A Structural-Equation GME Estimator

by

Thomas L. Marsh[a], Ron C. Mittelhammer[b], and N. Scott Cardell[c]

Selected Paper
1998 AAEA Annual Meeting
Salt Lake City

[a]Assistant Professor, Department of Agricultural Economics, Kansas State University; [b]Professor, Department of Agricultural Economics and Adjunct Professor, Program in Statistics, Washington State University; [c]Salford Systems, San Diego, CA

Abstract. A generalized maximum entropy estimator is developed for the linear simultaneous equations systems model. We provide results on large and small sample properties of the estimator. Empirical results illustrate efficiency advantages of the generalized maximum entropy estimator proposed in this study over traditional estimators (e.g., 2SLS and 3SLS).

Introduction

The simultaneous equation systems model (SESM) has been extensively applied in many areas of agricultural economics. For the linear SESM, the traditional estimators include two stage least squares, three stage least squares, and maximum likelihood methods that yield consistent estimates of structural parameters by correcting for simultaneity between the endogenous variables and the disturbance terms of the statistical model. However, in the presence of small samples or ill-posed problems, the traditional approaches may provide parameter estimates with high variance and/or bias, or provide no solution at all.

Recently generalized maximum entropy (GME) estimators were proposed for the general linear model in both the moment (Golan, Judge, and Miller 1996) and data-constrained forms (Mittelhammer and Cardell 1998), for the data-constrained form under autocorrelated error terms (Sever, Murat, and Cardell 1998), and for the linear SESM in several different ways (Golan, Judge, and Miller 1996; Golan and Judge 1997). Properties of the GME estimators in the context of the linear SESMs are indeed promising, judging from a limited number of Monte Carlo simulations of the estimators' performance. However a rigorous statistical foundation relating to the performance of the maximum entropy estimator - including the finite sample properties of the estimator, as well as the asymptotic properties of both the estimator and test statistics based on the estimator - has not yet been fully developed and is in need of further study before the procedure can be confidently adopted for widespread use by empirical economists.

Some Monte Carlo results have suggested that the GME estimator is superior to traditional estimators in the presence of small samples or when the underlying sampling method is incomplete or incorrectly specified (see Golan, Judge, and Miller 1996). In extreme cases of incomplete or incorrect sampling methods the problem is commonly referred to as an ill-posed problem. Maximum entropy provides a principle or formalism to cope with such ill-posed problems and provides a method for improving on estimation efficiency over traditional estimation approaches even for problems that are not ill-posed (see, for example, Shannon 1948 or Donoho 1992).

In this paper we propose a new generalized maximum entropy estimator for the linear SESM with contemporaneous correlation in the error structure. It is a data constrained structural-equation generalized maximum entropy estimator, or SGME, which accounts explicitly for simultaneity problems inherent in a system of equations. We provide results on the large sample properties of the estimator including consistency and asymptotic normality. We also provide a collection of asymptotically chisquare-distributed test procedures capable of testing all of the hypothesis tests typically performed in applied econometrics. For small sample properties, we provide empirical evidence from Monte Carlo sampling experiments.

As a basis for an intensive Monte Carlo sampling experiment, we analyze an overdetermined simultaneous systems model. The Monte Carlo simulations suggest that in larger sample situations the SGME is generally not dominated by two and three stage least squares estimators, and in small sample situations the SGME is the much more efficient estimator.

Overall, the results of this paper elevate the SGME to a mainstream status for applied econometricians that is comparable to traditional estimation techniques; including rigorous theorems on asymptotic properties of the estimator and test statistics, and empirical evidence suggesting the contexts in which the SGME is superior to the traditional approaches.

Linear Simultaneous Equations Statistical Model (SESM)

Consider the linear simultaneous equations model in matrix form

$$Y\Gamma + X\mathrm{B} + \mathrm{E} = 0 \qquad \textbf{(1)}$$

Here Γ and B represent the unknown true parameter values and E represents the unobserved true random errors of the system. For G equations, $Y=(\mathbf{y}_1 \ldots \mathbf{y}_G)$ is a (N x G) matrix of jointly determined endogenous variables, $\Gamma=(\gamma_1 \ldots \gamma_G)$ is an invertible (G x G) matrix of structural coefficients of the endogenous variables, $X=(\mathbf{x}_1 \ldots \mathbf{x}_K)$ is a (N x K) matrix of exogenous or predetermined variables that has full column rank, $\mathrm{B}=(\beta_1 \ldots \beta_G)$ is a (K x G) matrix of coefficients of the predetermined variables, and $\mathrm{E}=(\epsilon_1 \ldots \epsilon_G)$ is a (N x G) matrix of unobserved random errors. The standard stochastic assumptions of the error vectors are that $E[\epsilon_\mathbf{i}] = \mathbf{0}$ for i=1,...,G, $E[\epsilon_\mathbf{i} \epsilon_\mathbf{i}'] = \sigma_{ii} I_N$ for i=1,...,G, and $E[\epsilon_\mathbf{i} \epsilon_\mathbf{j}'] = \sigma_{ij} I_N$ for $i \neq j$ and i,j=1,...,G. If $\epsilon=\mathrm{vec}(\epsilon_1,\ldots,\epsilon_G)$[1], then this implies that the covariance matrix of ϵ is defined by $E[\epsilon \epsilon'] = \Sigma \otimes I_N$ with the (G x G) matrix Σ containing the unknown σ_{ij}'s for i,j=1,...,G.

An alternative form of the gth structural-equation is the column vectorized form

$$\boldsymbol{y}_g = Y_{(-g)} \gamma_g + X_g \beta_g + \epsilon_g \qquad \textbf{(2)}$$

where $Y_{(-g)}$ represents a (N x G_g) matrix of endogenous variables included in the gth equation explicitly excluding the $\mathbf{y}_g$ vector[2], γ_g is a (G_g x 1) vector of coefficients on the endogenous

---

[1]The notation $\epsilon=\mathrm{vec}(\epsilon_1,\ldots,\epsilon_G)$ represents the vertical concatenation of the G vectors $\epsilon_1,\ldots,\epsilon_G$, which in this case are each of dimension (N x 1), into a (NG x 1) vector.

[2]The subscript notation (-g) implies that $Y_{(-g)}$ is a (NxG_g) matrix formed by: (1) always removing the column vector $\mathbf{y}_g$ and (2) possibly removing other endogenous vectors due to zero exclusions. For example, if only $\mathbf{y}_g$ is removed from Y then $Y_{(-g)}$ is a (NxG-1) matrix. On the other hand, if $\mathbf{y}_g$ and $\mathbf{y}_{g+2}$ are removed from Y then $Y_{(-g)}$ is a (NxG-2) matrix.

variables contained in $Y_{(-g)}$, X_g is a (N x K_g) matrix that represents the exogenous variables included in the gth equation, β_g is a corresponding (K_g x 1) vector of coefficients on the exogenous variables contained in X_g, and ϵ_g is a (N x 1) residual vector of the gth equation.

The reduced form model is obtained by post-multiplying (1) by Γ^{-1}

$$Y = X(-\mathrm{B}\Gamma^{-1}) + (-\mathrm{E}\Gamma^{-1}) = X\Pi + \mathbf{V} \tag{3}$$

where Π=-BΓ^{-1}=(π_1,...,π_G) is a (K x G) matrix of reduced form coefficients and **V**= -EΓ^{-1} =(ν_1,...,ν_G) is a (N x G) matrix of reduced form disturbances. The vectorized reduced form of the gth equation can be written as

$$\mathbf{y}_g = X\pi_g + \nu_g \tag{4}$$

Equations (1)-(4) and the identification conditions each play important roles in formulating a consistent generalized maximum entropy estimator of the SESM, which is developed below.

A Structural-Equation GME Formulation

In this section we derive a structural generalized maximum entropy estimator of the SESM in equation (1). In its formulation, the objective function is based on Shannon's entropy function. Shannon (1948) used an axiomatic method to define the entropy of the discrete distribution of probabilities $\mathbf{p} = (p_1,...,p_K)'$ as the measure $h(\mathbf{p}) = -\sum_{i=1}^{K} p_i \ln(p_i)$ where $h(\mathbf{0})$=0. The measure $h(\mathbf{p})$ reaches a maximum at $p_1 = ... = p_K = 1/K$ when the probabilities are uniform. Under the maximum entropy principle $h(\mathbf{p})$ is maximized subject to data and other constraints, which yields a distribution of $\mathbf{p}$ that is closest to the uniform distribution and still consistent with the data.

To motivate the structural GME formulation of the SESM in (1), recall the vectorized version of the *g*th equation given in (2). Following Thiel (1971), it can be rewritten as

$$\begin{aligned}\mathbf{y}_{\mathrm{g}} &= E[Y_{(-g)}]\gamma_g + X_g\beta_g + \epsilon_{\mathrm{g}} + (Y_{(-g)} - E[Y_{(-g)}])\gamma_g \\ &= E[Y_{(-g)}]\gamma_g + X_g\beta_g + \boldsymbol{\mu}_{\mathrm{g}}\end{aligned} \tag{5}$$

As a result of this formulation, equation (5) has a nonstochastic observation matrix and the elements of $\boldsymbol{\mu}_{\mathrm{g}}$ have zero mean, constant covariance matrix, and are uncorrelated under the prevailing assumptions. From the reduced form equation (3) we have that

$$E[Y_{(-g)}] = E[X\Pi_{\boldsymbol{(-g)}} + \mathbf{V}_{\mathrm{g}}] = X\Pi_{\boldsymbol{(-g)}} \tag{6}$$

where $\Pi_{(\text{-g})}$ is a (K x G_{g}) matrix of reduced form coefficients that are associated with the endogenous variables in $Y_{(-g)}$. Combining the last two equations yields the structural model

$$\mathbf{y}_{\mathrm{g}} = X\,\Pi_{(-g)}\gamma_g + X_g\beta_g + \boldsymbol{\mu}_{\mathrm{g}} = Z_g\delta_g + \boldsymbol{\mu}_g \tag{7}$$

where $Z_g = \left(X\Pi_{(-g)} \quad X_g\right)$ and $\delta_g = \mathrm{vec}(\gamma_g, \beta_g)$. Equation (7) historically motivated 2SLS and 3SLS, whereby OLS is applied to obtain predicted values of $E[Y_{(-g)}]$ in (6), and then predicted values of $X\,\Pi_{(-g)}$ are used in (7).

Data-Constrained Structural GME model of the SESM

We formulate a structural generalized maximum entropy model (SGME) by reparameterizing the coefficients and error terms of (4) and (7) as: $\beta = S^{\beta}\mathbf{p}^{\beta}$, $\gamma = S^{\gamma}\mathbf{p}^{\gamma}$, $\pi = S^{\pi}\mathbf{p}^{\pi}$, $\nu = S^{z}\mathbf{z}$, and $\boldsymbol{\mu} = S^{w}\mathbf{w}$. The coefficient estimates of $\pi = \mathrm{vec}(\pi_1,...,\pi_G)$, $\gamma = \mathrm{vec}(\gamma_1,...,\gamma_G)$, and $\beta = \mathrm{vec}(\beta_1,...,\beta_G)$ can be derived from the following constrained GME problem

$$\max_{\mathbf{p}^{\pi},\mathbf{p}^{\gamma},\mathbf{p}^{\beta},\mathbf{w},\mathbf{z}} \{-\mathbf{p}^{\pi\prime}\ln\mathbf{p}^{\pi} - \mathbf{p}^{\gamma\prime}\ln\mathbf{p}^{\gamma} - \mathbf{p}^{\beta\prime}\ln\mathbf{p}^{\beta} - \mathbf{w}^{\prime}\ln\mathbf{w} - \mathbf{z}^{\prime}\ln\mathbf{z}\} \quad \textbf{(8)}$$

subject to:

$$\mathbf{y} = (I_G \otimes X)(S^{\pi}_{(-)}\mathbf{p}^{\pi})(S^{\gamma}\mathbf{p}^{\gamma}) + X^{\beta}S^{\beta}\mathbf{p}^{\beta} + S^{w}\mathbf{w} \quad \textbf{(9)}$$

$$\mathbf{y} = (I_G \otimes X)(S^{\pi}\mathbf{p}^{\pi}) + S^{z}\mathbf{z} \quad \textbf{(10)}$$

$$\sum_{m=1}^{M} \mathrm{p}^{\gamma}_{\mathrm{igm}} = 1, \quad \sum_{m=1}^{M} \mathrm{p}^{\beta}_{\mathrm{kgm}} = 1, \quad \sum_{m=1}^{M} \mathrm{p}^{\pi}_{\mathrm{kgm}} = 1, \quad \sum_{m=1}^{M} \mathrm{w}_{\mathrm{ngm}} = 1, \quad \sum_{m=1}^{M} \mathrm{z}_{\mathrm{ngm}} = 1 \quad \textbf{(11)}$$

The intuition behind incorporating both the reduced and structural components is based on several observations. First, without the reduced form component in (10), the parameters of the structural component in (9) are not identified. Second, the structural component in (9) draws the SGME estimates to the true parameter values as the sample size increases.

We define $\bar{K} = \sum_{g=1}^{G} K_g$ to denote the number of unknown β_{kg}'s and $\bar{G} = \sum_{g=1}^{G} G_g$ to denote the number of unknown γ_{ig}'s. Then together with the *KG* reduced form parameters π_{kg}'s the total number of unknown parameters of the system is given by $Q = \bar{K} + \bar{G} + KG$. In (9) the matrix X^{β} = diag(X_1,..., X_G) represents a (GN x $\bar{K}$) block diagonal matrix and $\mathbf{y}$=vec($\mathbf{y}_1$,...,$\mathbf{y}_G$) is a (GN x 1) vector of endogenous variables. The (NGM x 1) vectors $\mathbf{w}$=vec($\mathbf{w}_{11}$,...,$\mathbf{w}_{NG}$) and $\mathbf{z}$=vec($\mathbf{z}_{11}$,...,$\mathbf{z}_{NG}$) represent vertical concatenations of sets of (M x 1) subvectors for N observations (n=1, ... ,N) and G equations (g=1, ...,G), where each subvector $\mathbf{w}_{ng}$=(w_{ng1} ... w_{ngM})′ and $\mathbf{z}_{ng}$=(z_{ng1} ... z_{ngM})′ is made up of M support points (m=1,...,M) for the structural and reduced form residuals μ_{ng} and v_{ng}, respectively. Similarly, the vector $\mathbf{p}^{\pi} = \mathrm{vec}(\mathbf{p}_{11}^{\pi},...,\mathbf{p}_{KG}^{\pi})$ is a (KGM x 1) vector, $\mathbf{p}^{\gamma} = \mathrm{vec}(\mathbf{p}_{11}^{\gamma},...,\mathbf{p}_{KG}^{\gamma})$ is a ($M\bar{G}$ x 1) vector, and $\mathbf{p}^{\beta} = \mathrm{vec}(\mathbf{p}_{11}^{\beta},...,\mathbf{p}_{KG}^{\beta})$ is a ($\bar{K}$M x 1) vector.

Regarding the S^i support matrices (for i=π, γ, β, z,w), note for example, S^w is given by

$$S^{w} = \begin{pmatrix} S_1^{w} & 0 & ... & 0 \\ 0 & S_2^{w} & ... & 0 \\ . & . & . & . \\ 0 & 0 & . & S_G^{w} \end{pmatrix}_{(GN \times GNM)} \quad \text{where } S_g^{w} = \begin{pmatrix} \mathbf{s}_{1g}^{w\,\prime} & 0 & ... & 0 \\ 0 & \mathbf{s}_{2g}^{w\,\prime} & ... & 0 \\ . & . & . & . \\ 0 & 0 & . & \mathbf{s}_{Ng}^{w\,\prime} \end{pmatrix}_{(N \times NM)} \quad \text{and } \mathbf{s}_{ng}^{w} = \begin{pmatrix} s_{ng1}^{w} \\ s_{ng2}^{w} \\ . \\ s_{ngM}^{w} \end{pmatrix}_{(M \times 1)}$$

In (9), the matrix $S_{(-)}^{\pi}$ defines the supports for the matrix $\mathrm{diag}\left(\Pi_{(-1)},...,\Pi_{(-G)}\right)$.

Consistency and Asymptotic Normality

Regularity Conditions. To derive consistency and asymptotic normality results for the SGME estimator, we assume the following regularity conditions.

R1. The N rows of the (N x G) disturbance matrix Ε are independent random drawings from an G-dimensional population with zero mean vector and unknown but finite covariance matrix Σ.

R2. The (N x K) matrix X of predetermined variables has rank K and consists of nonstochastic elements. And, $\lim \frac{1}{N} X'X = \Omega$ as $N \to \infty$ where Ω is a p.d. matrix.

R3. The elements $\nu_{ng} = \mu_{ng}$ (n=1,...,N, g=1,...,G) of the vector $\boldsymbol{\mu}_g$ are independent and bounded such that $c_{g1} + \omega_g \leq \mu_{ng} \leq c_{gM} - \omega_g$ for some ω_g>0 and large enough positive $c_{gM} = -c_{g1}$. The pdf of $\boldsymbol{\mu}$ is assumed to be symmetric about the origin with a finite covariance matrix.

R4. $\pi_{kg} \in (\pi_{kgL}, \pi_{kgH})$, for finite π_{kgL} and π_{kgH}, $\forall$ k =1,...,K and g =1,...,G.
$\gamma_{jg} \in (\gamma_{jgL}, \gamma_{jgH})$, for finite γ_{jgL} and γ_{jgH}, $\forall$ ($j \neq g$) j,g =1,...,G; γ_{gg}=-1.
$\beta_{kg} \in (\beta_{kgL}, \beta_{kgH})$, for finite β_{kgL} and β_{kgH}, $\forall$ k =1,...,K and g=1,...,G.

R5. For the true Β, and nonsingular Γ there exists a p.d. matrices Ψ_g (g=1,...,G) ∋ $\lim_{N \to \infty} \frac{1}{N} Z_g' Z_g \to \Psi_g$ where $\Pi = \mathrm{B}\Gamma^{-1}$.

Below the theorems for consistency and asymptotic normality of the structural parameters

$\hat{\delta} = \text{vec}(\hat{\delta}_1,...,\hat{\delta}_G)$ are stated. Proofs are available from the authors upon request.

Theorem 1. *Under the regularity conditions R1-R5, the SGME estimator,* $\hat{\theta} = \text{vec}(\hat{\pi},\hat{\delta})$*, is a consistent estimator of the true coefficient values* $\theta=\text{vec}(\pi,\delta)$.

Theorem 2. *Under the conditions of Theorem 1, the SGME estimator,* $\hat{\delta} = \text{vec}(\hat{\delta}_1,...,\hat{\delta}_G)$*, is asymptotically normally distributed as* $\hat{\delta} \overset{a}{\sim} N(\delta, \frac{1}{N}\Omega_\xi^{-1}\Omega_\Sigma\Omega_\xi^{-1})$.

The elements of the asymptotic covariance matrix are $\Omega_\xi = \text{diag}\left(\xi_1\Psi_1,...,\xi_G\Psi_G\right)$ where $\xi_{ng}^w(\tau) = \frac{\partial\lambda^w(u_{ng}(\tau))}{\partial u_{ng}(\tau)} = \left(\sum_{m=1}^{M}(s_{ngm}^w)^2 w_{ngm}(\lambda^w(u_{ng}(\tau))) - (u_{ng}(\tau))^2\right)^{-1}$ and λ^w is the lagrangain multiplier with respect to the *n*th observation and *g*th structural equation. In addition $\frac{1}{N}Z(\Sigma_\lambda \otimes I)Z' \to \Omega_\Sigma$, where $Z = \text{diag}\left(Z_1',...,Z_G'\right)$ and Σ_λ is a (*G* x *G*) matrix of the covariances for the $\lambda_{ng}^{w\,\prime}$s .

Given the SGME estimator $\hat{\delta}$ is consistent and asymptotically normally distributed, then asymptotically valid normal and χ^2 test statistics can be used to test hypothesis about δ. For empirical implementation a consistent estimate of the asymptotic covariance of $\hat{\delta}$, or $\Omega_\xi^{-1}\Omega_\Sigma\Omega_\xi^{-1}$, is required. First, we define

$$\hat{\xi}_g = \frac{1}{N}\sum_{n=1}^{N}\frac{1}{\sum_{m=1}^{M}(s_{ngm}^w)^2 w_{ngm}(\lambda^w(u_{ng}(\hat{\delta})) - u_{ng}(\hat{\delta})^2} \text{ and } \hat{\Psi}_g = \frac{1}{N}Z_g'Z_g,$$

then $\hat{\Omega}_\xi = \text{diag}\left(\hat{\xi}_1\hat{\Psi}_1,...,\hat{\xi}_G\hat{\Psi}_G\right)$. Second, a straightforward estimate of Ω_Σ can be constructed using an estimate of the (*G* x *G*) matrix Σ_λ based on $\hat{\varsigma}_{ij} = \frac{1}{N}\lambda^w(\boldsymbol{u}_{\bullet i}(\hat{\delta}))'\lambda^w(\boldsymbol{u}_{\bullet j}(\hat{\delta}))$ for *i*,*j*=1,...,*G*. Finally, we define the estimated asymptotic covariance matrix as $\hat{\text{Cov}}(\hat{\delta}) = \frac{1}{N}\hat{\Omega}_\xi^{-1}\hat{\Omega}_\Sigma\hat{\Omega}_\xi^{-1}$.

Asymptotically Normal Test. Since $Z = \frac{\hat{\delta}_{ij} - \delta_{ij}^0}{\sqrt{\hat{\text{Var}}(\hat{\delta})_{ii}}}$ is asymptotically *N*(0,1) under the null hypothesis H_0: $\delta_{ij} = \delta_{ij}^0$, the statistic *Z* can be used to test hypothesis about the values of the δ_{ij}'s.

Wald Test. Let $H_0: R(\delta) = [0]$ be the null hypothesis to be tested. Here $R(\delta)$ is a continuously differentiable L-dimensional vector function with rank$\left(\frac{\partial R(\delta)}{\partial\delta}\right)=L\leq K$. For a linear null hypothesis $R(\delta)=R\delta$. The Wald statistic has a χ^2 limiting distribution with L degrees of freedom, as

$$W=(R(\hat{\delta})-r)^{\prime}\left(\frac{\partial R(\hat{\delta})}{\partial\delta}^{\prime}\hat{\text{V}}\text{ar}(\hat{\delta})\frac{\partial R(\hat{\delta})}{\partial\delta}\right)^{-1}(R(\hat{\delta})-r)\overset{d}{\rightarrow}\chi^2_L$$ under the null hypothesis.

Lagrange Multiplier Test. Let δ_R be a restricted SGME estimator of δ and $\hat{\delta}_{\boldsymbol{R}}=\underset{\tau:R(\delta)=r}{argmax}\{F(\tau)\}$.

$$LM=\nabla(\hat{\delta}_{\boldsymbol{R}})^{\prime}(\hat{\text{V}}\text{ar}(\hat{\delta}))\nabla(\hat{\delta}_{\boldsymbol{R}})\overset{d}{\rightarrow}\chi^2_L$$ under the null hypothesis H_0: $R(\delta)=[0]$.

Above $F(\tau)$ is the conditionally-maximized entropy function that is a solution to (8)-(11).

Monte Carlo Experiments

In this sampling experiment we set up an overdetermined simultaneous systems model that corresponds to the empirical model discussed in Golan and Judge (1997) in order to directly compare the results of the alternative GME estimators. This system was introduced by Tsurumi (1990) and is a modification of the model used by Cragg (1967). The Γ, Β and covariance matrices are

$$\boldsymbol{\Gamma}=\begin{pmatrix}-1 & .267 & .087\\ .222 & -1 & 0\\ 0 & .046 & -1\end{pmatrix}\quad \boldsymbol{B}=\begin{pmatrix}6.2 & 4.4 & 4.0\\ 0 & .74 & 0\\ .7 & 0 & .53\\ 0 & 0 & .11\\ .96 & .13 & 0\\ 0 & 0 & .56\\ .06 & 0 & 0\end{pmatrix}\quad \boldsymbol{\Sigma}=\begin{pmatrix}1 & -1 & -.125\\ -1 & 4 & .0625\\ -.125 & .0625 & 8\end{pmatrix}$$

The exogenous variables are drawn from a N(0,1) distribution, while the errors for the structural equations were drawn from a truncated multivariate normal with mean zero and covariance $\sum\otimes I$. The support spaces specified for the structural and reduced form models are $s_{ik}^{\beta}=s_{ik}^{\pi}=[-5,0,5]^{\prime}$ for k=2,...,7; $s_{ik}^{\beta}=s_{ik}^{\pi}=[-20,0,20]^{\prime}$ for k=1; and $s_{in}^{\gamma}=[-2,0,2]^{\prime}$ for i,n=1,2,3. The error supports for the reduced form and structural model are specified as $d_{in}=v_{in}=[-\omega_i-3\sigma_i,0,3\sigma_i+\omega_i]$, where σ_i is the standard deviation of the errors from the *i*th equation and from R3 we let $\omega_i=2.5$.

Results

Results of the Γ parameters from sampling experiment are presented in Table 1. The precision of the parameter estimate and the reduction of mean square error (MSE) is apparent over the traditional 2SLS and 3SLS estimates (with true covariance matrix) in small samples. In the tables below notice that the MSE are reported in the parentheses below the coefficient estimate.

From Table 1 we can infer several implications. First, the standard GME estimates are not converging to the true values, suggesting that standard GME estimator - like the OLS estimator - is not consistent when the $Y_{(-i)}$ are correlated with the error vector $\mathbf{e}_i$. Second, as the sample size increases the 2SLS, 3SLS, and the SGME estimates are converging to the true coefficient values. Third, predominately the SGME has the lowest MSE among the consistent estimators. For small samples of 20 observations, the MSE performance of the SGME estimator is vastly improved relative to the 3SLS estimator with the true covariance structure. As the sample size increases from 20 to 100 to 200 observations, the MSE of the 3SLS estimator approaches that of the SGME estimator.

Conclusions

In this paper we proposed a data constrained structural generalized maximum entropy estimator, or SGME model, for the linear simultaneous equations statistical model with contemporaneous covariance in the error terms. We have shown that the SGME estimator is consistent and asymptotically normal under the assumed regulatory conditions. Moreover, asymptotically distributed test statistics were derived that are capable of preforming all of the hypothesis tests typically used in applied econometrics.

Monte Carlo results indicate that like ordinary least squares the standard data constrained generalized maximum entropy estimator, or GME model, is not a consistent estimator for the SESM. This is because the correlation between endogenous variables and the errors are not accounted for in either model formulation. In regard to the other estimators of the linear SESM, the Monte Carlo simulations suggest that in larger sample situations, the SGME estimator is comparable to traditional estimators. In contrast, small sample simulations indicated that the SGME is likely to be a more efficient estimator than 2SLS or 3SLS.

Table 1. Monte Carlo results from 1000 repetitions using GME, 2SLS, 3SLS(Σ), and SGME. Structural parameter γ_{12} with true value .267.

Obs	GME	2SLS	3SLS(Σ)	SGME
20	.073 (.054)	.259 (.976)	.403 (.948)	.303 (.025)
100	-.064 (.006)	.277 (.157)	.311 (.141)	.287 (.034)
200	-.151 (.182)	.255 (.096)	.270 (.083)	.270 (.033)
400	----	----	----	.270 (.024)
Structural parameter γ_{13} with true value .087.				
20	.107 (.045)	.064 (.506)	.067 (.506)	.205 (.052)
100	.046 (.039)	.082 (.078)	.083 (.078)	.139 (.041)
200	.009 (.030)	.090 (.037)	.090 (.037)	.122 (.027)
400	----	----	----	.106 (.015)
Structural parameter γ_{21} with true value .222.				
20	.078 (.034)	.134 (.097)	.200 (.089)	.317 (.036)
100	-.053 (.082)	.211 (.018)	.229 (.017)	.264 (.016)
200	-.112 (.116)	.219 (.009)	.229 (.009)	.247 (.009)
400	----	----	----	.235 (.004)
Structural parameter γ_{32} with true value .046.				
20	.068 (.014)	.050 (.330)	.013 (.324)	.146 (.078)
100	.053 (.006)	.041 (.125)	.027 (.100)	.086 (.050)
200	.048 (.003)	.058 (.083)	.051 (.061)	.075 (.040)
400	----	----	----	.065 (.024)

References

Cragg, J. G. 1967. "On the Relative Small Sample Properties of Several Structural-Equation Estimators." *Econometrica*, 35.

Donoho, D. L., I. M. Johnstone, J. C. Hoch, and A. S. Stern (1992). "Maximum Entropy and the Nearly Black Object." *Journal of the Royal Statistical Society B*, 54, 41-81.

Golan, A., G. Judge, and D. Miller. 1997. Information Recovery in Simultaneous Equations Statitical Models. In *Handbook of Applied Economic Statistics*, edited by Aman Ullah and David E. A. Giles. Marcel Dekker, New York.

Golan, A., G. Judge, and D. Miller. 1996. *Maximum Entropy Econometrics: Robust Estimation with Limited Data*. John Wiley & Sons, New York.

Mittelhammer, R. C. and N. S. Cardell. 1998. "The Data-Constrained GME Estimator of the GLM: Asymptotic Theory and Inference." Mimeo, Washington State University.

Newey, W. K. and D. L. McFadden. 1994. "Large Sample Estimation and Hypothesis Testing", Ch. 36 in: R. F. Engle and D. L. Mcfadden, eds., *Handbook of Econometrics, Vol 4*, New York, Elsevier.

Rao, C. R. 1973. *Linear Statistical Inference and Its Applications*. 2d. Ed. J. Wiley & Sons, New York.

Rockafellar, R. T. 1970. *Convex Analysis*. Princeton: Princeton University Press.

Shannon, C. E. 1948. "A Mathematical Theory of Communication." *Bell System Technical Journal*, 27.

Sever, M., R. C. Mittelhammer, and N.S. Cardell. 1998. " The Data-Constrained GME Estimator of the GLM Under Autocorrelated Error Terms: Asymptotic Theory and Inference." Mimeo, Washington State University.

Theil, H. 1971. *Principles of Econometrics*. New York: Wiley.

Tsurumi, H. 1990. "Comparing Bayesian and Non-Bayesian Limited Information Estimators." S. Geisser, J.S. Hodges, S.J. Press, and A. Zellner (ed.), Bayesian and Likelihood Methods in Statistics and Econometrics, North Holland Publishing, Amsterdam.