	$+^{***}$	$+$
$\beta 4$	0.02	-0.740	0.994	-0.427	0.998	-0.355	0.997	$+$	$+$	$-$

Notes: Parameters refer to the main text (eq. 4). σ and ρ are respectively the standard error and correlation of the error term. Deviation mean (column 2, 4, 6 and 8) shows the percentage deviation of the mean over the 500 simualtions from the true value. Column 10, 11 and 12 show the sign of difference between column 5 and 3, 7 and 3, and 7 and 5, respectively. *, ** and *** indicate that the compared columns are significantly different at 10, 5 and 1 percent based on two sided t-tests.

Table 6: Simulation results. Large correlations, 3,000 observations (500 simulations).

Parameter	True value	SML Deviation mean (%)	SML MSE (x10000)	QML-T2 Deviation mean (%)	QML-T2 MSE (x10000)	QML-T1 Deviation mean (%)	QML-T1 MSE (x10000)	Comparison (5) − (3)	Comparison (7) − (3)	Comparison (7) − (5)
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
$\gamma 11$	-0.06	-0.614	0.948	-0.824	0.996	-0.897	1.018	$+^{***}$	$+^{***}$	$+^{***}$
$\gamma 12$	-0.03	-1.202	0.460	-1.158	0.459	-1.160	0.461	$-$	$+$	$+$
$\gamma 13$	0.05	-0.454	0.324	-0.366	0.334	-0.352	0.339	$+$	$+^{**}$	$+^{***}$
$\gamma 14$	0.02	-0.129	0.212	-0.001	0.219	0.008	0.220	$+^{**}$	$+^{**}$	$+^{*}$
$\gamma 22$	-0.01	-2.396	0.714	-0.887	0.740	-0.408	0.756	$+^{*}$	$+^{**}$	$+^{***}$
$\gamma 23$	0.02	-1.967	0.334	-2.269	0.342	-2.302	0.347	$+$	$+^{*}$	$+^{***}$
$\gamma 24$	0.01	-0.241	0.211	-0.792	0.218	-0.886	0.221	$+^{**}$	$+^{**}$	$+^{***}$
$\gamma 33$	-0.03	-1.113	0.498	-1.407	0.519	-1.440	0.527	$+^{*}$	$+^{**}$	$+^{***}$
$\gamma 34$	-0.02	-0.722	0.179	-0.776	0.187	-0.777	0.189	$+^{***}$	$+^{***}$	$+^{***}$
$\gamma 44$	0.01	-0.403	0.275	-0.415	0.281	-0.427	0.282	$+^{**}$	$+^{*}$	$+$
$\sigma 1$	0.6	-0.083	0.945	-0.080	0.982	-0.078	0.992	$+$	$+^{***}$	$+^{***}$
$\sigma 2$	0.5	-0.143	0.652	-0.146	0.651	-0.144	0.655	$-$	$+$	$+^{**}$
$\sigma 3$	0.4	-0.013	0.496	0.004	0.525	0.006	0.532	$+^{***}$	$+^{***}$	$+^{***}$
$\sigma 4$	0.3	-0.072	0.272	-0.063	0.276	-0.062	0.276	$+$	$+$	$+^{*}$
$\rho 12$	-0.4	0.032	2.660	-0.011	2.703	.	.	$+$	.	.
$\rho 13$	-0.3	-0.131	3.681	-0.198	3.893	.	.	$+^{***}$	.	.
$\rho 14$	-0.16	0.988	3.934	0.899	3.956	.	.	$+$	.	.
$\rho 23$	-0.3	0.044	3.332	-0.029	3.500	.	.	$+^{***}$	.	.
$\rho 24$	-0.14	-0.695	3.658	-0.698	3.766	.	.	$+^{**}$	.	.
$\rho 34$	-0.2	-0.291	3.959	-0.135	4.148	.	.	$+^{***}$	.	.
$\alpha 01$	0.3	0.202	1.310	0.191	1.326	0.190	1.333	$+$	$+^{*}$	$+$
$\alpha 02$	0.25	-0.043	0.935	-0.044	0.938	-0.046	0.941	$+$	$+$	$+$
$\alpha 03$	0.05	-1.589	0.762	-1.701	0.768	-1.710	0.771	$+$	$+$	$+$
$\alpha 04$	0.1	0.315	0.404	0.295	0.403	0.295	0.403	$-$	$-$	$+$
$\beta 1$	-0.025	0.094	0.994	0.458	1.000	0.562	1.004	$+$	$+$	$+^{**}$
$\beta 2$	0.03	0.621	0.773	0.669	0.782	0.698	0.786	$+$	$+$	$+$
$\beta 3$	-0.01	2.300	0.530	2.348	0.536	2.388	0.539	$+$	$+$	$+$
$\beta 4$	0.02	0.591	0.263	0.675	0.263	0.695	0.263	$+$	$+$	$+$

Notes: Parameters refer to the main text (eq. 4). σ and ρ are respectively the standard error and correlation of the error term. Deviation mean (column 2, 4, 6 and 8) shows the percentage deviation of the mean over the 500 simualtions from the true value. Column 10, 11 and 12 show the sign of difference between column 5 and 3, 7 and 3, and 7 and 5, respectively. *, ** and *** indicate that the compared columns are significantly different at 10, 5 and 1 percent based on two sided t-tests.