AGRICULTURAL AND RESOURCE ECONOMICS
AGRAR- UND RESSOURCENÖKONOMIK

Bayesian Bootstrap Analysis of Systems of Equations

Thomas Heckelei, Ron. C. Mittelhammer

University of Bonn
Universität Bonn

Thomas Heckelei, Ph.D., is a lecturer and research associate at the Institute for Agricultural Policy, Market research and Sociology. His current research areas are quantitative agricultural sector modelling and econometric methodology.

Mailing address: Institut für Agrarpolitik, Universität Bonn, Nußallee 21,
                     53115 Bonn, Germany
Phone: +49-228-732323
Fax: +49-228-9822923
E-mail: heckelei@agp.uni-bonn.de

Ron C. Mittelhammer, Ph.D, is Professor at the Department of Agricultural Economics and the Program in Statistics of Washington State University. He researches new statistical procedures for applied econometric analysis and constructs and analyzes econometric models for simulating and analyzing demand, supply, and price formation in agricultural commodity markets.

Mailing address: Department of Agricultural Economics, Washington State University,
                     Pullman, WA 99164, USA
Phone: +509-335-1706
Fax: +509-335-1173
E-mail: mittelha@coopext.cahe.wsu.edu

Bayesian Bootstrap Analysis of Systems of Equations

1 Introduction

This paper extends the Bayesian Bootstrap Regression (BBR) procedure developed and evaluated by Heckelei and Mittelhammer (1996) to a Bayesian Bootstrap Multivariate Regression (BBMR) framework that allows robust Bayesian analysis of traditional multivariate regression models. The application of 2SLS- and 3SLS-mappings to posterior distributions of reduced form coefficients derived via BBMR further allows robust Bayesian analysis of simultaneous equation systems. The BBMR approach is easily automated, widely applicable, robust with respect to the likelihood function, and flexible with respect to the choice of prior distribution. These characteristics remove impediments and contribute to a wider use of Bayesian techniques in applied econometrics research.

BBMR uses a form of Monte Carlo Integration (MCI) in order to analyze posterior distributions of model parameters. In this regard it is in the line of work by Kloek and van Dijk (1978), van Dijk and Kloek (1980), Zellner, Bauwens and van Dijk (1988), and Geweke (1989 and 1991). With continuously increasing computing power, MCI is a convincing solution to the problem of analytical tractability in multidimensional integration problems as is often encountered in Bayesian analysis of econometric models. It allows a totally flexible choice of prior distributions and can be implemented as a generic algorithm in standard statistical software independently of the actual choice of prior distribution (Geweke 1991).

In addition, BBMR does not require the specification of a parametric family for the likelihood function. Instead, it uses a bootstrapped likelihood based on the joint sampling distribution of location and scale estimators. The outcomes of this bootstrap procedure serve as outcomes of an importance function for the MCI evaluation of posterior expectations based on an importance sampling scheme (see Geweke 1989 and 1991 for importance sampling in the case of normality).

The paper is organized as follows: First, a short introduction to Bayesian analysis of the traditional multivariate regression model is given, followed by an outline of a possible MCI analysis of the posterior under normality. The presentation serves to establish notation and to set the inferential context for the BBMR approach. Then, the BBMR-algorithm is introduced as a robust alternative for performing posterior inference. The next section demonstrates how BBMR can be applied to perform Bayesian analysis of simultaneous equation models based on a generalized version of "unrestricted reduced form mappings" by Zellner, Bauwens and van Dijk (1988). Simulation results are presented in the penultimate section to assess the performance of the BBMR in the simultaneous equations context using the "Klein Model I" from Theil (1971). The final section presents conclusions and areas in need of further research.

2   Analytical Bayesian Analysis of the Traditional Multivariate Regression Model

We begin with a general analytical approach for performing Bayesian analysis of the traditional multivariate regression model. We deviate from standard textbook expositions (e.g. Zellner, 1971) by not assuming normality and arranging terms differently in order to provide better motivation for the BBMR procedure introduced in the next section. Since later sections will utilize multivariate regression formulations in the context of analyzing reduced forms of simultaneous equations systems, we use the conventional notation for the reduced form. Let

$$\mathbf{Y} = \mathbf{X}\mathbf{P} + \mathbf{V}, \tag{1}$$

where $\mathbf{Y}$ is a (n×m) matrix of observations on m endogenous variables, $\mathbf{X}$ is a (n×k) matrix of observations on k exogenous variables, $\mathbf{P}$ is a (k×m) matrix of regression coefficients, and $\mathbf{V}$ is a (n×m) matrix representing n iid outcomes of a 1×m disturbance vector having a joint density function g([**0**],$\mathbf{S}$), mean vector [**0**], and covariance matrix $\mathbf{S}$. Then $\mathbf{Y}$ has some corresponding probability density function f($\mathbf{Y}|\mathbf{P},\mathbf{S}$). Assuming that a prior probability density ("prior") on the model parameters, p($\mathbf{P},\mathbf{S}$), is independent regarding $\mathbf{P}$ and $\mathbf{S}$ such that p($\mathbf{P},\mathbf{S}$) = p($\mathbf{P}$) p($\mathbf{S}$) and letting

$L(\mathbf{P},\mathbf{S}|\mathbf{Y}) \equiv f(\mathbf{Y}|\mathbf{P},\mathbf{S})$ denote the likelihood function, the joint posterior probability density ("posterior") of the model parameters can be defined by Bayes theorem as

$$h(\mathbf{P}, \mathbf{S}| \mathbf{Y}) = \frac{p(\mathbf{P})p(\mathbf{S})\, f(\mathbf{Y}|\mathbf{P},\mathbf{S})}{m(\mathbf{Y})} = \frac{p(\mathbf{P})p(\mathbf{S})\, L(\mathbf{P},\mathbf{S}|\mathbf{Y})}{m(\mathbf{Y})} \tag{2}$$

$$\mu\ \ p(\mathbf{P})p(\mathbf{S})\, L(\mathbf{P},\mathbf{S}|\mathbf{Y})$$

where $m(\mathbf{Y})$ is the marginal probability density of the matrix $\mathbf{Y}$.

The main interest of the empirical analyst is most often the marginal posterior density of $\mathbf{P}$ which can be represented as

$$h(\mathbf{P}|\mathbf{Y}) \propto p(\mathbf{P}) \int_0^\infty \ldots \int_0^\infty p(\mathbf{S})\, L(\mathbf{P},\mathbf{S}|\mathbf{Y})\, d\mathbf{S} \tag{3}$$

$$\propto p(\mathbf{P})\ \frac{\int_0^\infty \ldots \int_0^\infty p(\mathbf{S})\, L(\mathbf{P},\mathbf{S}|\mathbf{Y})\, d\mathbf{S}}{\int_{-\infty}^\infty \ldots \int_{-\infty}^\infty \left[\int_0^\infty \ldots \int_0^\infty p(\mathbf{S})\, L(\mathbf{P},\mathbf{S}|\mathbf{Y})\, d\mathbf{S}\right] d\mathbf{P}}$$

$$= p(\mathbf{P})\ L_{\mathbf{S}}(\mathbf{P}|\mathbf{Y}),$$

where $L_{\mathbf{S}}(\mathbf{P}|\mathbf{Y})$ is interpreted as a marginal likelihood, derived via weighting $L(\mathbf{P},\mathbf{S}|\mathbf{Y})$ with the prior information on $\mathbf{S}$ and then normalizing to unit total mass. If an ignorance prior on $\mathbf{S}$ is employed, which will be the case for the remainder of this paper, then the representation of the posterior in (3) separates the available information on $\mathbf{P}$ into two parts, the prior information $p(\mathbf{P})$, and the information that comes from the data as interpreted through the model, $L_{\mathbf{S}}(\mathbf{P}|\mathbf{Y})$.

Expectations of various functions of $\mathbf{P}$ taken with respect to the posterior (3) are the typical measures used to summarize the information contained in the posterior. The posterior mean represents the Bayesian point estimate for the unknown true value of $\mathbf{P}$ that minimizes a quadratic loss function (Judge et al. 1988). Posterior variances and tail probabilities allow an evaluation of how precise the knowledge relating to $\mathbf{P}$ is. Posterior probabilities associated with specified subsets of the parameter space provide measures of confidence regarding the location of model parameters. Generally, posterior expectations of functions of $\mathbf{P}$, say $E[g(\mathbf{P})]$, can be represented as

$$E[g(\mathbf{P})] = \int_{-\infty}^\infty \ldots \int_{-\infty}^\infty g(\mathbf{P})\, h(\mathbf{P}|\mathbf{Y})\, d\mathbf{P} \tag{4}$$

3   Monte Carlo Integration and Importance Sampling

Monte Carlo Integration (MCI) is often the preferred method of evaluating the integrals in the preceding section, since flexibility in choosing prior densities, the use of complicated functions of $\mathbf{P}$ and/or high parameter dimensionality prohibit the use of analytical integration tools (Geweke 1989). Conceptually, Monte Carlo evaluation of expectations is straightforward. If N iid outcomes from $h(\mathbf{P}|\mathbf{Y})$ in equation (4) are available, say $\mathbf{P_i}$, i = 1,...,N and if $|E[g(\mathbf{P})]| = c < \infty$, then, by Kolmogorov's strong law of large numbers

$$\frac{1}{N} \sum_{i=1}^{N} g(\mathbf{P}_i) \overset{as}{\longrightarrow} E[g(\mathbf{P})] \tag{5}$$

where $\overset{as}{\rightarrow}$ denotes almost sure convergence. Thus for large enough N, a simple average of the outcomes of $g(\mathbf{P_i})$ provides an arbitrarily close approximation to $E[g(\mathbf{P})]$.

However, it is often the case that a flexible choice of prior and likelihood will result in a situation where random sampling from the posterior distribution of $\mathbf{P}$ is difficult or impossible. In this case, a particular variation of the MCI-technique, importance sampling, is especially useful (Geweke 1986 and 1989).   In this approach, a so-called "importance function" $I(\mathbf{P})$ is introduced, that replaces $h(\mathbf{P}|\mathbf{Y})$ as the sampling density. $I(\mathbf{P})$ must include the support of $h(\mathbf{P}|\mathbf{Y})$ and should be easy to sample from. Rewriting equation (4) as

$$E[g(\mathbf{P})] \;=\; \int_{-\infty}^{\infty}...\int_{-\infty}^{\infty} \left[\frac{g(\mathbf{P})\, h(\mathbf{P}|\mathbf{Y})}{I(\mathbf{P})} \right] I(\mathbf{P})\, d\mathbf{P} \tag{6}$$

expresses $E[g(\mathbf{P})]$ as the expectation of the bracketed term with respect to the sampling density $I(\mathbf{P})$. Again, Kolmogorov's strong law of large numbers allows $E[g(\mathbf{P})]$ to be approximated as

$$E[g(\mathbf{P})] \;=\; \frac{1}{N} \sum_{i=1}^{N} \frac{g(\mathbf{P}_i)\, h(\mathbf{P}_i|\mathbf{Y})}{I(\mathbf{P}_i)} \tag{7}$$

where $\mathbf{P_i}$, for i = 1,...,n, are iid outcomes from $I(\mathbf{P})$. The approximation can be made arbitrarily close by increasing n. The rate of convergence is determined by the variance of $g(\mathbf{P})$ as well as by the variability of the ratio of density values $h(\mathbf{P}|\mathbf{Y})/I(\mathbf{P})$, the latter being smaller the closer the two densities are in shape.

Normalizing $p(\mathbf{P})L_\Sigma(\mathbf{P}|\mathbf{Y})$ in (3) to unit mass in order to define the proper posterior density for $h(\mathbf{P}|\mathbf{Y})$ and substituting $L_\Sigma(\mathbf{P}|\mathbf{Y})$ for $I(\mathbf{P})$ in equation (6) then obtains

$$E[g(\mathbf{P})] \;=\; \int_{-\infty}^{\infty}\ldots\int_{-\infty}^{\infty}\left[\frac{g(\mathbf{P})\left[\dfrac{p(\mathbf{P})\,L_{\mathbf{s}}(\mathbf{P}|\mathbf{Y})}{\int_{-\infty}^{\infty}\ldots\int_{-\infty}^{\infty}p(\mathbf{P})\,L_{\mathbf{s}}(\mathbf{P}|\mathbf{Y})d\mathbf{P}}\right]}{L_{\mathbf{s}}(\mathbf{P}|\mathbf{Y})}\right]L_{\mathbf{s}}(\mathbf{P}|\mathbf{Y})\,d\mathbf{P}$$

$$=\;\int_{-\infty}^{\infty}\ldots\int_{-\infty}^{\infty}\left[\frac{g(\mathbf{P})\,p(\mathbf{P})}{\int_{-\infty}^{\infty}\ldots\int_{-\infty}^{\infty}p(\mathbf{P})\,L_{\mathbf{s}}(\mathbf{P}|\mathbf{Y})\,d\mathbf{P}}\right]L_{\mathbf{s}}(\mathbf{P}|\mathbf{Y})\,d\mathbf{P} \tag{8}$$

$$=\;\frac{\int_{-\infty}^{\infty}\ldots\int_{-\infty}^{\infty}g(\mathbf{P})\,p(\mathbf{P})\,L_{\mathbf{s}}(\mathbf{P}|\mathbf{Y})\,d\mathbf{P}}{\int_{-\infty}^{\infty}\ldots\int_{-\infty}^{\infty}p(\mathbf{P})\,L_{\mathbf{s}}(\mathbf{P}|\mathbf{Y})\,d\mathbf{P}}\;\;.$$

Note that both the numerator and denominator of the last expression in (9) can be interpreted as expectations taken with respect to $L_{\Sigma}(\mathbf{P}|\mathbf{Y})$. In particular, adopting this interpretation of the expectation operator, the numerator can be replaced by $E[g(\mathbf{P})p(\mathbf{P})]$ and the denominator replaced by $E[p(\mathbf{P})]$. Therefore, letting $\mathbf{P}_i$, $i=1,\ldots N$, be iid outcomes from $L_{\Sigma}(\mathbf{P}|\mathbf{Y})$, $E[g(\mathbf{P})]$ can be approximated by a prior-weighted average of the form

$$E[g(\mathbf{P})]\;=\;\frac{\dfrac{1}{N}\sum_{i=1}^{N}g(\mathbf{P}_i)\,p(\mathbf{P}_i)}{\dfrac{1}{N}\sum_{i=1}^{N}p(\mathbf{P}_i)}\;=\;\frac{\sum_{i=1}^{N}g(\mathbf{P}_i)\,p(\mathbf{P}_i)}{\sum_{i=1}^{N}p(\mathbf{P}_i)}\;\;. \tag{9}$$

with $\mathbf{P}_i$ being iid outcomes from $L_{\Sigma}(\mathbf{P}|\mathbf{Y})$, $i=1,\ldots,N$. In case of an ignorance prior on $\mathbf{P}$ ($p(\mathbf{P})=$ constant) it is clear that estimated expectations generated via (9) reduce to a simple average of functions of the outcomes $\mathbf{A}_i$ (see also Zellner, Bauwens, and van Dijk 1988, p. 47).

If a normal likelihood function is assumed then $L_{\Sigma}(\mathbf{P}|\mathbf{Y})$, defined in equation (3), is a matrix student-t distribution denoted as $T(v,\mathbf{P},\mathbf{S})$, where $v=n\text{-}k$, and , i.e.

$$L_{\mathbf{s}}(\mathbf{P}|\mathbf{Y})\;\propto\;|S+(\mathbf{P}-\hat{\mathbf{P}})X'X(\mathbf{P}-\hat{\mathbf{P}})|^{\frac{(v+k)}{2}} \tag{10}$$

Outcomes of this distribution can be easily generated (an algorithm is given, for example, in the appendix of Zellner, Bauwens and van Dijk 1988), so that it can serve as the importance function for the evaluation of posterior expectations.

We can conclude that for the normal multivariate regression model a general and easily implementable approach for analyzing posterior distributions of the regression coefficients is available. However, leaving the realm of normality requires the development of some other (importance) sampling approach that may or may not be as tractable as the preceding one. The need

to develop methodological solutions on a case-by-case basis is unappealing and certainly impedes the use of Bayesian techniques in econometric analyses. Furthermore, the error distribution family is usually unknown and requiring that one be chosen most likely results in a model specification error to some extent. As in Heckelei and Mittelhammer (1996) for the single equation case, we suggest an approach, BBMR, that is robust with respect to the underlying error distribution but preserves the advantages of the normal distribution-MCI approach regarding flexibility in choosing the prior and the possibility of being implemented generically, once and for all, in standard statistical software.

4 Bayesian Bootstrap Multivariate Regression

In order to obtain an approach to Bayesian analysis of the multivariate regression model that is robust with respect to the underlying probability model we first substitute for $L_\mathbf{s}(\mathbf{P}|\mathbf{Y})$ a "Multivariate Regression-Structure Likelihood," $L_\mathbf{s}(\mathbf{P}|\hat{\Pi},\mathbf{S})$, conditioned on the joint outcome of location and scale estimators representing information contained in the data. We then approximate this likelihood by an empirical likelihood defined via a bootstrap procedure that simultaneously provides a random sample from the empirical likelihood and replaces outcomes of $L_\mathbf{s}(\mathbf{P}|\mathbf{Y})$ in (10) to approximate posterior expectations.

4.1 Multivariate Regression-Structure Likelihood

Analogous to $L_\mathbf{s}(\mathbf{P}|\mathbf{Y})$ in equation (3) we define the normalized (to unit mass) Multivariate Regression-Structure-Likelihood of $\mathbf{P}$ as

$$L_\mathbf{s}(\mathbf{P}|\hat{\mathbf{P}},\mathbf{S}) = \frac{\int_0^\infty \ldots \int_0^\infty p(\mathbf{S})\, L(\mathbf{P},\mathbf{S}|\hat{\mathbf{P}},\mathbf{S})\, d\mathbf{S}}{\int_{-\infty}^\infty \ldots \int_{-\infty}^\infty \left[\int_0^\infty \ldots \int_0^\infty p(\mathbf{S})\, L(\mathbf{P},\mathbf{S}|\hat{\mathbf{P}},\mathbf{S})\, d\mathbf{S}\right] d\mathbf{P}} \tag{11}$$

The likelihood function $L(\mathbf{P},\mathbf{S}|\hat{\mathbf{P}},\mathbf{S})$ in (11) is conditioned on the usual least squares estimators of the parameters $(\mathbf{P},\mathbf{S})$ that relate to the structure of the multivariate regression problem (1), hence the name "Multivariate Regression-Structure-Likelihood." The idea of constructing a likelihood conditional on estimators of unknown parameters has been used previously in related contexts by Boos and Monahan (1986), Doksum and Lo (1990), Davison, Hinkley, and Worton (1992), and

Pettit (1982, 1983), among others, and in the same context for the univariate linear model by Heckelei and Mittelhammer (1996). Monahan and Boos (1992) have presented criteria that can be used to determine when such likelihoods are defensible via the probability calculus for Bayesian posterior inference in the case where proper prior densities are used. Essentially, such likelihoods for Θ are defensible when they are based on the densities of statistics that are parameterized entirely by Θ.

It is apparent that in the current context information contained in the data is now exclusively represented via the information contained in the estimators $\hat{\mathbf{P}}$ and $\mathbf{S}$. In the case where $\hat{\mathbf{P}}$ and $\mathbf{S}$ are sufficient statistics for $(\mathbf{P},\mathbf{S})$, as under normality, $L_{\mathbf{S}}(\mathbf{P}|\mathbf{Y})$ and $L_{\mathbf{S}}(\mathbf{P}|\hat{\mathbf{P}},\mathbf{S})$ are informationally identical and in fact are both equal to the aforementioned matrix student-t distribution. Whenever $\hat{\Pi}$ and $\mathbf{S}$ are not sufficient statistics the use of $L_{\mathbf{S}}(\mathbf{P}|\hat{\mathbf{P}},\mathbf{S})$ leads to some loss of information on the parameter vector $\mathbf{P}$. However, information loss is to some extent inevitable in empirical analyses, being completely avoided only in cases of perfect knowledge regarding the form of the underlying error distribution. Moreover, the common assumption of normally-distributed errors also represents all data information via the statistics $\hat{\mathbf{P}}$ and $\mathbf{S}$, and thus precipitates information loss, as well as constitutes a specification error, when normality does not hold.

4.2 Mixing Algorithm for Likelihood Sampling

Analogous to the analysis of the multivariate regression model under normality, one can approximate posterior expectations using (9) if a random sample from the Multivariate Regression-Structure-Likelihood, $L_{\mathbf{S}}(\mathbf{P}|\hat{\mathbf{P}},\mathbf{S})$, can be drawn. In order to define a sampling algorithm we express $L_{\mathbf{S}}(\mathbf{P}|\hat{\mathbf{P}},\mathbf{S})$ in terms of a mixed distribution involving the marginal posterior distribution of $\mathbf{S}$ (which does the mixing) and the marginal likelihood function of $\mathbf{P}$, conditional on $\mathbf{S}$ (which is mixed over $\mathbf{S}$):

$$
\begin{aligned}
L_{\mathbf{S}}(\mathbf{P}|\hat{\mathbf{P}},\mathbf{S}) \;&\propto\; \int_0^\infty \ldots \int_0^\infty p(\mathbf{S})\, L(\mathbf{P},\mathbf{S}|\hat{\mathbf{P}},\mathbf{S})\, d\mathbf{S} \\
&=\; \int_0^\infty \ldots \int_0^\infty p(\mathbf{S})\, L(\mathbf{P}|\hat{\mathbf{P}},\mathbf{S},\mathbf{S})\, L(\mathbf{S}|\hat{\mathbf{P}},\mathbf{S})\, d\mathbf{S} \\
&\propto\; \int_0^\infty \ldots \int_0^\infty L(\mathbf{P}|\hat{\mathbf{P}},\mathbf{S},\mathbf{S})\, h(\mathbf{S}|\hat{\mathbf{P}},\mathbf{S})\, d\mathbf{S},
\end{aligned}
$$

(12)

where $h(\mathbf{S}|\hat{\mathbf{P}},S) \propto p(\mathbf{S})\,L(\mathbf{S}|\hat{\mathbf{P}},S)$ is the marginal posterior of $\mathbf{S}$. It follows from this representation that a random outcome of $I_{\mathbf{S}}(\mathbf{P}|\hat{\mathbf{P}},S)$ can be obtained by first drawing a random outcome of $\mathbf{S}$ from $h(\mathbf{S}|\hat{\mathbf{P}},S)$, say $\mathbf{S}_*$, and then drawing a random outcome of $\mathbf{P}$ from $L(\mathbf{P}|\hat{\mathbf{P}},S,\mathbf{S}_*)$ (normalized to unit mass).

In order to operationalize this mixing algorithm, we first show how it relates to the sampling distributions of $\hat{\mathbf{P}}$ and $\mathbf{S}$. We consider a slightly modified version of model (1):

$$\mathbf{Y} = \mathbf{X}\mathbf{P} + \mathbf{U}\mathbf{S}^{1/2}, \tag{1'}$$

where the rows of the $(n \times m)$ matrix of errors, $\mathbf{U}$, are iid outcomes from $g([\mathbf{0}],\mathbf{I})$ having a mean of $[\mathbf{0}]$ and a covariance matrix of $\mathbf{I}$, the density of $\mathbf{V[i,.]} = \mathbf{U[i,.]}\mathbf{Q}$ is $g([\mathbf{0}],\mathbf{Q'Q})$ for any conformable $\mathbf{Q}$ with full column rank, and the $(m \times m)$ scale parameter matrix $\mathbf{S}^{1/2}$ is such that $\mathbf{U[i,.]}\mathbf{S}^{1/2} = \mathbf{V[i,.]} \sim g([\mathbf{0}],\mathbf{S})$ $\forall i$, so that the exponent $(1/2)$ denotes the calculation of a "matrix square root" fulfilling this condition. Everything else is defined as in (1). Note that the foregoing distributional assumptions characterize the class of error density families for which the ensuing mixing algorithm will be robust. In particular, when $m=1$, the robust class includes all symmetric density families with mean zero, as well as any family of skewed densities that can be defined via scaling of a random variable having a parameterless density with mean zero. For $m \geq 2$, all density families in the elliptically contoured class having mean vector $[\mathbf{0}]$ and covariance matrix Σ are in the robust class, including families such as Pearson II, Pearson VII, multivariate T, LaPlace, Bessel, Uniform (elliptical), and multivariate normal (Johnson, 1987, chapter 6; Johnson and Kotz, 1972, p. 297). This follows straightforwardly from the fact that the characteristic function of an elliptically contoured random vector V_i with mean vector $[\mathbf{0}]$ and covariance matrix Σ is given by $\phi_{v_i}(t) = \psi(ct'\mathbf{S}t)$ for some function $\Psi(\cdot)$, where c is a known numerical constant that is specific to a density family (Cambanis, et. al, 1981, p. 368 and Theorem 4).

Given the preceding assumptions, the outcomes of the estimators $\hat{\Pi}$ and $\mathbf{S}$ are distributionally equivalent, in the sense of their marginal distributions, to the outcomes of

$$\left\{ \begin{array}{l} \hat{\mathbf{P}} = \mathbf{P} + (\mathbf{X'X})^{-1}\mathbf{X'}\,\mathbf{U}\mathbf{S}^{1/2} \\[2mm] \mathbf{S} = \mathbf{T'}(\mathbf{U'MU})\mathbf{T} \end{array} \right\}, \tag{13}$$

where $\mathbf{T}$ is a matrix such that $\mathbf{T'T} = \mathbf{S}$ and $\mathbf{M} = \mathbf{I\text{-}X(X'X)^{-1}X'}$. Fixing $\hat{\mathbf{P}}$ and $\mathbf{S}$ at their observed values and treating $\mathbf{P}$ and $\mathbf{S}$ as random variables, outcomes of $\mathbf{P}$ and $\mathbf{S}$ can be solved for as

$$\left\{ \begin{array}{l} \mathbf{P} = \hat{\mathbf{P}} - (\mathbf{X'\,X})^{-1}\mathbf{X'\,US}^{1/2} \\ \mathbf{S} = \mathbf{S}^{1/2}(\mathbf{U'\,MU})^{-1}\mathbf{S}^{1/2} \end{array} \right\}. \tag{14}$$

To show that the second equation in (14) holds consider the following sequence of steps: First, pre- and postmultiply the second equation in (13) by $\mathbf{S}^{-1/2}$ to get

$$\mathbf{I} = \mathbf{S}^{-1/2}\mathbf{T'}(\mathbf{U'MU})\mathbf{TS}^{-1/2}. \tag{15}$$

It follows that $\mathbf{S}^{-1/2}\mathbf{T'} = (\mathbf{U'MU})^{-1/2}$ will satisfy (15), so that

$$\mathbf{T} = \mathbf{S}^{1/2}(\mathbf{U'MU})^{-1/2} \tag{16}$$

and therefore

$$\mathbf{S} = \mathbf{T'T} = \mathbf{S}^{1/2}(\mathbf{U'MU})^{-1}\mathbf{S}^{1/2}. \tag{17}$$

If the parametric family of the underlying error distribution were known and new samples could be drawn from it, then (14) represents an empirical recipe for implementing the mixing concept presented in (12). Specifically, one could draw a random sample of size n from g($[\mathbf{0}],\mathbf{I}$), resulting in the (n×m) matrix $\mathbf{U}_*$, and then calculate, using (14) and the given value of $\mathbf{S}$, an outcome of $\mathbf{S}$, say $\mathbf{S}_*$, which is interpretable as an outcome from h($\mathbf{S}|\hat{\mathbf{P}},\mathbf{S}$) (further motivated below). Then $\mathbf{U}_*$ can be scaled by the square root of $\mathbf{S}_*$, $\mathbf{S}_*^{1/2}$, to condition an outcome of $\mathbf{P} = \hat{\mathbf{P}} - (\mathbf{X'X})^{-1}\mathbf{X'U}_*\mathbf{S}_*^{1/2}$ on the covariance matrix $\mathbf{S}_*$. This outcome of $\mathbf{P}$ is effectively an outcome from (the normalized to unit total mass) L($\mathbf{P}|\hat{\mathbf{P}},\mathbf{S},\mathbf{S}_*$). Repeating the procedure N times produces a random sample of size N from $\mathrm{L}_\Sigma(\mathbf{P}|\hat{\mathbf{P}},\mathbf{S})$, which is equivalent to mixing L($\mathbf{P}|\hat{\mathbf{P}},\mathbf{S},\mathbf{S}$) over h($\mathbf{S}|\hat{\mathbf{P}},\mathbf{S}$) and random sampling from the mixture.

We now examine the type of marginal posterior distribution for $\mathbf{S}$ that is implied by the transformation in (17) and find that it incorporates the standard ignorance prior on Σ that is typically used in Bayesian analyses of the multivariate regression model. Under the conditions given for the distributional model in (1') and denoting the sampling distribution of $\mathbf{S}$ as f($\mathbf{S}|\mathbf{S}$), we have

$$\mathbf{S} = \mathbf{T'}(\mathbf{U'\,MU})\mathbf{T} \sim \mathrm{f}(\mathbf{S}|\mathbf{S}) = \mathrm{f}(\mathbf{S}|\mathbf{T}) \tag{18}$$

since the distribution of $\mathbf{UT}$ depends only on the value of $\mathbf{T'T} = \mathbf{S}$. Letting $\mathbf{W} = \mathbf{U'MU} = (\mathbf{T'})^{-1}\mathbf{ST}^{-1}$ and noting that $|\partial\mathbf{W}/\partial\mathbf{S}| = |\mathbf{T}|^{-(m+1)}$, because of the symmetry of $\mathbf{S}$ (Deemer and Olkin, 1951) the distribution of $\mathbf{W}$ is given via change of variables as

$$\eta(\mathbf{W}) = f(\mathbf{T'WT}|\mathbf{T})|\mathbf{T}|^{m+1}. \tag{19}$$

For a given $\mathbf{S}$ it follows from (17) that $\mathbf{W} = \mathbf{S}^{1/2}\mathbf{S}^{-1}\mathbf{S}^{1/2}$ and $|\partial\mathbf{W}/\partial\mathbf{S}| = |\mathbf{S}|^{(m+1)/2}|\mathbf{S}|^{-(m+1)}$ (Deemer and Olkin, 1951). The change of variable transformation from the distribution of $\mathbf{W}$ to the distribution of $\mathbf{S}$ then yields

$$\begin{aligned}
h(\mathbf{S}|\mathbf{S}) &\propto f(\mathbf{T'S}^{1/2}\mathbf{S}^{-1}\mathbf{S}^{1/2}\mathbf{T}|\mathbf{T})|\mathbf{T}|^{m+1}|\mathbf{S}|^{-(m+1)} \\
&\propto f(\mathbf{S}|\mathbf{S})|\mathbf{S}|^{(m+1)/2}|\mathbf{S}|^{-(m+1)} = f(\mathbf{S}|\mathbf{S})|\mathbf{S}|^{-(m+1)/2}
\end{aligned} \tag{20}$$

since $|\mathbf{T}| = |\mathbf{T'}| = |\mathbf{T'T}|^{1/2} = |\mathbf{S}|^{1/2}$, which implies that the regression structure likelihood for $\mathbf{S}$, represented by $f(\mathbf{S}|\mathbf{S})$, is postmultiplied by the standard ignorance prior $p(\mathbf{S}) \propto |\mathbf{S}|^{-(m+1)/2}$ to obtain the marginal posterior for $\mathbf{S}$, $h(\mathbf{S}|\hat{\mathbf{P}},\mathbf{S}) \equiv h(\mathbf{S}|\mathbf{S})$. It can be straightforwardly verified that in the multivariate normal case, (20) yields the appropriate inverted Wishart distribution.

4.3 Bootstrapping the Mixing Algorithm: The BBMR Algorithm

If the error distribution family were known, the analytical Bayesian approach were intractable, and random sampling from the error distribution family were reasonably straightforward, then the mixing algorithm, described in the preceding section together with (10), would represent a feasible method of conducting Bayesian analysis of the multivariate regression model. Of course, other computationally intensive and possibly more efficient approaches might be available for performing Bayesian inference in this case, such as MCI and importance sampling. However, the mixing approach will approximate the results of the analytical approach arbitrarily closely for large enough samples if $\hat{\mathbf{P}}$ and $\mathbf{S}$ are sufficient statistics for the parameters of the regression model. Furthermore, even if there were some information loss associated with the mixing approach, computational simplicity may favor its use.

In the more typical case where the error distribution family is unknown, a robust variation on the mixing approach can be defined using the empirical distribution function of the regression

residuals, $EDF(\hat{\mathbf{V}})$, to approximate the true error distribution, $g([\mathbf{0}],\mathbf{S})$, and ultimately to approximate $L_{\mathbf{S}}(\mathbf{P}|\hat{\mathbf{P}},\mathbf{S})$ by a nonparametric estimate $L_{\mathbf{S}*}(\mathbf{P}|\hat{\mathbf{P}},\mathbf{S})$. In this case the bootstrap is the relevant sampling technique, where new n×m residual matrices are constructed by random sampling residual vectors, with replacement, from the observed vectors of residuals in $\hat{\mathbf{V}} = \mathbf{M}\mathbf{U}\mathbf{S}^{1/2}$.

Denote a bootstrap sample of the OLS-residual vectors as $\mathbf{V}_*=\mathbf{U}_*(\mathbf{S}/n)^{1/2}$, where $\mathbf{U}_* = \mathbf{V}_*(\mathbf{S}/n)^{-1/2}$ then approximates a random sample of size n from $g([\mathbf{0}], I)$. According to (14) a bootstrapped outcome of $\mathbf{P}$, $\mathbf{P}_*$, can be generated by first computing a bootstrapped outcome of $\mathbf{S}$, $\mathbf{S}_*$ (second equation in (14)) by substituting U_* in place of U, and then substituting the square root $\Sigma_*^{1/2}$ into the first equation in (14) in place of $\mathbf{S}^{1/2}$. Considering this procedure in detail allows substitutions to be made that eliminate the need for sequential calculations of $\mathbf{S}_*$ and $\mathbf{P}_*$ as:

$$\left\{\begin{array}{l} \mathbf{P}_* = \hat{\mathbf{P}} - (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{V}_*\mathbf{S}^{-1/2}(\mathbf{S}\mathbf{S}_*^{-1}\mathbf{S})^{1/2} \\[2mm] \mathbf{S}_* = \mathbf{S}^{1/2}[(\mathbf{V}_*(\mathbf{S}/n)^{-1/2})'\mathbf{M}(\mathbf{V}_*(\mathbf{S}/n)^{-1/2})]^{-1}(\mathbf{S})^{1/2} = n^{-1}\mathbf{S}\mathbf{S}_*^{-1}\mathbf{S} \end{array}\right\} \tag{21}$$

with $\mathbf{S}_* = \mathbf{V}_{*'}\mathbf{M}\mathbf{V}_*$.

We can now define the BBMR algorithm as follows (compare to the BBR-algorithm presented by Heckelei and Mittelhammer, 1996):

1. Calculate $\hat{\mathbf{P}} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{Y}$.

2. Calculate $\hat{\mathbf{V}} = \mathbf{Y} - \mathbf{X}\hat{\mathbf{P}}$.

3. Calculate $\mathbf{S}^{1/2} = \left(\hat{\mathbf{V}}'\hat{\mathbf{V}}\right)^{1/2}$

4. Draw a random sample (with replacement) of size n from $\hat{\mathbf{V}}_1,\ldots,\hat{\mathbf{V}}_n$, with the subscripts indicating the rows of the $\hat{\mathbf{V}}$ matrix, resulting in the (n×m) matrix $\mathbf{V}_*$.

5. Calculate $\mathbf{S}_* = \left(\mathbf{V}_*'\mathbf{M}\mathbf{V}_*\right)$.

6. Calculate $\mathbf{V}_{**} = \mathbf{V}_*\mathbf{S}^{-1/2}(\mathbf{S}\mathbf{S}_*^{-1}\mathbf{S})^{1/2}$.

7. Calculate $\mathbf{P}_* = \hat{\mathbf{P}} - (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{V}_{**}$.

8. Repeat steps 4-7 N times and collect outcomes of $\mathbf{P}_*$.

9. In combination with a specification of the prior distribution, $p(\mathbf{P})$, use the N outcomes of $\mathbf{P}_*$ to approximate posterior expectations according to equation (9).

As the data sample size increases, and EDF($\hat{\mathbf{V}}$) $\rightarrow$ g([0], $\mathbf{S}$), then bootstrap sampling in the BBMR algorithm becomes equivalent to random sampling from the true error distribution. As the bootstrap sample size increases, posterior expectations calculated via the BBMR algorithm become equivalent to analytical expectations based on EDF($\hat{\mathbf{V}}$). Thus, for large enough data and bootstrap sample sizes, the BBMR algorithm will produce posterior expectations based on the true regression structure likelihood for the parameters of the regression model.

5 BBMR and Robust Bayesian Analysis of Simultaneous Equations

A considerable part of econometric modeling work involves structural equations which have endogenous variables on the right hand side, i.e. simultaneous equation systems. Zellner, Bauwens and van Dijk (1988) developed, among other things, several mappings of unrestricted reduced form coefficients that allow for Bayesian estimation and specification analysis of structural equations based on random samples from the posterior distribution of reduced form coefficients. In this section their "2SLS-Mapping" is extended to a "3SLS-Mapping" to accommodate cases where more than one structural equation is of interest. In combination with the BBMR algorithm developed above, these mappings can be used to perform robust Bayesian analysis of simultaneous equation systems.

5.1 2SLS and 3SLS Mappings of the Reduced Form

Consider the following representation of a system of m structural equations

$$\mathbf{Y}\mathbf{G} + \mathbf{X}\mathbf{B} + \mathbf{U} = [\mathbf{0}] \tag{22}$$

where $\mathbf{Y}$ is a n$\times$m matrix of endogenous variables, $\mathbf{X}$ is a (n$\times$k) matrix of predetermined variables, $\mathbf{G}$ and $\mathbf{B}$ are (m$\times$m) and (k$\times$m) matrices of coefficients of endogenous and predetermined variables, respectively, and [$\mathbf{0}$] is a n$\times$m matrix of zeros. Assume the system is normalized so that $\mathbf{G}$ has negative unit values on the diagonal. The reduced form of (22) can then be written as

$$\mathbf{Y} = \mathbf{X}\mathbf{P} + \mathbf{V} \tag{23}$$

with $\mathbf{V} = -\mathbf{U}\mathbf{G}^{-1}$ and

$$\mathbf{P} = -\mathbf{B}\mathbf{G}^{-1}. \tag{24}$$

Now let $\mathbf{G}_0 = \mathbf{G} + \mathbf{I}$, with $\mathbf{I}$ being a m×m identity matrix, and rearrange equation (24) to get

$$\mathbf{P}(\mathbf{G}_0 - \mathbf{I}) = -\mathbf{B} \iff \mathbf{P} = \mathbf{P}\mathbf{G}_0 + \mathbf{B} \tag{25}$$

Premultiplying by $\mathbf{X}$ yields a generalized version of equation 2.24 in Zellner, Bauwens and van Dijk (1988) as

$$\mathbf{X}\mathbf{P} = \mathbf{X}\mathbf{P}\mathbf{G}_0 + \mathbf{X}\mathbf{B} = \overline{Z}\delta, \tag{26}$$

where $\overline{Z} = [\mathbf{X}\mathbf{P} \ \ \mathbf{X}]$ and $\mathbf{d} = \begin{bmatrix} \mathbf{G}_0 \\ \mathbf{B} \end{bmatrix}$. To allow for possible errors in the exact restrictions implied by (26) a (n×m) discrepancy matrix $\mathbf{D}$ (instead of a vector in the case of single equation analyses) is introduced leading to

$$\mathbf{X}\mathbf{P} = \overline{Z}\delta + \mathbf{D}. \tag{27}$$

Let $\mathbf{D}_i$ be the ith column of $\mathbf{D}$, which represents discrepancies in the exact restrictions of (26) corresponding to the ith structural equation. Minimizing each of the discrepancy functions $\mathbf{D}_i'\mathbf{D}_i$, for $i = 1,...,m$, separately with respect to $\mathbf{d}_i$ (the ith column of $\mathbf{d}$) defines what Zellner, Bauwens and van Dijk call the 2SLS-mapping, because it resembles the way 2SLS-estimation maps $\mathbf{P}$ into δ, as

$$\delta_i^{2SLS} = (\overline{\mathbf{Z}}_i'\overline{\mathbf{Z}}_i)^{-1}\overline{\mathbf{Z}}_i'\mathbf{X}\mathbf{P}_i, \ i = 1,\ldots,m. \tag{28}$$

Here $\overline{\mathbf{Z}}_i$ denotes a matrix of the columns of $\mathbf{X}\mathbf{P}$ and $\mathbf{X}$ that appear in the ith equation, and $\mathbf{P}_i$ denoting the ith column of $\mathbf{P}$.

Accordingly, a 3SLS Mapping can be defined as

$$\delta^{3SLS} = (\widetilde{\mathbf{Z}}'(\widehat{\mathbf{W}}\otimes I)^{-1}\widetilde{\mathbf{Z}})^{-1} \ \widetilde{\mathbf{Z}}'(\widehat{\mathbf{W}}\otimes I)^{-1} \operatorname{vec}(\mathbf{X}\mathbf{P}) \tag{29}$$

with

$$\widetilde{\mathbf{Z}} = \begin{bmatrix} \overline{\mathbf{Z}}_1 & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \overline{\mathbf{Z}}_2 & \cdots & \mathbf{0} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \cdots & \overline{\mathbf{Z}}_m \end{bmatrix}, \qquad \widehat{\mathbf{W}} = \begin{bmatrix} \hat{\omega}_{11} & \hat{\omega}_{12} & \cdots & \hat{\omega}_{1m} \\ \hat{\omega}_{21} & \hat{\omega}_{22} & \cdots & \hat{\omega}_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ \hat{\omega}_{m1} & \hat{\omega}_{m2} & \cdots & \hat{\omega}_{mm} \end{bmatrix},$$

and $\hat{\omega}_{ij} = (X\mathbf{P}_i - \overline{\mathbf{Z}}_i \delta_i^{2SLS})'(X\mathbf{P}_j - \overline{\mathbf{Z}}_j \delta_j^{2SLS}) / n$, so that this mapping takes covariances between the different errors in (27) into account. The 3SLS-mapping, $\mathbf{d}^{3SLS}$, is the solution for $\mathbf{d}$ in (27) that minimizes the quadratic discrepancy function defined by $\text{vec}(\mathbf{D})'(\hat{\mathbf{W}} \otimes I)^{-1}\text{vec}(\mathbf{D})$.

5.2 BBMR and Robust Structural Equation Analysis

The 2SLS- and 3SLS-mappings can be combined with the BBMR algorithm to perform robust Bayesian estimation and specification analysis of structural coefficients. In the first step BBMR is used to provide outcomes from a robust ignorance prior-based posterior of the reduced form coefficients, replacing the random sample from the matrix student t-distribution in Zellner, Bauwens and van Dijk (1988). In the second step these reduced form coefficient outcomes are substituted for $\mathbf{P}_i$ and $\mathbf{P}$ in (28) and (29) to calculate outcomes of the 2SLS- and 3SLS-mappings. The mappings represent information about structural coefficients that is contained in the data and bootstrapped outcomes can be interpreted as being drawn from an approximation to the marginal likelihood function of the structural coefficients. Finally, posterior expectations of structural coefficients or functions thereof ($g(\mathbf{d})$) are evaluated using a prior ($p(\mathbf{d})$)-weighted average of the mapping outcomes analogous to equation (9), e.g. for the 3SLS-mapping as

$$\frac{\sum_{i=1}^{N} g(\delta_i^{3SLS}) \, p(\delta_i^{3SLS})}{\sum_{i=1}^{N} p(\delta_i^{3SLS})} \; . \tag{30}$$

where N denotes the bootstrap sample size.

It should be noted here that an important use of the unrestricted reduced form mappings is in analyzing the reasonableness of the identifying restrictions of the structural model. This can be accomplished through an examination of the posterior distribution of the discrepancy functions described above (see Zellner, Bauwens and van Dijk for the 2SLS-mapping case).

6 Motivation, Design, and Results of Monte Carlo Simulations

The theoretical validity of BBMR is based on the bootstrap's ability to consistently estimate the sampling distribution of $\hat{\mathbf{P}}$ and $\mathbf{S}$. As the data sample size n approaches infinity, the empirical distribution function of the observed residuals, EDF($\hat{\mathbf{V}}$), converges to the true distribution $g([\mathbf{0}]|\mathbf{S})$ and accordingly $L_{\mathbf{S}*}(\mathbf{P}|\hat{\mathbf{P}},\mathbf{S})$ converges to $L_{\mathbf{S}}(\mathbf{P}|\hat{\mathbf{P}},\mathbf{S})$. Finite sample properties of the bootstrap, however, are generally unknown. Monte Carlo simulations of the single equation regression model have been promising, even for rather small data sample sizes (see Heckelei and Mittelhammer 1996) but performance in the multivariate regression setting requires investigation. Furthermore, the suggested use of the BBMR-outcomes to analyze structural equations involves nonlinear 2SLS- and 3SLS-mappings that might render the approximation error of the bootstrap more significant. The performance of the BBMR should be evaluated within a simultaneous equations structure that is typical in applied econometrics. The design of the Monte Carlo simulations should allow an assessment of the crucial properties of a robust estimator, namely the efficiency loss relative to the normal approach when the normal probability model is true, and the ability of the procedure to accurately represent characteristics of the true posterior distribution under different probability regimes.

In order to measure the approximation error of BBMR regarding posterior expectations, we contrast means, variances, and tail probabilities of the bootstrapped marginal posterior distributions $h_*(\delta_j|\hat{\mathbf{P}},\mathbf{S})$ based on 2SLS and 3SLS-mappings with their parametric counterparts from $h(\delta_j|\hat{\mathbf{P}},\mathbf{S})$. Note that the efficiency loss resulting from the use of $\hat{\mathbf{P}}$ and $\mathbf{S}$ as the only source of data information relative to the use of a known likelihood function $L(\mathbf{P},\mathbf{S}|\mathbf{Y})$ is not evaluated with this approach when $\hat{\mathbf{P}}$ and S are not sufficient statistics. Such efficiency loss is in a sense unavoidable in empirical work since the analyst rarely knows the true functional form of the likelihood function and, consequently, it is of less practical interest in our context.

Two important sources of approximation error remain in the current simulation context: (1) A finite collection of bootstrapped error samples $\mathbf{V}_*$ from EDF() does not completely represent the full informational content of EDF($\hat{\mathbf{V}}$), so that the appropriate bootstrap sample size is an issue and (2) EDF($\hat{\mathbf{V}}$) itself is only an approximation to the true error distribution $g([\mathbf{0}]|\mathbf{S})$ raising questions

regarding required data sample sizes and also regarding the usefulness of certain transformations of the observed residuals $\hat{\mathbf{V}}$ suggested in the literature.

Heckelei and Mittelhammer (1996) paid considerable attention to techniques that have been proposed to mitigate the preceding types of approximation errors. The design of the simulations reported here are based on those findings in the single equation case and we do not reiterate all of the comparisons between different bootstrap correction techniques and residual transformations. We therefore use sampling with replacement of the untransformed OLS-residuals of the reduced form equations exactly as described in the BBMR-algorithm above. Nevertheless, we do analyze whether a system version of a promising second order correction technique, introduced by Heckelei and Mittelhammer, is useful device in the BBMR context. The technique transforms the collection of bootstrap samples in order to achieve second moment characteristics of the "infinite" bootstrap.

Letting $\mathbf{V}_{*i}$, $i = 1,...,N$, be the collection of N bootstrapped (n×m) residual matrices, the transformation is defined as

$$\mathbf{V}_{*i}^S = \mathbf{V}_{*i}\,\mathbf{P}\,\mathbf{L}^{-1/2}\,\hat{\mathbf{S}}^{1/2},\ i = 1,...,N \tag{31}$$

where P and Λ are the eigenvector and diagonal eigenvalue matrices corresponding to the matrix

$$\overline{\mathbf{Y}} = \frac{1}{N}\sum_{i=1}^{N}\mathbf{V}_{*i}'\,\mathbf{V}_{*i}, \quad \text{and} \quad \hat{\mathbf{S}}^{1/2} \quad \text{is} \quad \text{the} \quad \text{symmetric} \quad \text{square} \quad \text{root} \quad \text{of}$$

$\hat{\mathbf{S}} = n^{-1}S = (\mathbf{Y}-\mathbf{X}\hat{\mathbf{P}})'(\mathbf{Y}-\mathbf{X}\hat{\mathbf{P}})/n$. It follows that

$$\frac{1}{N}\sum_{i=1}^{N}\mathbf{V}_{*i}^S{}'\,\mathbf{V}_{*i}^S = \hat{\mathbf{S}}^{1/2}\,\mathbf{L}^{-1/2}\mathbf{P}'\left[\frac{1}{N}\sum_{i=1}^{N}\mathbf{V}_{*i}'\,\mathbf{V}_{*i}\right]\mathbf{P}\,\mathbf{L}^{-1/2}\,\hat{\mathbf{S}}^{1/2} = \hat{\mathbf{S}}, \tag{32}$$

which is the appropriate covariance matrix associated with EDF($\hat{\mathbf{V}}$) and the limiting covariance matrix that would be estimated from bootstrapped residuals if the bootstrap sample size were increased toward ∞. This technique is a more generally applicable alternative to ?Second Order Balancing? (Graham et al. 1990), and places none of the latter approach?s restrictions on bootstrap and data sample sizes.

6.1 Simulation Design

The Monte Carlo simulations are based on Klein's Model I as reported in Theil (1971). For variable definitions and additional information about the model that is not reported here, the reader can consult the reference. The simulation results are generated via the following sequence of steps:

1. A data sample of n=21 is drawn from Klein's model using the 3SLS-estimates reported by Theil as the "true" values of the model parameters, i.e. the data is drawn from a simultaneous equation system with three behavioral equations (first three equations) and three identities (last three equations) of the following form:

$$\mathbf{Y}_t\, \mathbf{G} \,=\, \mathbf{X}_t\, \mathbf{B} \,+\, \mathbf{U}_t \tag{33}$$

with

$$\mathbf{Y}_t \,=\, \left\{\, C_t \;\; I_t \;\; W_t^I \;\; X_t \;\; P_t \;\; D_t \,\right\},$$

$$\mathbf{G} = \left\{ \begin{array}{cccccc} 1 & 0 & 0 & -1 & 0 & 0 \\ 0 & 1 & 0 & -1 & 0 & -1 \\ -0.7901 & 0 & 1 & 0 & 1 & 0 \\ 0 & 0 & -0.4005 & 1 & -1 & 0 \\ -0.1249 & 0.0131 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{array} \right\},$$

$$\mathbf{X}_t \,=\, \left\{\, 1 \;\; t - 1931 \;\; W_t^G \;\; T_t \;\; G_t \;\; P_{t\text{-}1} \;\; K_{t\text{-}1} \;\; X_{t\text{-}1} \,\right\}$$

$$\mathbf{B} = \left\{ \begin{array}{cccccc} 16.44 & 28.18 & 1.8 & 0 & 0 & 0 \\ 0 & 0 & 0.1497 & 0 & 0 & 0 \\ 0.7901 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0.1631 & 0.7557 & 0 & 0 & 0 & 0 \\ 0 & 0.1948 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0.1813 & 0 & 0 & 0 \end{array} \right\}.$$

The structural errors **U** have some multivariate probability density with mean $[\mathbf{0}]$ and a covariance submatrix for the three behavioral equations (all other entries of the complete covariance matrix are zero)

$$\mathbf{W} = \begin{Bmatrix} 4.459 & 2.057 & -1.968 \\ 2.057 & 10.47 & 2.015 \\ -1.968 & 2.015 & 2.600 \end{Bmatrix}$$

which is five times the contemporaneous covariance matrix estimated from the 3SLS residuals. The additional variation was added in order to insure that any observed accuracy of the BBMR is not due to the relatively good fit of Klein's model. However, we also show simulation results based on the original smaller contemporaneous covariance matrix estimated from the 3SLS residuals for comparison with the results of the standard simulation case described here (Table 2).

The data sample was generated sequentially (because of lagged endogenous variables) using $\mathbf{Y}_t = \mathbf{X}_t \mathbf{B}\mathbf{G}^{-1} + \mathbf{U}_t \mathbf{G}^{-1}$. All reported results are restricted to the first model equation, i.e the consumption function.

2. The means, $\bar{\delta}_j$, variances, $\mathrm{Var}(\delta_j)$, and values of δ_j corresponding to the 2nd, 5th, 10th, 90th, 95th, and 98th percentile, denoted as δ_j^{ith}, are calculated from the parametric marginal posterior distributions of the structural coefficients, $h(\delta_j | \hat{\mathbf{P}}, \mathbf{S})$. For the case of normally distributed errors this is done using the sampling procedure from the appropriate matrix T-distribution as described in Zellner, Bauwens and van Dijk (1988). In the case of T-distributed errors we use a procedure equivalent to the BBMR algorithm except that the errors are drawn from the known parametric family parameterized by $\mathbf{S}$. For simplicity we employ an ignorance prior on **d** noting that the nonlinear character of the mappings of the reduced form coefficients should sufficiently challenge the ability of the BBMR to approximate multivariate distributions. All "parametric solutions" are based on 100000 error samples to minimize noise.

3. In a sequence of $n_{sim} = 50$ simulations of the BBMR procedure, the BBMR-outcomes of the posterior expectations based on a bootstrap sample size of $n_b = 1000$ (5000), are compared with their parametric counterparts. The reported distance measures for each marginal posterior $h_j(\delta_j | \hat{\mathbf{P}}, \mathbf{S})$, are calculated as follows (subscript j is suppressed below, bootstrapped outcomes have a "*" subscript):

(a) Root Mean Square Error of bootstrapped posterior mean estimate (denoted in Tables as $(MSE)^{1/2}$, Mean):

$$\sqrt{\frac{1}{n_{sim}} \sum_{i=1}^{n_{sim}} (\bar{\delta}_{*i} - \bar{\delta})^2} \;.$$

b) Bias of posterior mean estimate (Bias, Mean):

$$\frac{1}{n_{sim}} \sum_{i=1}^{n_{sim}} (\bar{\delta}_{*i} - \bar{\delta}) \;.$$

c) Root Mean Square Error of bootstrapped posterior variance estimate($(MSE)^{1/2}$, Variance):

$$\sqrt{\frac{1}{n_{sim}} \sum_{i=1}^{n_{sim}} \left[Var(\delta_{*i}) - Var(\delta) \right]^2} \;.$$

d) Bias of bootstrapped posterior variance estimate (Bias, Variance):

$$\frac{1}{n_{sim}} \sum_{i=1}^{n_{sim}} (Var(\delta_{*i}) - Var(\delta)) \;.$$

e) The average bootstrapped probabilities (ith percentiles):

$$\frac{1}{n_{sim}} \sum_{i=1}^{n_{sim}} prob_{*i} \;,$$

where $prob_{*i}$ is calculated as the proportion of the bootstrapped δ_*'s in the ith bootstrapped sample that are below δ_j^{ith}, relative to the bootstrap sample size.

f) Standard deviation of bootstrapped probabilities (STDV):

$$\sqrt{\frac{1}{n_{sim}-1} \sum_{i=1}^{n_{sim}} (prob_{*i} - \frac{1}{n_{sim}} \sum_{i=1}^{n_{sim}} prob_{*i})^2} \;.$$

A single simulation sequence evaluating the distance measures for a specific error distribution repeats steps one to three 10 times in order to make the results less dependent on a specific data sample. The results reported in each table are consequently averages over 10 different simulated data sets and required between 15 and 25 hours on a IBM-compatible PC with a Pentium-90 CPU. Note however, that the simulations involve three layers of sampling--the bootstrap itself, the n_{im} repetitions of the bootstrap, and the resampling of the data set. When BBMR is used as an estimation technique in empirical work, only bootstrap sampling will be involved. Simple Bayesian point estimates and variances of a three equation model can be obtained in less than a minute even with a bootstrap sample size of 5000.

6.2 Simulation Results

Table 1 shows distance measures for 2SLS-mappings based on the standard simulation procedure described in the previous section. The accuracy of the nonparametric mean and variance estimates as measured by the root mean square error is very promising. Since the respective bias measures are smaller than the root mean square errors by an order of magnitude, the largest part of $MSE^{1/2}$ must be variance-related. Therefore, one would expect that the general approximation accuracy of BBMR will increase for higher bootstrap sample sizes, second order corrected samples or lower error variances in the data generating process. Results presented below will partly confirm this hypothesis. Comparing these results to the single equation BBR algorithm in Heckelei and Mittelhammer (1996, Table 1), however, suggests there may be less accuracy in the multivariate case. This may reflect the higher demand put on the empirical distribution function in the multivariate setting with considerably more parameters to be approximated. Moreover, the nonlinear mappings of the reduced form coefficients may accentuate the approximation errors of the BBMR. It is somewhat surprising in this context that the average of the bootstrapped probabilities and their standard deviations do not significantly differ regarding their accuracy from the single equation results reported in Heckelei and Mittelhammer. One should keep in mind, however, that, contrary to the single equation case, averages over 10 data samples are reported here. More detailed results (not reported here) show that the accuracy of the probability estimates differ somewhat across data samples, but remain comparable in terms of order of magnitude. The standard deviations of the probabilities are rather stable across different data samples.

Table 1:   **Distance Measures Between 2SLS-Mappings Based on BBMR and Parametric Posterior - Normal Errors**

	Structural Coefficients (True Value)			
Distance Measure	δ_1(16.44)	δ_2(0.1249)	δ_3(0.1631)	δ_4(0.7901)
$MSE^{1/2}$, Mean	0.2693	0.0603	0.0591	0.0482
Bias, Mean	0.0043	-3.0E-04	0.0051	1.1E-04
$MSE^{1/2}$, Variance	0.3672	0.0014	8.3E-04	4.1E-04
Bias, Variance	0.0382	8.8E-05	-3.5E-05	5.3E-05
2% probability	1.92	1.93	2.00	2.04
(STDV)	(0.4357)	(0.4662)	(0.4292)	(0.4513)
5% "	4.94	4.89	5.03	5.05

	(0.6910)	(0.7177)	(0.6815)	(0.6611)
10% "	9.96	9.84	10.07	9.99
	(0.9628)	(0.9504)	(0.9507)	(0.9218)
90% "	89.80	89.82	90.02	89.95
	(0.9180)	(0.9921)	(0.9363)	(0.9718)
95% "	94.87	94.83	95.00	94.96
	(0.6972)	(0.7736)	(0.7108)	(0.6649)
98% "	97.87	97.88	98.01	97.93
	(0.4591)	(0.4804)	(0.4466)	(0.4406)

NOTE: $n_b = 1000$, $n_{sim} = 50$, MSE = Mean Square Error, STDV = Standard Deviation. The coefficients of government wages (W^G) and industry wages (W^I) are set equal in model estimation (δ_4) so that only four instead of five coefficients are reported here.

Table 2 is based on the original covariance matrix estimated from 3SLS residuals as reported by Theil. Scaling the error covariance by a factor of 1/5 relative to the one underlying Table 1 improves accuracy, as expected. Root Mean Square Error and bias measures of mean and variance estimates in Table 2 are respectively smaller than in Table 1 for all coefficients. The smaller error variance had less impact on the accuracy of the average bootstrapped probabilities and the size of their standard deviations. Across all coefficients and percentiles, the accuracy of the probabilities is virtually indistinguishable from the higher variance scenario. Standard deviations are on average slightly lower but the differences are small.

Table 2: **Distance Measures Between 2SLS-Mappings Based on BBMR and Parametric Posterior - Normal Errors, Small Error Variance**

	Structural Coefficients (True Value)			
Distance Measure	δ_1(16.44)	δ_2(0.1249)	δ_3(0.1631)	δ_4(0.7901)
$MSE^{1/2}$. Mean	0.1942	0.0508	0.0482	0.0325
Bias, Mean	0.0019	-1.1E-04	-9.7E-04	-6.4E-05
$MSE^{1/2}$, Variance	0.1002	5.4E-04	4.4E-04	8.3E-05
Bias, Variance	0.0197	7.4E-05	-1.7E-05	1.1E-05
2% probability	2.04	2.01	1.92	1.98
(STDV)	(0.4114)	(0.4767)	(0.4006)	(0.4506)
5% "	5.01	5.02	4.86	4.98
	(0.6492)	(0.7434)	(0.6575)	(0.7043)
10% "	10.06	10.02	9.85	9.97

		(0.9314)	(0.9826)	(0.9185)	(0.9277)
90%	"	89.89	89.92	89.97	89.90
		(0.9607)	(0.9618)	(0.9340)	(0.9140)
95%	"	94.95	94.91	94.97	94.90
		(0.6884)	(0.7068)	(0.6992)	(0.6531)
98%	"	97.93	97.88	97.95	97.93
		(0.4287)	(0.4427)	(0.4425)	(0.4303)

NOTE: See Table 1.

Of more interest to the empirical researcher are impacts of changing characteristics of the estimation problem that are actually controllable, which can generally not be said about the error variance of the data generating process.   Two approaches promising an improvement in approximation accuracy are the aforementioned second order sample correction technique and an increase in the bootstrap sample size. Table 3 allows these two cases to be compared with the "reference" scenario in Table 1 and with each other for the coefficients δ_2 and δ_4. Both approaches improve upon the accuracy of the mean and variance estimates presented in Table 1.

Table 3: **Distance Measures Between 2SLS-Mappings Based on BBMR and Parametric Posterior - Normal Errors, Comparison Between Second Order Corrected Bootstrap Samples and Large Bootstrap Sample Size**

| | Structural Coefficients (True Value) | | | |
| | Second Order Corrected | | Large Bootstrap Sample Size | |
Distance Measure	$\delta_2(0.1249)$	$\delta_4(0.7901)$	$\delta_2(0.1249)$	$\delta_4(0.7901)$
$MSE^{1/2}$, Mean	0.0575	0.0471	0.0358	0.0213
Bias, Mean	6.7E-05	9.7E-05	8.1E-05	-1.8E-05
$MSE^{1/2}$, Variance	0.0010	2.9E-04	3.0E-04	3.6E-05
Bias, Variance	-4.5E-05	-6.9E-06	7.7E-05	4.7E-05
2% probability	1.95	1.97	1.93	1.97
(STDV)	(0.4261)	(0.4274)	(0.2029)	(0.1952)
5% "	4.94	4.95	4.94	4.96
	(0.6325)	(0.6406)	(0.3141)	(0.2927)
10% "	9.91	9.95	9.96	9.98
	(0.9181)	(0.8157)	(0.4343)	(0.4078)
90% "	90.01	89.92	89.88	89.86
	(0.8590)	(0.8745)	(0.4359)	(0.4378)
95% "	95.02	94.93	94.85	94.88
	(0.6337)	(0.6262)	(0.3137)	(0.3224)
98% "	97.99	97.95	97.88	97.89
	(0.4261)	(0.4143)	(0.2037)	(0.2028)

NOTE: $n_b = 1000$ for first two columns and $n_b = 5000$ for last two columns, $n_{sim} = 50$, MSE = Mean Square Error, STDV = Standard Deviation. The coefficients of government wages (W^G) and industry wages (W^I) are set equal in model estimation (δ_4).

Increasing the bootstrap sample size from 1000 to 5000 clearly outperforms the second order correction technique if measured with respect to root mean square errors whereas the correction technique yields a lower bias for the variance estimates. Correcting the covariance matrix of the reduced form errors alone apparently is not as effective in stabilizing the expectations of the nonlinear mappings as is the large, and for applications certainly achievable, increase in bootstrap sample size. This is also confirmed looking at the standard deviations of the bootstrapped probabilities. Here the reduction for the increased sample size is around 50% compared to the reference scenario. The second order correction hardly improves at all on these measures. These results are somewhat in disagreement with the single equation case in Heckelei and Mittelhammer (1996): There, the second

order correction had a stronger impact on distance measures at comparable bootstrap sample sizes (n_b = 900) and the impact of increasing the bootstrap sample size was less significant. This again may be explained by the multivariate setting and the nonlinear mappings employed here that place higher demands on the bootstrap. Higher bootstrap sample sizes in the range between 1000 and 5000 still seem to improve upon the bootstrap?s ability to represent the information contained in the multivariate empirical distribution function.

Another important trait of a robust estimation approach is its performance under different probability regimes. Table 4 presents distance measures between BBMR based 2SLS-Mappings and those calculated via a parametric regression structure likelihood using the mixing algorithm described above when the disturbances are assumed to be multivariate T-distributed instead of normally distributed. Compared with the results for the normal distribution in Table 1, root mean square errors of mean and variance estimates and standard deviations of bootstrapped probabilities are slightly reduced for the T-distributed errors. The comparison of mean and variance bias is ambiguous whereas the average bootstrapped probabilities are somewhat less accurate than those for normal errors. Overall one can say that the general approximation accuracy of the BBMR-based 2SLS-mappings for a multivariate T-distribution is quite good and the accuracy did not notably differ from the normal case.

Table 4: **Distance Measures Between 2SLS-Mappings Based on BBMR and Parametric Posterior - Errors Distributed as Multivariate T with 5 d.f.**

	Structural Coefficients (True Value)			
Distance Measure	δ_1(16.44)	δ_2(0.1249)	δ_3(0.1631)	δ_4(0.7901)
$MSE^{1/2}$. Mean	0.2436	0.0582	0.0559	0.0439
Bias, Mean	-0.0387	-4.8E-04	0.0015	2.7E-04
$MSE^{1/2}$, Variance	0.2484	0.0010	7.7E-04	2.9E-04
Bias, Variance	-0.0325	-3.0E-04	3.1E-05	-1.0E-04
2% probability	2.05	1.96	1.94	1.70
(STDV)	(0.4344)	(0.4173)	(0.4245)	(0.3997)
5% "	5.17	5.00	4.92	4.57
	(0.6984)	(0.6765)	(0.6575)	(0.6287)
10% "	10.38	10.05	9.93	9.54
	(0.9629)	(0.9260)	(0.9033)	(0.9156)
90% "	90.36	90.40	89.69	89.95
	(0.8988)	(0.9132)	(0.9356)	(0.9141)
95% "	95.31	95.32	94.81	94.96
	(0.6524)	(0.6362)	(0.6921)	(0.6580)
98% "	98.15	98.18	97.94	97.93
	(0.4177)	(0.3980)	(0.4477)	(0.4296)

NOTE: See Table 1.

All of the types of simulations on 2SLS-mappings reported here were also done for 3SLS-mappings, but only for a bootstrap sample size of 1000. The absolute approximation accuracy as well as the findings regarding variations on error variance, second order correction and error distribution paralleled those for the 2SLS-mappings. Therefore, in Table 5, only distance measures for the standard simulation scenario, analogous to the simulation context of Table 1, are presented. Comparisons with Table 1 demonstrates the similarity of the results.

Table 5: **Distance Measures Between 3SLS -Mappings Based on BBMR and Parametric Posterior - Normal Errors**

Distance Measure	Structural Coefficients (True Value)			
	$\delta_1(16.44)$	$\delta_2(0.1249)$	$\delta_3(0.1631)$	$\delta_4(0.7901)$
$MSE^{1/2}$. Mean	0.2568	0.0612	0.0565	0.0457
Bias, Mean	0.0082	1.8E-04	0.0013	-1.8E-04
$MSE^{1/2}$, Variance	0.2777	0.0015	0.0007	0.0004
Bias, Variance	-0.0385	-1.1E-05	-1.5E-05	-5.0E-05
2% probability	1.90	1.94	1.95	1.94
(STDV)	(0.4301)	(0.4276)	(0.4494)	(0.4338)
5% "	4.79	4.95	4.91	4.93
	(0.6851)	(0.6996)	(0.6503)	(0.6844)
10% "	9.72	9.92	9.92	9.91
	(0.9153)	(0.9674)	(0.9127)	(0.9262)
90% "	89.86	89.90	89.92	90.12
	(0.9428)	(0.9425)	(0.9068)	(0.9254)
95% "	94.97	94.83	94.96	95.08
	(0.6806)	(0.7045)	(0.7013)	(0.6877)
98% "	97.94	97.86	97.94	98.03
	(0.4460)	(0.4581)	(0.4557)	(0.4400)

NOTE: See Table 1.

7 Conclusions

The Bayesian Bootstrap Regression (BBR) procedure developed by Heckelei and Mittelhammer (1996) has been generalized to a Bayesian Bootstrap Multivariate Regression (BBMR) approach. This allows for a generic, algorithmic Bayesian analysis of the traditional multivariate regression model without specification of a likelihood function and without restrictions on the form of prior densities. Combining BBMR with 2SLS- and 3SLS- mappings allows Bayesian analysis of simultaneous equation systems based on unrestricted reduced forms. Monte Carlo results

for the simultaneous equation setting presented in the paper indicate that the efficiency loss relative to a parametric approach under normality is small. Simulation results with multivariate T-distributed error terms suggest that the BBMR-algorithm may also be robust for elliptically contoured distributions, given that the data information is represented exclusively via a regression structure likelihood defined in terms of the sampling distribution of location and scale estimators. Since the underlying probability model for the regression disturbances is generally unknown, BBMR can be a robust and useful alternative to assuming normality and performing parametric Bayesian analysis with a potentially incorrect likelihood function.

Several areas in need of further research on Robust Bayesian Analysis in the system context can be identified. First of all, Monte Carlo studies are always of limited generality and additional simulations across other error distributions would provide additional insights regarding the robustness of the BBMR. Second, in the context of non-normal error distributions it would be interesting to compare the parametric normality-based approach with the BBMR both to evaluate the robustness of the normality assumption, and to assess possible relative improvements that BBMR can provide over incorrectly assuming normality. Third, the development of a bootstrap algorithm for performing restricted reduced form analysis of simultaneous equation systems would be desirable for cases in which the analyst felt that over identifying restrictions could be imposed with certainty. Finally, one might consider BBMR based on other robust estimators of location and scale to generate samples from likelihood functions that are possibly more robust than the Regression-Structure-Likelihood.

Summary

A Bayesian Bootstrap Multivariate Regression (BBMR) procedure is presented that allows robust Bayesian posterior analyses of traditional multivariate regression models. The procedure is then extended via 2SLS- and 3SLS-mappings of reduced form posterior distributions to facilitate robust posterior analyses of simultaneous equations systems. BBMR does not require the specification of a parametric family for the likelihood function and instead uses a bootstrapped likelihood based on the sampling distribution of location and scale estimators. It also allows a flexible choice of prior distributions and can be implemented

as a generic algorithm in standard statistical software independently of the actual choice of prior distribution.

Zusammenfassung

Der Beitrag stellt die Multivariate Bayes-Bootstrap -Regression (BBMR) vor, die eine robuste Analyse der A-posteriori Verteilungen traditioneller multivariater Regressionsmodelle erlaubt. Die Methode wird erweitert durch die Anwendung zweier Projektionen (verwandt mit der zweistufigen und dreistufigen Kleinstquadratmethode), die eine robuste Bayes'sche Analyse simultaner Gleichungssysteme ermöglichen. BBMR benötigt keine Spezifikation der Likelihood Funktion, sondern benutzt eine "bootstrap-Likelihood" auf der Grundlage der Stichprobenverteilungen von Lage- und Varianz-Schätzern. Gleichzeitig erlaubt BBMR eine flexible Spezifikation der A-priori Verteilungen und kann, unabhängig von der spezifischen Wahl der A-priori Verteilung, als standardisierter Algorithmus in Ökonometriesoftware-paketen implementiert werden.

References

Boos, D.D., and J.F. Monahan, 1986. Bootstrap methods using prior information. Biometrika 73, 77-83.

Cambanis, S., S. Huang, and G. Simons, 1981. On the theory of elliptically contoured distributions. J. of Multivariate Analysis 11, 368-385.

Davison, A.C., D.V. Hinkley, and B.J. Worton, 1992. Bootstrap likelihoods. Biometrika 79, 113-30.

Deemer, W.L. and I. Olkin, 1951. The jacobians of certain matrix transformations useful in multivariate analysis. Biometrika 38, 345-67.

Doksum, K.A. and A. Y. Lo, 1990. Consistent and robust Bayes procedures for location based on partial information. Ann. Stat. 18, 443-453.

Geweke, J., 1986. Exact inference in the inequality constrained normal linear regression model. J. of Appl. Econometrics 1, 127-41.

_____, 1989. Bayesian inference in econometric models using Monte Carlo integration. Econometrica 57, 1317-39.

_____, 1991. Generic, algorithmic approaches to Monte Carlo integration in Bayesian inference. Contemporary Mathematics 115, 117-34.

Graham, R.L., D.V. Hinkley, P.W.M. John, and S. Shi, 1990. Balanced design of bootstrap simulations. J. R. Statist. Soc. B. 52, 185-202.

Heckelei, Th. and R.C. Mittelhammer, 1996. Bayesian bootstrap inference via regression structure likelihood, in: T.B. Fomby and R.C. Hill, eds., Advances in Econometrics, Vol. 11A.

Herrmann, M., 1990. An economic analysis of world salmon markets: Effects of salmon farming. Ph. D. Dissertation, Dept. of Ag. Economics, Washington State Univ., Pullman.

Herrmann, M.L., R.C. Mittelhammer, and B.-H. Lin, 1992. Import demands for Norwegian farmed Atlantic salmon and wild Pacific salmon in North America, Japan and the EC. Can. J. Agr. Econ. 41, 111-25.

Johnson, M.E., 1987. Multivariate statistical simulation. John Wiley & Sons (New York).

Johnson, N.L. and S. Kotz, 1972. Distributions in statistics: Continuous multivariate distributions. John Wiley & Sons (New York).

Judge, G.G., R.C. Hill, W.E. Griffiths, H. Lütkepohl, and T.C. Lee, 1988. Introduction to the Theory and Practice of Econometrics, 2nd ed., John Wiley & Sons (New York).

Kloek, T. and H.K. van Dijk, 1978. Bayesian estimates of equation system parameters: An application of integration by Monte Carlo. Econometrica 46, 1-19.

Mittelhammer, R.C, 1990. "Bayesian estimation of the linear model made easy: Some thoughts on Bayesian bootstrapping without kernels." Seminar presented to the Program in Statistics, University of Idaho.

Monahan, J.F. and D.B. Boos, 1992. Proper likelihoods for Bayesian analysis. Biometrika 79, 271-8.

Pettit, A.N., 1982. Inference for the linear model using a likelihood based on ranks. J.R. Statist. Soc. B 44, 234-43.

_____, 1983. Likelihood based inference using signed ranks for matched pairs. J.R. Statist. Soc. B 45, 287-96.

Theil, H., 1971. Principles of Econometrics. John Wiley & Sons (New York).

van Dijk, H.K. and T. Kloek, 1980. Further experience in Bayesian analysis using Monte Carlo integration. J. of Econometrics 14, 307-28.

Zellner, A., 1971. An introduction to Bayesian inference in econometrics. John Wiley & Sons (New York).

Zellner, A., Bauwens, L., and H.K. van Dijk, 1988. Bayesian Specification analysis and estimation of simultaneous equation models using Monte Carlo methods. J. of Econometrics 38, 39-72.