ECONOMETRIC INSTITUTE

BAYESIAN SIMULTANEOUS EQUATIONS ANALYSIS USING REDUCED RANK STRUCTURES

FRANK KLEIBERGEN AND HERMAN K. VAN DIJK

REPORT 9714/A

ERASMUS UNIVERSITY ROTTERDAM - P.O. BOX 1738 - 3000 DR ROTTERDAM-THE NETHERLANDS
Internet-URL: http://www.eur.nl/few/ei/papers

Bayesian Simultaneous Equations Analysis using Reduced Rank Structures

Frank Kleibergen* Herman K. van Dijk†

April 10, 1997

Abstract

Diffuse priors lead to pathological posterior behavior when used in Bayesian analyses of Simultaneous Equation Models (SEMs). This results from the local nonidentification of certain parameters in SEMs. When this, a priori known, feature is not captured appropriately, an a posteriori favor for certain specific parameter values results which is not the consequence of strong data information but of local nonidentification. We show that a proper consistent Bayesian analysis of a SEM explicitly has to consider the reduced form of the SEM as a standard linear model on which nonlinear (reduced rank) restrictions are imposed, which result from a singular value decomposition. The priors/posteriors of the parameters of the SEM are therefore proportional to the priors/posteriors of the parameters of the linear model under the condition that the restrictions hold. This leads to a framework for constructing priors and posteriors for the parameters of SEMs. The framework is used to construct priors and posteriors for one, two and three structural equation SEMs. These examples jointly with a theorem, which states that the reduced forms of SEMs accord with sets of reduced rank restrictions on standard linear models, show how Bayesian analyses of generally specified SEMs are conducted.

---

*Corresponding author: Econometric Institute, Erasmus University Rotterdam, P.O. Box 1738, 3000 DR Rotterdam, the Netherlands, email:kleibergen@few.eur.nl.

†Econometric and Tinbergen Institute, Erasmus University Rotterdam, P.O. Box 1738, 3000 DR Rotterdam, the Netherlands. Part of this research was carried out while the first author was hosted at the department of Econometrics at Tilburg University, and during his PPS fellowship of the Netherlands Foundation for Scientific Research (N.W.O). We like to thank four anonymous referees, Luc Bauwens, Peter Phillips, Dale Poirier, Arnold Zellner, and other participants of the seminars discussing this paper for helpful comments and suggestions. Paper also available at: http://www.eur.nl/few/ei/faculty/kleibergen.

1 Introduction

Since the early forties people have focussed on the development of statistical methods for analyzing Simultaneous Equation Models (SEMs), see a.o. [16] and [2]. This shows the importance of models which are able to generate variables simultaneously as it is a stylized fact of many economic time series. The SEM is not only an important model but also a rather complicated one because of the problems which can occur regarding the identification of the different structural form parameters. The identification of the structural form parameters is reflected in the rank and order conditions which result from the reduced form, see [18]. The order condition reflects overall identification while the rank condition reflects local (non) identification. This latter phenomenon, local nonidentification, is shown to lead to pathological posterior behavior when diffuse priors are used in Bayesian analyses of the SEM. This behavior occurs in the traditional Bayesian analyses of SEMs documented in the literature, see a.o. [8], [10] and [11]. We show its occurence in a limited information (one equation) analysis of the SEM. Similar behavior can be found in other specifications of the SEM as well since the origin of the pathological posterior behavior, the local nonidentification of parameters, is exemplary to SEMs.

In order to obtain a consistent Bayesian analysis of a SEM, which does not suffer from different kind of pathologies, we construct a framework in which the reduced form of a SEM is specified as a multivariate linear model with nonlinear (reduced rank) restrictions on its parameters. Using singular value decompositions we specify the restrictions such that an one-to-one correspondence with a linear model is obtained when the restrictions do not hold; and the reduced form of the SEM is obtained when they hold. The proposed framework leads to invariance of the priors and posteriors with respect to the specification of the model. The resulting posteriors of the parameters of the SEM accord with the posterior of the embedding linear model. Our analysis is similar to the construction of the Savage-Dickey density ratio, see [7]. That is, we construct the priors/posteriors in the points where the hypothesis (restriction) holds. In contrast, the posterior of the parameters of a SEM, derived in the usual way using a diffuse prior, is inconsistent in the sense that the resulting posterior of its embedding linear model is not a member of the standard class of posteriors of linear models, see [19].

The contents of the paper is organized as follows. In section 2, we show the before mentioned pathologies arising in the posteriors of the parameters of an incomplete (one structural equation analysis of a) SEM when diffuse priors are used. Sections 3 and 4 show how an incomplete SEM is rewritten as a multivariate linear model with nonlinear parameter restrictions, which leads to a framework for prior and posterior analysis. Singular value decompositions are used in this framework which are similar to the canonical correlations used in a limited information maximum likelihood analysis, see [2]. In section 5, posterior simulators are constructed to sample from the posterior of the parameters of an incomplete SEM. Section 6 extends the one structural equation analysis to a full system analysis by showing that a fully specified SEM accords with a set of reduced rank restrictions on a linear model. Different subsections then show what the framework for prior/posterior analysis then amounts to for two and three structural equation examples and also show that the order condition for a full system analysis of a SEM can differ from the order condition resulting from an one structural equation analysis. Finally the seventh section concludes.

2 Nonidentification and Pathological Posterior Behavior

To show the consequences of local nonidentification of parameters of SEMs for their posteriors, we analyze, as an example, the one (set of) structural equation(s) model also known as INcomplete Simultaneous Equations Model (INSEM). As the results for the posteriors of the INSEM are exemplary for other specifications of the SEM, the importance of a proper treatment of the issue of local nonidentification is shown by the analysis of the INSEM.

We use as specification of the INSEM, see [38],

$$\begin{aligned} y_1 &= Y_2\beta + Z_1\gamma + \varepsilon_1, \\ Y_2 &= Z_1\Pi_{12} + Z_2\Pi_{22} + \varepsilon_2, \end{aligned} \tag{1}$$

where $y_1 : T \times 1$, and $Y_2 : T \times (m-1)$, are endogenous and $Z_1 : T \times k_1$, and $Z_2 : T \times k_2$, $k = k_1 + k_2$, contain the (weakly) exogenous, see [12], and lagged dependent variables, $\beta : (m-1) \times 1$, $\gamma : k_1 \times 1$, $\Pi_2 = (\Pi_{12}' \; \Pi_{22}')' : k \times (m-1)$ and we assume that $(\varepsilon_1 \; \varepsilon_2) \sim n(0, \Sigma \otimes I_T)$. The identification problems arise when the parameter $\Pi_{22} = 0$ (or has reduced rank) as (parts of) the structural form parameter β is then nonidentified. This is easily seen when we construct the reduced form of the INSEM from equation (1),

$$\begin{aligned} y_1 &= Z_1\pi_{11} + Z_2\Pi_{22}\beta + \xi_1, \\ Y_2 &= Z_1\Pi_{12} + Z_2\Pi_{22} + \varepsilon_2, \end{aligned} \tag{2}$$

where $\pi_{11} = \gamma + \Pi_{12}\beta$, $\xi_1 = \varepsilon_1 + \varepsilon_2\beta$, $(\xi_1 \; \varepsilon_2) \sim n(0, \Omega)$, $\Sigma = B'\Omega B$, $B = \begin{pmatrix} 1 & 0 \\ -\beta & I_{m-1} \end{pmatrix}$. When $\Pi_{22} = 0$, β drops out of the model in equation (2) and the disturbances ξ_1 are not affected by the value of β. So, the likelihood is flat and nonzero in the direction of β when $\Pi_{22} = 0$. If we use flat (diffuse) priors in a Bayesian analysis of the INSEM, such that the joint posterior is proportional to the likelihood, also the joint posterior of the different parameters will be flat and nonzero in the direction of β for zero values of Π_{22}. This property is passed on to the marginal posteriors, which are the integrals of the joint posterior over the

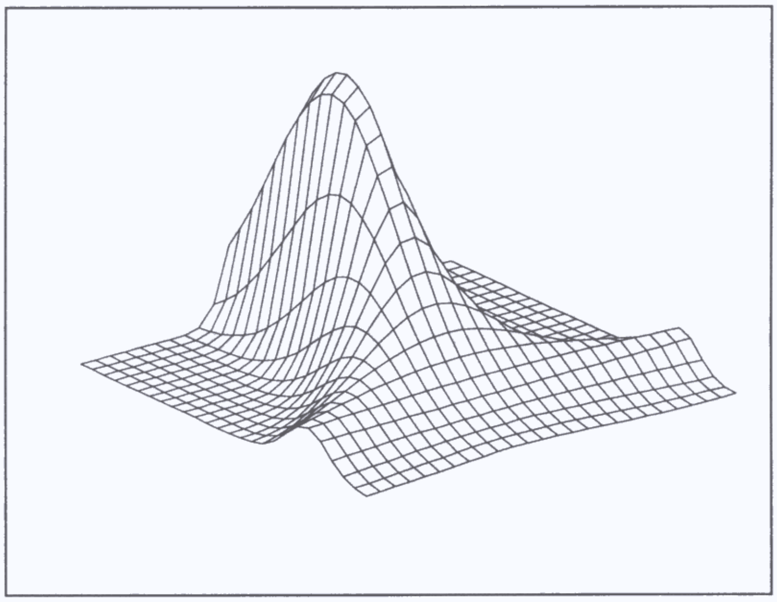

Figure 1: Bivariate Posterior (β, π_{22}) demand equation Tintner model

different parameters. To show the consequences for the marginal posteriors in practice, we calculated the marginal posteriors of the parameters of the demand equation of the "Tintner meat market" model, see [35]. In this exact identified model, y_1 reflects quantity of meat consumed, y_2 is the price of meat, z_1 is national income per capita, z_2 is the cost of processing meat (all variables are in deviation from their mean).

In figure 1, the joint posterior of β and Π_{22} is drawn for the Tintner meat market dataset and figure 2 contains the contourlines of this bivariate posterior. The functional form of this posterior is obtained by using a flat prior ($\propto 1$) and integrating out $(\Sigma, \pi_{11}, \pi_{12})$, and reads,

$$\begin{aligned} p(\beta, \Pi_{22}|Y, Z) \;\; \propto \;\; & |(y_1 - Z_2\Pi_{22}\beta)' M_{(Z_1\,\varepsilon_2)}(y_1 - Z_2\Pi_{22}\beta)|^{-\frac{1}{2}(T-k_1-m-1)} \\ & |(Y_2 - Z_2\Pi_{22})' M_{Z_1}(Y_2 - Z_2\Pi_{22})|^{-\frac{1}{2}(T-k_1-m-1)}, \end{aligned} \quad (3)$$

as $Y_2 = Z_1\Pi_{12} + Z_2\Pi_{22} + \varepsilon_2$ and $M_V = I_T - V(V'V)^{-1}V'$, $V = Z_1$, $V = (Z_1\, \varepsilon_2)$. Both figures 1, 2, and the functional form of the posterior in equation (3) show that the marginal posterior does not depend on β when $\Pi_{22} = 0$ as it is flat and nonzero in the direction of β for zero values of Π_{22}. This implies that the marginal posterior of Π_{22}, which is the integral of the posterior in equation (3) over β, will be infinite at $\Pi_{22} = 0$ as at this particular value of Π_{22}, we construct an integral of a function over an infinite parameter region while the function itself does not depend on the parameter β over which we integrate. So, the integral will be proportional to the size of the parameter region, i.e. infinity. Both the

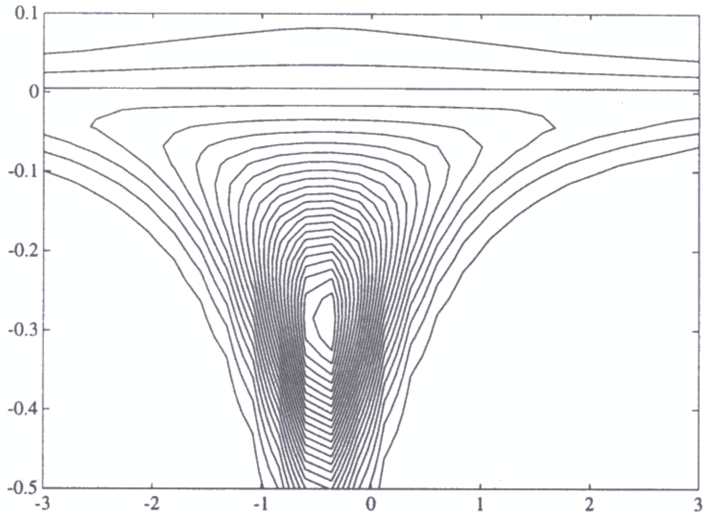

Figure 2: Contourlines marginal posterior (β, Π_{22}) demand equation

functional form of the marginal posterior of Π_{22},

$$p(\Pi_{22}|Y,Z) \propto |\Pi_{22}' Z_2' M_{(Z_1\,\varepsilon_2)} Z_2 \Pi_{22}|^{-\frac{1}{2}} [\frac{|\Pi_{22}' Z_2' M_{(Z_1\,Y_2)} Z_2 \Pi_{22}|}{|\Pi_{22}' Z_2' M_{(Z_1\,Y_1\,Y_2)} Z_2 \Pi_{22}|}]^{-\frac{1}{2}(T-k_1-2(m-1))}$$
$$|(Y_2 - Z_2\Pi_{22})' M_{Z_1} (Y_2 - Z_2\Pi_{22})|^{-\frac{1}{2}(T-k_1-m-1)}, \qquad (4)$$

and the marginal posterior of Π_{22} for the Tintner dataset from figure 3 show this phenomenon and consequently the value of the posterior of Π_{22} is infinite at $\Pi_{22} = 0$.

The nonidentification of β has also consequences for its own marginal posterior. The marginal posterior of β, which belongs to the class of 1-1 poly t densities, see [3], [8], [9], [11], and [31] for an efficient algorithm to calculate the moments of this class of densities, reads

$$p(\beta|Y,Z) \propto |(y_1 - Y_2\beta)' M_{(Z_1\,Z_2)} (y_1 - Y_2\beta)|^{\frac{1}{2}(T-k_1-k_2-m-1)} \qquad (5)$$
$$|(y_1 - Y_2\beta)' M_{Z_1} (y_1 - Y_2\beta)|^{-\frac{1}{2}(T-k_1-m-1)},$$

and this posterior has fat tails resulting from the flat nonzero conditional posterior of β given $\Pi_{22} = 0$. For the case of the Tintner model, the marginal posterior is even nonintegrable which is also plausible given the fat tails of the marginal posterior of β shown in figure 4. In general, the moments of the posterior in equation (5) exist up to/including the degree of overidentification minus 1 implying that exact identified models lead to nonintegrable posteriors when diffuse priors are used.

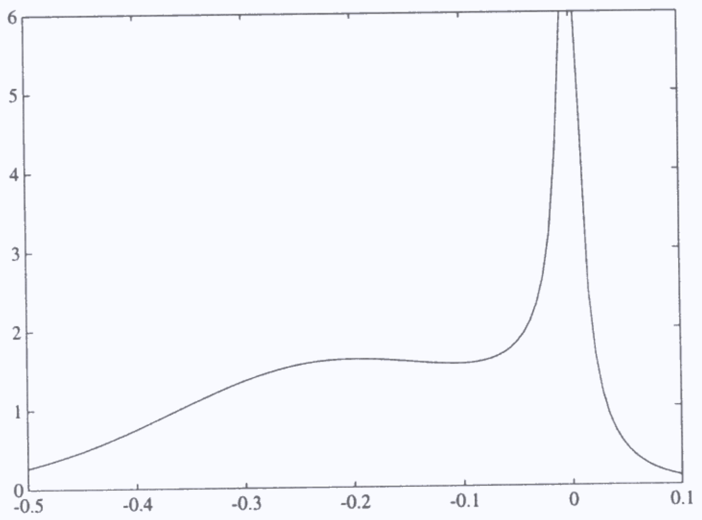

Figure 3: Marginal posterior π_{22} demand equation Tintner model

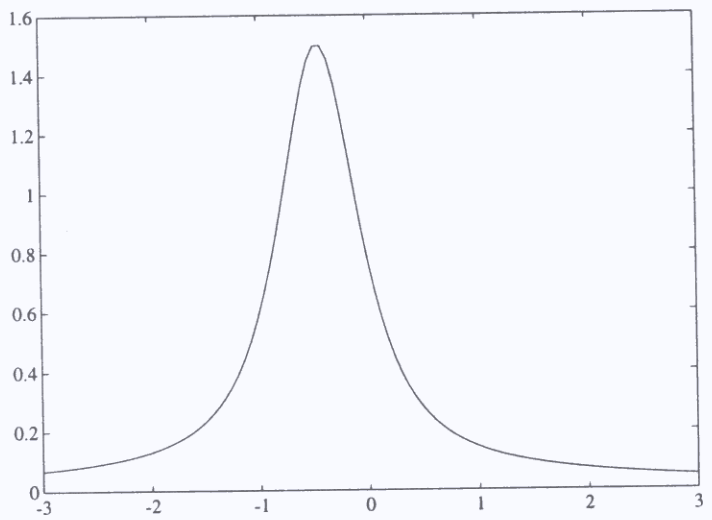

Figure 4: Marginal posterior β demand equation Tintner model

A popular method for numerical calculation of posterior densities is by constructing the conditional posteriors and to perform Gibbs Sampling, see [13] and [33]. When this Markov Chain Monte Carlo (MC^2) algorithm is used to compute the marginal posteriors of the parameters of the INSEM, as in [15], the local nonidentification problems lead to a reducible Markov Chain since when a locally nonidentified parameter value is drawn, the sampler continues drawing nonidentified parameter values, such that the region of locally nonidentified parameter values is an absorbing state in the Markov chain. The posterior violates, therefore, the convergence conditions for Gibbs Samplers as outlined in [32]. A solution to this problem is to use informative priors but the validity of this approach is questionable when priors are used which are not in accordance with the likelihood.

The integrability problems of the posteriors discussed previously result from the dependence of the structural form parameter β on Π_{22} in the INSEM. In classical statistical analysis, see [1], [28] and [30], the parameter β is analyzed conditional on a so-called concentration parameter, which is essentially a Wald test for $\Pi_{22} = 0$ and shows whether the information in the likelihood is concentrated around $\Pi_{22} = 0$. When this concentration parameter tends to infinity in the limit, normal asymptotic theory can be applied, see [1] and [28]. If $\Pi_{22} = 0$, however, estimators of β, like 2SLS, converge to random variables, see [29]. The integrability problems outlined above show that also in a Bayesian analysis β has to be analyzed given Π_{22}, which is natural given that the identification problems in the likelihood result from model inadequacies, i.e. the nonidentification of β at $\Pi_{22} = 0$, and are not the result of inferior data. As we know a priori that these integration problems will arise, there is a need for a framework which formalizes the way the parameters are analyzed conditional on one another and which leads to nonpathological posteriors. In the next sections, this framework is constructed.

3 Priors for the INSEM parameters

In the previous section, we showed that the parameters which suffer from local nonidentification problems should be analyzed conditional on the value of their identifying parameters. This is one of the main properties obtained through the priors constructed in this section. In previous versions of this paper, see [22], and also [5] Jeffreys' priors are used to obtain this property. The resulting posterior, when this prior is used, is, however, not nested within the assumed posterior of the embedding unrestricted linear model. This is a key property of the priors constructed in this section. The prior we construct in this section is based on [19], where it is shown that a whole range of models can be considered as nonlinear restrictions of the parameters of a standard linear model. This gives a general framework for the analysis of a large class of models, see for example [20] and [21].

3.1 SEMs as linear models with nonlinear parameter restrictions

Overidentified SEMs can be seen as a nonlinear restriction on the parameters of a multivariate linear model. It is well known how priors/posteriors for the standard linear model are constructed. If we explicitly consider the SEM as a nonlinear restriction on the parameters of a linear model, the priors/posteriors of the parameters of the SEM result, straightforwardly, as proportional to the priors/posteriors of the parameters of the linear model under the condition that the restrictions on these parameters hold.

To show the restrictions imposed by a SEM on the parameters of a linear model consider the INSEM (1) and its reduced form (2) which is a linear model with restrictions on its parameters. To show these restrictions, we add a parameter λ to this model which is such that when it is nonzero, (*i.*) there is an one-to-one correspondence with a standard linear model and when it equals zero both (*ii.*) the reduced form of the INSEM results and (*iii.*) it is locally uncorrelated with specific other parameters. This latter property is needed to obtain priors/posteriors of the parameters of the INSEM which are invariant with respect to the specification of the model. Several restrictions imposed on the linear model namely lead to the reduced form of the INSEM but only one restriction also leads to priors/posteriors which are invariant with respect to parameter transformations. The non invariance is a consequence of the Borel-Kolmogorov Paradox, see [4] and [11], and for more details on this posterior invariance, see [19] and [21]. The resulting model, which we call the unrestricted SEM, reads,

$$\begin{aligned} \begin{pmatrix} y_1 & Y_2 \end{pmatrix} &= Z_1 \begin{pmatrix} \pi_{11} & \Pi_{12} \end{pmatrix} + Z_2 \Pi_{22} \begin{pmatrix} \beta & I_{m-1} \end{pmatrix} \\ &\quad + Z_2 \Pi_{22\perp} \lambda \begin{pmatrix} \beta & I_{m-1} \end{pmatrix}_{\perp} + \begin{pmatrix} \xi_1 & \varepsilon_2 \end{pmatrix}, \end{aligned} \tag{6}$$

where $\lambda : (k_2 - m + 1) \times 1$ and $\Pi_{22\perp}$, $\begin{pmatrix} \beta & I_{m-1} \end{pmatrix}_{\perp}$ are the orthogonal complements of Π_{22}, $\begin{pmatrix} \beta & I_{m-1} \end{pmatrix}$ resp., such that $\Pi_{22}' \Pi_{22\perp} = 0$, $\begin{pmatrix} \beta & I_{m-1} \end{pmatrix} \begin{pmatrix} \beta & I_{m-1} \end{pmatrix}_{\perp}' = 0$, and $\Pi_{22\perp}' \Pi_{22\perp} = I_{k_2 - m + 1}$, $\begin{pmatrix} \beta & I_{m-1} \end{pmatrix}_{\perp} \begin{pmatrix} \beta & I_{m-1} \end{pmatrix}_{\perp}' = 1$ (i.e. $\Pi_{22\perp} = \begin{pmatrix} -\Pi_{222} \Pi_{221}^{-1} & I_{k_2 - m - 1} \end{pmatrix}' (I_{k_2 - m - 1} + \Pi_{222} \Pi_{221}^{-1} \Pi_{221}^{-1\prime} \Pi_{222}')^{-\frac{1}{2}}$, when $\Pi_{22} = \begin{pmatrix} \Pi_{221}' & \Pi_{222}' \end{pmatrix}'$, $\Pi_{221} : (m-1) \times (m-1)$, $\Pi_{222} : (k_2 - m + 1) \times (m-1)$, and $\begin{pmatrix} \beta & I_{m-1} \end{pmatrix}_{\perp} = (1 + \beta'\beta)^{-\frac{1}{2}} \begin{pmatrix} 1 & -\beta' \end{pmatrix}$). Note that the orthogonal complements used in other parts of the paper are defined identical to the ones stated above.

It is clear that when $\lambda = 0$, (6) is identical to (2) and since λ is multiplied by the orthogonal complements of the matrices containing β and Π_{22}, the information matrix is block diagonal at $\lambda = 0$. We therefore say that λ is locally uncorrelated with β and Π_{22} at $\lambda = 0$. The one-to-one correspondence between the parameters of (6) and a standard linear model,

$$\begin{pmatrix} y_1 & Y_2 \end{pmatrix} = \begin{pmatrix} Z_1 & Z_2 \end{pmatrix} \begin{pmatrix} \pi_{11} & \Pi_{12} \\ \phi_1 & \Phi_2 \end{pmatrix} + \begin{pmatrix} \xi_1 & \varepsilon_2 \end{pmatrix}, \tag{7}$$

where $\phi_1 : k_2 \times 1$, $\Phi_2 : k_2 \times (m-1)$, can be shown using a Singular Value Decomposition (SVD) of $\Phi = \begin{pmatrix} \phi_1 & \Phi_2 \end{pmatrix}$, see [26],

$$\Phi = USV', \tag{8}$$

where $U : k_2 \times m$, $U'U = I_m$; $V : m \times m$, $V'V = I_m$; and $S : m \times m$ is a diagonal matrix containing the (nonnegative) singular values (in decreasing order). If we write,

$$U = \begin{pmatrix} U_{11} & u_{12} \\ U_{21} & u_{22} \end{pmatrix}, \; S = \begin{pmatrix} S_1 & 0 \\ 0 & s_2 \end{pmatrix} \text{ and } V = \begin{pmatrix} v_{11} & v_{12} \\ V_{21} & v_{22} \end{pmatrix}, \tag{9}$$

where U_{11}, S_1, $V_{21} : (m-1) \times (m-1)$; s_2, $v_{12} : 1 \times 1$; v_{11}', v_{22}, $u_{12} : (m-1) \times 1$; $U_{21} : (k_2 - m + 1) \times (m-1)$; $u_{22} : (k_2 - m + 1) \times 1$, the following relationship between $(\Pi_{22}, \beta, \lambda)$ and (U, S, V) results,

$$\Pi_{22} = \begin{pmatrix} U_{11} \\ U_{21} \end{pmatrix} S_1 V_{21}', \; \beta = V_{21}'^{-1} v_{11}', \text{ and } \lambda = (u_{22}' u_{22})^{-\frac{1}{2}} u_{22} s_2 v_{12}' (v_{12}' v_{12})^{-\frac{1}{2}}. \tag{10}$$

Furthermore the SVD shows that λ is identified by the smallest singular value of Φ contained in s_2.

The above implies that the INSEM can be considered as a nonlinear (reduced rank) restriction, $\lambda = 0$, on the parameters of the linear model in (7). We can therefore construct the priors/posteriors of the INSEM (1) as proportional to the priors/posteriors of the parameters of the linear model (7) evaluated in $\lambda = 0$. In the following subsection we discus a framework to construct priors and posteriors which is based on this property. The resulting framework for prior and posterior analysis can also be used in a full system analysis in which SVDs have to be applied recursively. As this becomes notationally more complicated we discus this in a later section. Note also that the analysis for exact identified SEMs directly results from the standard linear model since there is an one-to-one correspondence between the parameters of the structural form and the linear model in that case.

3.2 Prior Framework for SEMs

As shown previously, the INSEM can be considered as a nonlinear restriction of a standard linear model. It is, however, not possible to analytically derive the conditional posterior of the parameters, π_{11}, β, Π_{12} and Π_{22}, given the parameter reflecting the restrictions, λ, and Ω, see also [19]. To show this let $\theta = (\pi_{11}, \beta, \Pi_{12}, \Pi_{22})$ and $\eta = (\Phi, \pi_{11}, \Pi_{12})$, then

$$p_{unsem}(\theta, \lambda | \Omega, Y, Z) \propto p_{lin}(\eta(\theta, \lambda) | \Omega, Y, Z) | \frac{\partial \eta}{\partial(\theta, \lambda)} |, \tag{11}$$

where η is a function of θ and λ, *unsem* stands for unrestricted SEM and *lin* for linear model. Assume that the posterior of η is well behaved, which is typically the case for the posterior of the parameters of a standard linear model, then we cannot give an exact expression of the conditional posterior of θ given λ, $p_{unsem}(\theta | \lambda, \Omega, Y, Z)$, including its normalizing constants. Consequently to obtain a consistent analysis, in the sense that the INSEM has to be comparable with its embedding linear model, we cannot ignore that the INSEM is a linear model with nonlinear restrictions and just proceed by deriving a posterior as done in section 2. This derivation namely implicitly assumes that the posterior is proportional to $p_{unsem}(\theta, \lambda | \Omega, Y, Z)|_{\lambda=0}$, which would imply a posterior for the parameters of the linear model in $\lambda = 0$,

$$p_{lin}(\eta | \Omega, Y, Z)|_{\lambda=0} \propto p_{unsem}(\eta(\theta, \lambda) | \Omega, Y, Z)|_{\lambda=0} | (\frac{\partial(\theta, \lambda)}{\partial \eta})|_{\lambda=0} |. \tag{12}$$

As shown in section 2 the posterior $p_{unsem}(\theta, \lambda | \Omega, Y, Z)|_{\lambda=0}$ is badly behaved and the resulting behavior of $p_{lin}(\eta | \Omega, Y, Z)|_{\lambda=0}$ will only be worse. This posterior of the parameters of a linear model does not belong to (or is nested within) the standard class of posteriors of parameters of linear models and therefore the analysis is inconsistent. Also slight modifications of the INSEM, to for example an INSEM which is nested in the original INSEM, lead to a completely different posterior of the parameters of the embedding linear model which emphasizes the inconsistency even more. We therefore use the priors/posteriors of the parameters of the linear model as a base to construct the priors/posteriors of the parameters of the INSEM. So, we specify a prior for the parameters of the linear model, for example a diffuse or natural-conjugate prior, see [37], and we evaluate this prior in $\lambda = 0$ to obtain the prior for the INSEM, see also [21],

$$\begin{aligned} p_{insem}(\theta, \Omega) &\propto p_{unsem}(\theta, \lambda, \Omega)|_{\lambda=0} \\ &\propto p_{lin}(\eta(\theta, \lambda), \Omega)|_{\lambda=0} | (\frac{\partial \eta}{\partial(\theta, \lambda)})|_{\lambda=0} |. \end{aligned} \tag{13}$$

where *insem* stands for INSEM.

3.2.1 Diffuse Prior

Using the framework resulting from (13), a diffuse (Jeffreys') prior for the parameters of the linear model, $(\pi_{11}, \Pi_{12}, \Phi, \Omega)$,

$$p_{lin}(\pi_{11}, \Pi_{12}, \Phi, \Omega) \propto |\Omega|^{-\frac{1}{2}(m+1)} |\Omega^{-1} \otimes Z'Z|^{\frac{1}{2}}, \tag{14}$$

implies the prior for the parameters of the INSEM, $(\beta, \pi_{11}, \Pi_{12}, \Pi_{22}, \Omega)$,

$$
\begin{aligned}
& p_{insem}(\beta, \pi_{11}, \Pi_{12}, \Pi_{22}, \Omega) \qquad (15)\\
\propto\ & |\Omega|^{-\frac{1}{2}(m+1)} |\Omega^{-1} \otimes Z'Z|^{\frac{1}{2}} |J(\Phi, (\Pi_{22}, \beta, \lambda))|_{\lambda=0}| \\
\propto\ & |\Omega|^{-\frac{1}{2}(k_1+m+1)} |Z_1'Z_1|^{\frac{1}{2}m} |B_\perp \Omega^{-1} B_\perp'|^{\frac{1}{2}(k_2-m+1)} |\Pi_{22\perp}' Z_2' M_{Z_1} Z_2 \Pi_{22\perp}|^{\frac{1}{2}} \\
& |\begin{pmatrix} B\Omega^{-1}B' \otimes Z_2' M_{Z_1} Z_2 & B_\perp \Omega^{-1} e_1 \otimes Z_2' M_{Z_1} Z_2 \Pi_{22} \\ e_1'\Omega^{-1}B' \otimes \Pi_{22}' Z_2' M_{Z_1} Z_2 & e_1'\Omega^{-1} e_1 \otimes \Pi_{22}' Z_2' M_{Z_1} Z_2 \Pi_{22} \end{pmatrix}|^{\frac{1}{2}}
\end{aligned}
$$

where $B = \begin{pmatrix} \beta & I_{m-1} \end{pmatrix}$, $|J(\Phi, (\Pi_{22}, \beta, \lambda))| = |\frac{\partial \eta}{\partial(\theta, \lambda)}|$ and is constructed in appendix a.

The prior (15) shows that β is analyzed conditional on the value of Π_{22} as it should be according to the local nonidentification of β for lower rank values of Π_{22}. Furthermore, the prior shows the functional form of a diffuse prior for the parameters of the INSEM. This accords with our conclusions from the previous section that diffuseness for models like the INSEM has to defined in a different way.

3.2.2 Natural Conjugate Prior

In case of a natural conjugate prior for the parameters of the linear model, we specify an inverted-Wishart prior for Ω and a matrix normal prior for $(\pi_{11}, \Pi_{12}, \Phi)$ given Ω,

$$
\begin{aligned}
p_{lin}(\Omega) &\propto |G|^{\frac{1}{2}h} |\Omega|^{-\frac{1}{2}(h+m+1)} \exp[-\frac{1}{2} tr(\Omega^{-1}G)] \qquad (16)\\
p_{lin}(\pi_{11}, \Pi_{12}, \Phi|\Omega) &\propto |\Omega|^{-\frac{1}{2}m} |A|^{\frac{1}{2}k} \exp[-\frac{1}{2} tr(\Omega^{-1}(\begin{pmatrix} \pi_{11} & \Pi_{12} \\ \varphi_1 & \Phi_2 \end{pmatrix} - P)' \\
& \qquad A(\begin{pmatrix} \pi_{11} & \Pi_{12} \\ \varphi_1 & \Phi_2 \end{pmatrix} - P))],
\end{aligned}
$$

where $G : m \times m$, $A : k \times k$, G and A are positive definite symmetric (pds) matrices, $P : k \times m$, and h is the prior degrees of freedom parameter. The matrix A can be decomposed as

$$
A = \begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix}, \qquad (17)
$$

where $A_{11} : k_1 \times k_1$, $A_{12} = A_{21}' : k_2 \times k_1$, $A_{22} : k_2 \times k_2$. The prior of the parameters of the INSEM resulting from $p_{lin}(\pi_{11}, \Pi_{12}, \Phi, \Omega)$ can again be constructed using (13),

$$
\begin{aligned}
& p_{insem}(\beta, \pi_{11}, \Pi_{12}, \Pi_{22}, \Omega) \qquad (18)\\
\propto\ & p_{lin}(\pi_{11}, \Pi_{12}, \Phi(\beta, \Pi_{22}, \lambda), \Omega)|_{\lambda=0} |J(\Phi, (\Pi_{22}, \beta, \lambda))|_{\lambda=0}| \\
\propto\ & |G|^{\frac{1}{2}h} |\Omega|^{-\frac{1}{2}(h+k_1+m+1)} |A_{11}|^{\frac{1}{2}m} |B_\perp \Omega^{-1} B_\perp'|^{\frac{1}{2}(k_2-m+1)} |\Pi_{22\perp}' A_{22.1} \Pi_{22\perp}|^{\frac{1}{2}}
\end{aligned}
$$

$$|\begin{pmatrix} B\Omega^{-1}B' \otimes A_{22.1} & B\Omega^{-1}e_1 \otimes A_{22.1}\Pi_{22} \\ e_1'\Omega^{-1}B' \otimes \Pi_{22}'A_{22.1} & e_1'\Omega^{-1}e_1 \otimes \Pi_{22}'A_{22.1}\Pi_{22} \end{pmatrix}|^{\frac{1}{2}}$$
$$\exp[-\frac{1}{2}tr(\Omega^{-1}(G+(\begin{pmatrix} \pi_{11} & \Pi_{12} \\ \Pi_{22}\beta & \Pi_{22} \end{pmatrix} - P)'A(\begin{pmatrix} \pi_{11} & \Pi_{12} \\ \Pi_{22}\beta & \Pi_{22} \end{pmatrix} - P)))],$$

where $A_{22.1} = A_{22} - A_{21}A_{11}^{-1}A_{12}$ and the specification of (18) is not unique.

It may be that we have more knowledge about possible values of the parameters of the INSEM than about the parameters of the linear model. This knowledge can be used in the construction of the prior of the parameters of the linear model as these parameters are an exact function of the parameters of the INSEM when the restriction $\lambda = 0$ holds.

The prior (18) does not belong to a known class of probability density functions and we do not know analytical expressions of its moments or normalizing constant. These properties can be calculated using Monte-Carlo simulation and in the fifth section we construct this simulation algorithm.

4 Posteriors of the INSEM parameters

The framework for constructing the priors of the parameters of the INSEM can directly be applied to construct the posteriors of the parameters of the INSEM. Since the likelihood of the INSEM is a continuous function of the parameters, it follows that the posterior, which is proportional to the product of the prior and the likelihood, can be evaluated in the same way as the prior,

$$\begin{aligned} & p_{insem}(\beta, \pi_{11}, \Pi_{12}, \Pi_{22}, \Omega|Y, Z) && (19) \\ \propto\ & p_{insem}(\beta, \pi_{11}, \Pi_{12}, \Pi_{22}, \Omega)L_{insem}(Y|\beta, \pi_{11}, \Pi_{12}, \Pi_{22}, \Omega, Z) \\ \propto\ & p_{unsem}(\beta, \lambda, \pi_{11}, \Pi_{12}, \Pi_{22}, \Omega)|_{\lambda=0}L_{unsem}(Y|\beta, \lambda, \pi_{11}, \Pi_{12}, \Pi_{22}, \Omega, Z)|_{\lambda=0} \\ \propto\ & p_{lin}(\pi_{11}, \Pi_{12}, \Phi(\beta, \lambda, \Pi_{22}), \Omega)|_{\lambda=0}|J(\Phi, (\Pi_{22}, \beta, \lambda))|_{\lambda=0}| \\ & L_{lin}(Y|\pi_{11}, \Pi_{12}, \Phi(\beta, \lambda, \Pi_{22}), \Omega, Z)|_{\lambda=0}. \end{aligned}$$

In the following two subsections, we construct the posteriors for different specifications of the prior, i.e. a diffuse and natural conjugate prior.

4.1 Posterior INSEM using Diffuse Prior

Using the diffuse prior (15), the joint posterior of the parameters of the INSEM can directly be constructed from this prior and the likelihood using (19),

$$\begin{aligned} & p_{insem}(\beta, \pi_{11}, \Pi_{12}, \Pi_{22}, \Omega|Y, Z) && (20) \\ \propto\ & p_{insem}(\beta, \pi_{11}, \Pi_{12}, \Pi_{22}, \Omega)L(Y|\beta, \pi_{11}, \Pi_{12}, \Pi_{22}, \Omega, Z) \\ \propto\ & |\Omega|^{-\frac{1}{2}(T+k_1+m+1)}|Z_1'Z_1|^{\frac{1}{2}m}|B_\perp\Omega^{-1}B_\perp'|^{\frac{1}{2}(k_2-m+1)}|\Pi_{22\perp}'Z_2'M_{Z_1}Z_2\Pi_{22\perp}|^{\frac{1}{2}} \end{aligned}$$

$$
\left| \begin{pmatrix} B\Omega^{-1}B' \otimes Z_2'M_{Z_1}Z_2 & B\Omega^{-1}e_1 \otimes Z_2'M_{Z_1}Z_2\Pi_{22} \\ e_1'\Omega^{-1}B' \otimes \Pi_{22}'Z_2'M_{Z_1}Z_2 & e_1'\Omega^{-1}e_1 \otimes \Pi_{22}'Z_2'M_{Z_1}Z_2\Pi_{22} \end{pmatrix} \right|^{\frac{1}{2}}
$$

$$
\exp[-\frac{1}{2}tr(\Omega^{-1}((\begin{array}{cc} y_1 & Y_2 \end{array}) - (\begin{array}{cc} Z_1 & Z_2 \end{array}) \begin{pmatrix} \pi_{11} & \Pi_{12} \\ \Pi_{22}\beta & \Pi_{22} \end{pmatrix})'
$$

$$
((\begin{array}{cc} y_1 & Y_2 \end{array}) - (\begin{array}{cc} Z_1 & Z_2 \end{array}) \begin{pmatrix} \pi_{11} & \Pi_{12} \\ \Pi_{22}\beta & \Pi_{22} \end{pmatrix}))].
$$

The posterior (20) does not belong to a known class of probability density functions nor do any of the conditional posteriors, apart from the conditional posterior of (π_{11}, Π_{12}) given $(\beta, \Pi_{22}, \Omega)$, which is matrix-normal, belong to a known class of probability density functions. So, we can only analytically integrate out (π_{11}, Π_{12}) to obtain the marginal posterior of $(\beta, \Pi_{22}, \Omega)$,

$$
\begin{aligned}
&p_{insem}(\beta, \Pi_{22}, \Omega|Y, Z) \propto |\Omega|^{-\frac{1}{2}(T+m+1)}|B_\perp\Omega^{-1}B_\perp'|^{\frac{1}{2}(k_2-m+1)} \\
&\left| \begin{pmatrix} B\Omega^{-1}B' \otimes Z_2'M_{Z_1}Z_2 & B\Omega^{-1}e_1 \otimes Z_2'M_{Z_1}Z_2\Pi_{22} \\ e_1'\Omega^{-1}B' \otimes \Pi_{22}'Z_2'M_{Z_1}Z_2 & e_1'\Omega^{-1}e_1 \otimes \Pi_{22}'Z_2'M_{Z_1}Z_2\Pi_{22} \end{pmatrix} \right|^{\frac{1}{2}} \\
&|\Pi_{22\perp}'Z_2'M_{Z_1}Z_2\Pi_{22\perp}|^{\frac{1}{2}} \exp[-\frac{1}{2}tr(\Omega^{-1}((\begin{array}{cc} y_1 & Y_2 \end{array}) - Z_2\Pi_{22}(\begin{array}{cc} \beta & I_{m-1} \end{array}))' \\
&M_{Z_1}((\begin{array}{cc} y_1 & Y_2 \end{array}) - Z_2\Pi_{22}(\begin{array}{cc} \beta & I_{m-1} \end{array})))]. \qquad (21)
\end{aligned}
$$

The posterior (21) is proportional to the marginal posterior of (Φ, Ω) evaluated in $\lambda = 0$,

$$
p_{insem}(\beta, \Pi_{22}, \Omega|Y, Z) \propto p_{lin}(\Phi(\beta, \lambda, \Pi_{22}), \Omega|Y, Z)|_{\lambda=0}|J(\Phi, (\Pi_{22}, \beta, \lambda))|_{\lambda=0}|. \qquad (22)
$$

In section 5, we construct Importance and Metropolis-Hastings samplers for calculating the marginal posteriors and moments of (21) which exploit (22).

4.2 Posterior INSEM using Natural Conjugate Prior

Identical to the posterior of the parameters of the INSEM using a diffuse prior (20), we can construct the posterior of the parameters of the INSEM when we use the natural conjugate prior (18),

$$
\begin{aligned}
&p_{insem}(\beta, \pi_{11}, \Pi_{12}, \Pi_{22}, \Omega|Y, Z) \qquad (23) \\
\propto\ &|\Omega|^{-\frac{1}{2}(T+k_1+m+1)}|(A+Z'Z)_{11}|^{\frac{1}{2}m}|B_\perp\Omega^{-1}B_\perp'|^{\frac{1}{2}(k_2-m+1)} \\
&\left| \begin{pmatrix} B\Omega^{-1}B' \otimes (A+Z'Z)_{22.1} & B\Omega^{-1}e_1 \otimes (A+Z'Z)_{22.1}\Pi_{22} \\ e_1'\Omega^{-1}B' \otimes \Pi_{22}'(A+Z'Z)_{22.1} & e_1'\Omega^{-1}e_1 \otimes \Pi_{22}'(A+Z'Z)_{22.1}\Pi_{22} \end{pmatrix} \right|^{\frac{1}{2}} \\
&|\Pi_{22\perp}'(A+Z'Z)_{22.1}\Pi_{22\perp}|^{\frac{1}{2}} \exp[-\frac{1}{2}tr(\Omega^{-1}(\tilde{G} + (\begin{pmatrix} \pi_{11} & \Pi_{12} \\ \Pi_{22}\beta & \Pi_{22} \end{pmatrix} - \tilde{\Pi})' \\
&(A+Z'Z)(\begin{pmatrix} \pi_{11} & \Pi_{12} \\ \Pi_{22}\beta & \Pi_{22} \end{pmatrix} - \tilde{\Pi})))].
\end{aligned}
$$

where $\tilde{\Pi} = (A + Z'Z)^{-1}(Z'Y + A'P)$, $\tilde{G} = G + Y'Y - \tilde{\Pi}'(A + Z'Z)\tilde{\Pi}$, $Y = \begin{pmatrix} y_1 & Y_2 \end{pmatrix}$. Again similar to the posterior using a diffuse prior (20), only the conditional posterior of (π_{11}, Π_{12}) given $(\beta, \Pi_{22}, \Omega)$ belongs to a known class of probability density functions and (π_{11}, Π_{12}) are the only parameters which can be integrated out analytically to obtain the marginal posterior of $(\beta, \Pi_{22}, \Omega)$,

$$
\begin{aligned}
&p_{insem}(\beta, \Pi_{22}, \Omega|Y, Z) \propto |\Omega|^{-\frac{1}{2}(T+m+1)}|B_{\perp}\Omega^{-1}B_{\perp}'|^{\frac{1}{2}(k_2-m+1)} \\
&|\begin{pmatrix} B\Omega^{-1}B' \otimes (A + Z'Z)_{22.1} & B\Omega^{-1}e_1 \otimes (A + Z'Z)_{22.1}\Pi_{22} \\ e_1'\Omega^{-1}B' \otimes \Pi_{22}'(A + Z'Z)_{22.1} & e_1'\Omega^{-1}e_1 \otimes \Pi_{22}'(A + Z'Z)_{22.1}\Pi_{22} \end{pmatrix}|^{\frac{1}{2}} \\
&|\Pi_{22\perp}'(A + Z'Z)_{22.1}\Pi_{22\perp}|^{\frac{1}{2}} \exp[-\frac{1}{2}tr(\Omega^{-1}(\tilde{G} + (\Pi_{22}\begin{pmatrix} \beta & I_{m-1} \end{pmatrix} - \tilde{\Pi}_2)' \\
&(A + Z'Z)_{22.1}(\Pi_{22}\begin{pmatrix} \beta & I_{m-1} \end{pmatrix} - \tilde{\Pi}_2)))]. \qquad (24)
\end{aligned}
$$

where $\tilde{\Pi} = \begin{pmatrix} \tilde{\Pi}_1' & \tilde{\Pi}_2' \end{pmatrix}'$, $\tilde{\Pi}_1 : k_1 \times m$, $\tilde{\Pi}_2 : k_2 \times m$.

Also for this posterior it holds that (22) applies and this is used in the following section to construct a posterior simulator.

5 Simulating Posteriors

As mentioned before the posteriors (21) and (24) do not belong to a standard class of probability density functions nor do their conditional posteriors. We can therefore not perform Gibbs sampling as the conditional posteriors are non-standard. In the simulation algorithms constructed in this section, we generate drawings from a probability density function which approximates the true posterior. To correct for not drawing from the true posterior, weights are attached to each drawing of the parameters proportional to the ratio of the posterior and the approximating density in the generated parameter points. These weights can both be used in an Importance, see [24] and [14], and Metropolis-Hastings, see [27] and [17], algorithm to draw from the posterior which explains why we discus them at first. Later on we briefly discuss the two different simulation algorithms.

We use the posterior of the unrestricted SEM, $p_{unsem}(\beta, \lambda, \Pi_{22}, \Omega|Y, Z)$, as approximating density of the posterior of the INSEM, $p_{insem}(\beta, \Pi_{22}, \Omega|Y, Z)$. The posterior of the unrestricted SEM contains the parameter λ, however, which is not present in the posterior of the INSEM. In order to obtain a density which both accords with the posterior of the INSEM and contains λ, we assume that λ is generated given $(\beta, \Pi_{22}, \Omega)$ from a proper conditional density $g(\lambda|\beta, \Pi_{22}, \Omega)$, which we specify ourselves, see [6], [36] and [21]. Furthermore, we assume that β, Π_{22} and Ω are generated from $p_{insem}(\beta, \Pi_{22}, \Omega|Y, Z)$. So, as density function to be approximated by $p_{unsem}(\beta, \lambda, \Pi_{22}, \Omega|Y, Z)$ we have,

$$
g(\lambda|\beta, \Pi_{22}, \Omega)p_{insem}(\beta, \Pi_{22}, \Omega|Y, Z) \propto g(\lambda|\beta, \Pi_{22}, \Omega)(p_{unsem}(\beta, \lambda, \Pi_{22}, \Omega|Y, Z)|_{\lambda=0}). \qquad (25)
$$

The weight function then becomes,

$$w(\beta,\lambda,\Pi_{22},\Omega)=\frac{g(\lambda|\beta,\Pi_{22},\Omega)(p_{unsem}(\beta,\lambda,\Pi_{22},\Omega|Y,Z)|_{\lambda=0})}{p_{unsem}(\beta,\lambda,\Pi_{22},\Omega|Y,Z)}. \tag{26}$$

The quality of the approximating density $p_{unsem}(\beta,\lambda,\Pi_{22},\Omega|Y,Z)$ crucially depends on the chosen specification of $g(\lambda|\beta,\Pi_{22},\Omega)$. In case we use the diffuse prior for the parameters of the INSEM (15), the natural choice of $g(\lambda|\beta,\Pi_{22},\Omega)$ is,

$$\begin{aligned} g(\lambda|\beta,\Pi_{22},\Omega) &= (2\pi)^{-\frac{1}{2}(k_2-m+1)}|B_\perp\Omega^{-1}B_\perp'|^{\frac{1}{2}(k_2-m+1)}|\Pi_{22\perp}'Z_2'M_{Z_1}Z_2\Pi_{22\perp}|^{\frac{1}{2}} \\ &\exp[-\frac{1}{2}tr(B_\perp\Omega^{-1}B_\perp'(\lambda-\hat{\lambda})'\Pi_{22\perp}'Z_2'M_{Z_1}Z_2\Pi_{22\perp}(\lambda-\hat{\lambda}))], \end{aligned} \tag{27}$$

where $\hat{\lambda}=(\Pi_{22\perp}'Z_2'M_{Z_1}Z_2\Pi_{22\perp})^{-1}\Pi_{22\perp}'Z_2'M_{Z_1}Y\Omega^{-1}B_\perp'(B_\perp\Omega^{-1}B_\perp')^{-1}$, while

$$\begin{aligned} g(\lambda|\beta,\Pi_{22},\Omega) &= (2\pi)^{-\frac{1}{2}(k_2-m+1)}|B_\perp\Omega^{-1}B_\perp'|^{\frac{1}{2}(k_2-m+1)}|\Pi_{22\perp}'(A+Z'Z)_{22.1}\Pi_{22\perp}|^{\frac{1}{2}} \\ &\exp[-\frac{1}{2}tr(B_\perp\Omega^{-1}B_\perp'(\lambda-\hat{\lambda})'\Pi_{22\perp}'(A+Z'Z)_{22.1}\Pi_{22\perp}(\lambda-\hat{\lambda}))], \end{aligned} \tag{28}$$

where $\hat{\lambda}=(\Pi_{22\perp}'(A+Z'Z)_{22.1}\Pi_{22\perp})^{-1}\Pi_{22\perp}'(AP+Z'Y)_2\Omega^{-1}B_\perp'(B_\perp\Omega^{-1}B_\perp')^{-1}$, $AP+Z'Y=((AP+Z'Y)_1'\ (AP+Z'Y)_2')'$, $(AP+Z'Y)_1: k_1\times m$, $(AP+Z'Y)_2: k_2\times m$, is the natural choice of $g(\lambda|\beta,\Pi_{22},\Omega)$ when we use the natural conjugate prior (16). The weight function resulting from these choices of g read in both cases,

$$w(\beta,\lambda,\Pi_{22},\Omega)=\frac{|J(\Phi,(\beta,\lambda,\Pi_{22}))|_{\lambda=0}|}{|J(\Phi,(\beta,\lambda,\Pi_{22}))|}g(\lambda|\beta,\Pi_{22},\Omega)|_{\lambda=0}, \tag{29}$$

where $g(\lambda|\beta,\Pi_{22},\Omega)$ should be chosen from (27) and (28) according to the prior involved.

We summarize the different steps involved in obtaining the weight function, attached to the $i-$th drawing, $i=1,...,N$, in a simulation algorithm as follows, see also [21],

- Draw Ω^i from $p_{lin}(\Omega|Y,Z)$

  Draw Φ^i from $p_{lin}(\Phi|\Omega^i,Y,Z)$.
- Perform a singular value decomposition of $\Phi^i=U^iS^iV^{i\prime}$
- Compute β, λ, Π_{22} according to (9)-(10)
- Compute $w(\beta^i,\lambda^i,\Pi_{22}^i,\Omega^i)$ according to (29)
- Draw π_{11}^i, Π_{12}^i from $p(\pi_{11},\Pi_{12}|\Phi(\beta^i,\lambda,\Pi_{22}^i),\Omega^i,Y,Z)|_{\lambda=0}$

The posteriors of the linear model parameters, Ω and Φ, used in the first step, are standard density functions, i.e. inverted-Wishart and matrix normal respectively, in case of diffuse or natural conjugate priors. The exact functional specification of these densities depends on the specification of the involved priors and is straightforward to construct, i.e.,

$$p_{lin}(\Omega|Y,Z) \propto |\Omega|^{-\frac{1}{2}(T+l+m+1)} \exp[-\frac{1}{2}tr(\Omega^{-1}Q)], \quad (30)$$

where $l=0$, $Q=Y'M_Z Y$ in case of the diffuse prior, and $l=h$ and $Q=\tilde{G}$ in case of the natural conjugate prior, and

$$p_{lin}(\Phi|\Omega,Y,Z) \propto |\Omega|^{-\frac{1}{2}k_2} \exp[-\frac{1}{2}tr(\Omega^{-1}(\Phi-\hat{\Phi})'W(\Phi-\hat{\Phi})], \quad (31)$$

where $\hat{\Phi}=(Z_2'M_{Z_1}Z_2)^{-1}Z_2'M_{Z_1}Y$, $W=Z_2'M_{Z_1}Z_2$, in case of the diffuse prior, and $\hat{\Phi}=\tilde{\Pi}_2$, $W=(A+Z'Z)_{22.1}$ in case of the natural conjugate prior.

The weight function can either be used in an Importance or Metropolis-Hastings Sampling algorithm to calculate the marginal posteriors or moments of these. Using the Importance Sampling algorithm, see [24] and [14], we approximate the moment $E(f(\pi_{11},\Pi_{12},\beta,\Pi_{22},\Omega))$ by

$$E(f(\pi_{11},\Pi_{12},\beta,\Pi_{22},\Omega)) = \frac{\sum_{i=1}^N w(\beta^i,\lambda^i,\Pi_{22}^i,\Omega^i) f(\pi_{11}^i,\Pi_{12}^i,\beta^i,\Pi_{22}^i,\Omega^i)}{\sum_{i=1}^N w(\beta^i,\lambda^i,\Pi_{22}^i,\Omega^i)}. \quad (32)$$

In [14] it is shown that under quite general conditions central limit theorems can be used to prove the convergence of the approximation (32) to its true value. Statistics which show the numerical accuracy of the approximation (32) are also constructed using these central limit theorems.

The weights (26) can also be used in a Metropolis-Hastings (M-H) algorithm, see [27] and [17], known as the independence sampler, see [34]. This algorithm constructs a Markov Chain from the drawn $(\pi_{11}^i,\Pi_{12}^i,\beta^i,\Pi_{22}^i,\Omega^i)$'s. The $(\pi_{11}^i,\Pi_{12}^i,\beta^i,\Pi_{22}^i,\Omega^i)$'s in this Markov Chain are accepted as drawings from the posterior. This is achieved using the following steps,

0. $i=1$

1. Draw $(\pi_{11}^{i+1},\Pi_{12}^{i+1},\beta^{i+1},\Pi_{22}^{i+1},\Omega^{i+1})$ using the simulation scheme stated previously. Given that $(\pi_{11}^i,\Pi_{12}^i,\beta^i,\Pi_{22}^i,\Omega^i)$ is accepted as drawing from the posterior, $(\pi_{11}^{i+1},\Pi_{12}^{i+1},\beta^{i+1},\Pi_{22}^{i+1},\Omega^{i+1})$ is accepted as the $(i+1)-$th drawing from the posterior with probability, $\min(\frac{w(\beta^i,\lambda^i,\Pi_{22}^i,\Omega^i)}{w(\beta^{i+1},\lambda^{i+1},\Pi_{22}^{i+1},\Omega^{i+1})},1)$, otherwise $(\pi_{11}^{i+1},\Pi_{12}^{i+1},\beta^{i+1},\Pi_{22}^{i+1},\Omega^{i+1})=(\pi_{11}^i,\Pi_{12}^i,\beta^i,\Pi_{22}^i,\Omega^i)$.

2. $i=i+1$. Go to 1.

When the resulting Markov Chain, $(\pi_{11}^i, \Pi_{12}^i, \beta^i, \Pi_{22}^i, \Omega^i)$, $i = 1, ...;$ has converged to its equillibrium distribution, say after H drawings, we can record $(\pi_{11}^i, \Pi_{12}^i, \beta^i, \Pi_{22}^i, \Omega^i)$, $i = H+1, ...;$ as simulated values of the parameters from the posterior.

These simulation algorithms can also be used to calculate Bayes Factors and Bayesian Lagrange Muliplier Statistics, see [21].

The algorithms can also be used to obtain drawings from the natural conjugate prior (18). In that case, the natural choice of the involved $g(\lambda|\beta, \Pi_{22}, \Omega)$ reads,

$$g(\lambda|\beta, \Pi_{22}, \Omega) = (2\pi)^{-\frac{1}{2}(k_2-m+1)}|B_\perp \Omega^{-1} B_\perp'|^{\frac{1}{2}(k_2-m+1)}|\Pi_{22\perp}' A_{22.1} \Pi_{22\perp}|^{\frac{1}{2}}$$
$$\exp[-\frac{1}{2}tr(B_\perp \Omega^{-1} B_\perp' (\lambda - \hat{\lambda})' \Pi_{22\perp}' A_{22.1} \Pi_{22\perp} (\lambda - \hat{\lambda}))], \tag{33}$$

where $\hat{\lambda} = (\Pi_{22\perp}' A_{22.1} \Pi_{22\perp})^{-1} \Pi_{22\perp}' A_{22.1} P_2 \Omega^{-1} B_\perp' (B_\perp \Omega^{-1} B_\perp')^{-1}$, $P = (P_1' \ P_2')'$, $P_1 : k_1 \times m$, $P_2 : k_2 \times m$, and $p_{lin}(\Omega|Y, Z)$, $p_{lin}(\Phi|\Omega, Y, Z)$, both result from (16). This also shows the conjugateness of this prior as it equals the posterior using a diffuse prior of some arbitrary set of observations which doesnot hold for the extended natural conjugate priors, which are also specifieds for SEMs, used by [10] and [11]. Note that the simulation algorithms do not calculate γ, as $\gamma = \pi_{11} + \Pi_{12}\beta$, we can easily incorporate γ into these algorithms.

6 Full System Analysis

The INSEM is a reduced rank restriction on a parameter matrix of a linear model. A full system analysis of a SEM can also be specified as a linear model with nonlinear restrictions on its parameters. Again these restrictions are reduced rank restrictions but the difference with the INSEM is that they can depend on one another in a recursive way. Theorem 1 states that the reduced form of a SEM is a linear model with reduced rank restrictions on its parameter matrices.

Theorem 1 *Assume that a SEM has the following specification,*

$$\begin{pmatrix} Y_{\bar{m}} & Y_m \end{pmatrix} \begin{pmatrix} B_{\bar{m}\bar{m}} & B_{\bar{m}m} \\ B_{m\bar{m}} & B_{mm} \end{pmatrix} \tag{34}$$
$$= \begin{pmatrix} Z_{\bar{m}} & Z_{\bar{m}m} & Z_m \end{pmatrix} \begin{pmatrix} \Gamma_{\bar{m}\bar{m}} & 0 \\ \Gamma_{m\bar{m}} & \Gamma_{\bar{m}m} \\ 0 & \Gamma_{mm} \end{pmatrix} + \begin{pmatrix} \varepsilon_{\bar{m}} & \varepsilon_m \end{pmatrix}$$

where the number of variables contained in Y_m is chosen such that $\Gamma_{mm} : i_m \times j_m$ $(i_m \geq j_m)$ and $\Gamma_{\bar{m}m} : l_m \times j_m$ are unrestricted, the parameter matrices, $\Gamma_{\bar{m}\bar{m}} : l_{\bar{m}} \times j_{\bar{m}}$, $\Gamma_{m\bar{m}} : l_m \times j_{\bar{m}}$, $B_{\bar{m}\bar{m}} : j_{\bar{m}} \times j_{\bar{m}}$, $B_{\bar{m}m} : j_{\bar{m}} \times j_m$, $B_{m\bar{m}} : j_m \times j_{\bar{m}}$, $B_{\bar{m}\bar{m}} : j_{\bar{m}} \times j_{\bar{m}}$, contain (some) parameters which are restricted to zero except for $B_{\bar{m}\bar{m}}$, which has all diagonal elements equal to one and some offdiagonal elements *equal to zero, and $B_{mm} = I_{j_m}$; then the reduced form of the SEM from equation (34) is equal to a set of reduced rank restrictions on the standard linear model,*

$$\begin{pmatrix} Y_{\bar{m}} & Y_m \end{pmatrix} = \begin{pmatrix} Z_{\bar{m}} & Z_{\bar{m}m} & Z_m \end{pmatrix} \Phi + \xi,$$

where $\Phi : (l_{\bar{m}} + l_m + i_m) \times (j_{\bar{m}} + j_{\bar{m}})$.

Proof: see appendix B.

Theorem 1 shows that we can also use the framework for prior/posterior analysis, constructed in the previous sections, in a full system analysis of a SEM. An important difference with this analysis is, however, the dependence of the different reduced rank restrictions on one another. For the INSEM we can either analyze Φ conditional on (π_{11}, Π_{12}) or vice versa. So, the conditionalization of these parameters on one another does not matter. This does not hold for the full system analysis as we can conclude from the proof of theorem 1. This gives a strict ordering in which the reduced rank restrictions have to be imposed and hence how the parameters have to be analyzed conditional on one another. The reduced form of the SEM constructed in appendix b shows already some important conditionalization rules for the parameters of the SEM. For example, the structural form parameter $\beta_{m\bar{m}}$ is analyzed conditional on the structural form parameter $\beta_{\bar{m}m}$. More of these conditionalization rules will appear when the reduced form is constructed further.

The conditionalization rules also imply rank and order conditions, which can differ from the INSEM based conditions generally used. This is essentially the point made in [25]. Regarding the conditionalization rules, the reduced form, constructed in appendix b, shows that $\beta_{\bar{m}m}$ is identified when $\Pi_{\bar{m}\bar{m}}$ has full rank (or when that part of $\Pi_{\bar{m}\bar{m}}$ which is multiplied by the nonzero parts of $\beta_{\bar{m}m}$ has full rank). When the INSEM based conditions are used, it is assumed that no restrictions are imposed on $\Pi_{\bar{m}\bar{m}}$. If restrictions are imposed, however, the resulting rank and order conditions can become different. In the following, an example of this will be discussed. It can also be seen in $\beta_{m\bar{m}}$, which is identified jointly by $\Pi_{\bar{m}m}$, $\Pi_{\bar{m}\bar{m}}\beta_{\bar{m}m}$ and Π_{mm}, and its rank and order conditions therefore depend on the specification of the SEM.

As mentioned before, the framework for prior/posterior analysis constructed in the previous sections can also be used to construct the posteriors of the parameters in a full system analysis of a SEM. When we apply this framework we have to give an exact specification of the reduced form and its (hyper) parameters reflecting the restrictions which obey the three conditions, that (*i.*) when these (hyper) parameters are nonzero, the model is observationally equivalent with a standard linear model and when these (hyper) parameters are zero, (*ii.*) both the reduced from of the SEM results and (*iii.*) these (hyper) parameters are locally uncorrelated with specific other parameters. This enables us to construct the prior/posterior of the parameters of the SEM as proportional to the prior/posterior of the parameters of the linear model under the restriction that the (hyper) parameters are zero which is identical to the construction of priors/posteriors for the INSEM. As there are differences involved compared to the analysis of the INSEM, because the reduced form has a more complicated structure and the number of additional parameters we have to simulate in the posterior simulator increases, see (25), we give two detailed examples, a two and three (sets of) equation(s) model, to indicate all these differences. These examples jointly with theorem 1 show how a full system analysis of a kind of SEM is to be conducted.

6.1 Two (sets of) equations

We specify the structural form of the two (sets of) equation(s) model by,

$$\begin{aligned} Y_1 &= Y_2\beta_1 + Z_1\Gamma_{11} + Z_2\Gamma_{21} + \varepsilon_1, \\ Y_2 &= Y_1\beta_2 + Z_1\Gamma_{12} + Z_3\Gamma_{32} + \varepsilon_2, \end{aligned} \tag{35}$$

where $Y_1 : T \times m_1$, $Y_2 : T \times m_2$; contain the endogenous variables, $Z_1 : T \times k_1$, $Z_2 : T \times k_2$, $Z_3 : T \times k_3$; contain (weakly) exogenous and lagged dependent variables; $k_2 \geq m_1$, $k_3 \geq m_2$, $m = m_1 + m_2$, $(\varepsilon_1\ \varepsilon_2) \sim n(0, \Sigma \otimes I_T)$, $\beta_1 : m_2 \times m_1$, $\beta_2 : m_1 \times m_2$, $\Gamma_{11} : k_1 \times m_1$, $\Gamma_{12} : k_1 \times m_2$, $\Gamma_{21} : k_2 \times m_1$, $\Gamma_{32} : k_3 \times m_2$. The reduced form of (35), which can be constructed using the proof of theorem 1, reads

$$\begin{aligned} Y_1 &= Z_1\Pi_{11} + Z_2\Pi_{21} + Z_3\Pi_{32}\beta_1 + \xi_1, \\ Y_2 &= Z_1\Pi_{12} + Z_2\Pi_{21}\beta_2 + Z_3\Pi_{32} + \xi_2, \end{aligned} \tag{36}$$

where $\Pi_{11} = (\Gamma_{11} + \Gamma_{12}\beta_1)(I_{m_1} - \beta_2\beta_1)^{-1}$, $\Pi_{21} = \Gamma_{21}(I_{m_1} - \beta_2\beta_1)^{-1}$, $\Pi_{12} = (\Gamma_{12} + \Gamma_{11}\beta_2)(I_{m_2} - \beta_1\beta_2)^{-1}$, $\Pi_{32} = \Gamma_{32}(I_{m_2} - \beta_1\beta_2)^{-1}$, $\xi_1 = (\varepsilon_1 + \varepsilon_2\beta_1)(I_{m_1} - \beta_2\beta_1)^{-1}$, $\xi_2 = (\varepsilon_2 + \varepsilon_1\beta_2)(I_{m_2} - \beta_1\beta_2)^{-1}$, $(\xi_1\ \xi_2) \sim n(0, \Omega \otimes I_T)$, $\Sigma = B'\Omega B$, $B = \begin{pmatrix} I_{m_1} & -\beta_2 \\ -\beta_1 & I_{m_2} \end{pmatrix}$.
Similar to the reduced form of the INSEM (2) and as indicated in the proof of theorem 1, we add parameters to the reduced form to obtain a model, which we call unrestricted SEM (UNSEM), which is observationally equivalent with a linear model and when these added parameters are zero both the reduced form (36) results and the added parameters are locally uncorrelated with specific other parameters,

$$\begin{aligned} \begin{pmatrix} Y_1 & Y_2 \end{pmatrix} &= Z_1 \begin{pmatrix} \Pi_{11} & \Pi_{12} \end{pmatrix} + Z_2\Pi_{21} \begin{pmatrix} I_{m_1} & \beta_2 \end{pmatrix} + Z_3\Pi_{32} \begin{pmatrix} \beta_1 & I_{m_2} \end{pmatrix} \\ &\quad + Z_2\Pi_{21\perp}\lambda_2 \begin{pmatrix} I_{m_1} & \beta_2 \end{pmatrix}_\perp + Z_3\Pi_{32\perp}\lambda_1 \begin{pmatrix} \beta_1 & I_{m_2} \end{pmatrix}_\perp + \begin{pmatrix} \xi_1 & \xi_2 \end{pmatrix}, \end{aligned} \tag{37}$$

where $\lambda_2 : (k_2 - m_1) \times m_2$, $\lambda_3 : (k_3 - m_2) \times m_1$, and the orthogonal complements $\Pi_{21\perp}$, $\Pi_{32\perp}$, $\begin{pmatrix} I_{m_1} & \beta_2 \end{pmatrix}_\perp$ and $\begin{pmatrix} \beta_1 & I_{m_2} \end{pmatrix}_\perp$ are defined similar to the ones used in (6), see appendix c. It is clear that when $\lambda_2 = 0$, $\lambda_3 = 0$, the reduced form (36) results and that λ_2 and λ_3 are locally uncorrelated, when they are equal to zero, with (Π_{21}, β_2) and (Π_{32}, β_1) respectively. When $\lambda_2 \neq 0$, $\lambda_3 \neq 0$, again similar to (6), (37) is observationally equivalent with the linear model,

$$\begin{pmatrix} Y_1 & Y_2 \end{pmatrix} = \begin{pmatrix} Z_1 & Z_2 & Z_3 \end{pmatrix} \begin{pmatrix} \Phi_1 \\ \Phi_2 \\ \Phi_3 \end{pmatrix} + \begin{pmatrix} \xi_1 & \xi_2 \end{pmatrix}, \tag{38}$$

where $\Phi_1 = \begin{pmatrix} \Pi_{11} & \Pi_{12} \end{pmatrix}$, $\Phi_2 : k_2 \times m$, $\Phi_3 : k_3 \times m$. Using a SVD, the equality between (37) and (38) can be shown. SVDs are also used to obtain $(\beta_2, \lambda_2, \Pi_{21})$ from Φ_2 and $(\beta_1, \lambda_3, \Pi_{32})$ from Φ_3, see appendix c. The resulting relationships are similar to (8)-(10) and straightforward to derive. The SEM (35) is consequently a linear model with nonlinear restrictions on its parameters, $\lambda_2 = 0$, $\lambda_3 = 0$. The framework for prior/posterior analysis of the INSEM discussed in sections 3-4 can, therefore, directly be extended to the two equation SEM (35). So, we specify a prior for the parameters of the linear model $(\Phi_1, \Phi_2, \Phi_3, \Omega)$, for example a diffuse or natural conjugate prior, and this implies a prior for the parameters of the SEM (36) as this SEM equals the linear model evaluated in $\lambda_2 = 0$, $\lambda_3 = 0$ (Note that we use the reduced form (36) but this model is observationally equivalent with the SEM (35)),

$$\begin{aligned} & p_{sem}(\Pi_{11}, \Pi_{12}, \beta_1, \beta_2, \Pi_{21}, \Pi_{32}, \Omega) \\ \propto\ & p_{unsem}(\Pi_{11}, \Pi_{12}, \beta_1, \beta_2, \lambda_2, \lambda_3, \Pi_{21}, \Pi_{32}, \Omega)|_{\lambda_2=0,\lambda_3=0} \\ \propto\ & p_{lin}(\Phi_1, \Phi_2(\beta_2, \lambda_2, \Pi_{21}), \Phi_3(\beta_1, \lambda_3, \Pi_{32}), \Omega)|_{\lambda_2=0,\lambda_3=0} \\ & |J(\Phi_2, (\beta_2, \lambda_2, \Pi_{21}))|_{\lambda_2=0}||J(\Phi_3, (\beta_1, \lambda_3, \Pi_{32}))|_{\lambda_3=0}| \end{aligned} \tag{39}$$

where *sem* stands for SEM, *unsem* for UNSEM, and *lin* for linear model and the jacobians $J(\Phi_2, (\beta_2, \lambda_2, \Pi_{21}))$, $J(\Phi_3, (\beta_1, \lambda_3, \Pi_{32}))$ are straightforward to derive given the derivation of the jacobian of the transformation in case of the INSEM and are stated in appendix c. Using (39) and the expressions of diffuse and natural conjugate priors for the linear model, (14) and (16), we can again construct the functional expressions of diffuse and natural conjugate priors for SEMs like (36). For reasons of compactness and similarity with section 3 we do not give the exact functional expressions.

For the posterior exactly the same reasoning as for the prior applies, i.e. the posterior of the parameters of the SEM (36) is proportional to the posterior of the parameters of the linear model under the imposed restriction. We can decompose the posterior of the linear model into a product of marginal and conditional posteriors which belong to a standard class of density functions, i.e. normal or inverted-Wishart, see a.o. [37]. This property can directly be used to decompose the posterior of the SEM,

$$p_{sem}(\Pi_{11}, \Pi_{12}, \beta_1, \beta_2, \Pi_{21}, \Pi_{32}, \Omega|Y, Z) \tag{40}$$

$$
\begin{aligned}
&\propto p_{unsem}(\Pi_{11}, \Pi_{12}, \beta_1, \beta_2, \lambda_2, \lambda_3, \Pi_{21}, \Pi_{32}, \Omega|Y, Z)|_{\lambda_2=0,\lambda_3=0} \\
&\propto p_{lin}(\Phi_1, \Phi_2(\beta_2, \lambda_2, \Pi_{21}), \Phi_3(\beta_1, \lambda_3, \Pi_{32}), \Omega|Y, Z)|_{\lambda_2=0,\lambda_3=0} \\
&\quad |J(\Phi_2, (\beta_2, \lambda_2, \Pi_{21}))|_{\lambda_2=0}||J(\Phi_3, (\beta_1, \lambda_3, \Pi_{32}))|_{\lambda_3=0}| \\
&\propto p_{lin}(\Phi_1|\Phi_2(\beta_2, \lambda_2, \Pi_{21}), \Phi_3(\beta_1, \lambda_3, \Pi_{32}), \Omega, Y, Z)|_{\lambda_2=0,\lambda_3=0} \\
&\quad p_{lin}(\Phi_2(\beta_2, \lambda_2, \Pi_{21})|\Phi_3(\beta_1, \lambda_3, \Pi_{32}), \Omega, Y, Z)|_{\lambda_2=0,\lambda_3=0} \\
&\quad |J(\Phi_2, (\beta_2, \lambda_2, \Pi_{21}))|_{\lambda_2=0}| \\
&\quad p_{lin}(\Phi_3(\beta_1, \lambda_3, \Pi_{32})|\Omega, Y, Z)|_{\lambda_3=0}|J(\Phi_3, (\beta_1, \lambda_3, \Pi_{32}))|_{\lambda_3=0}| \\
&\quad p_{lin}(\Omega|Y, Z).
\end{aligned}
$$

Note that we can also use other orderings in this decomposition. To simulate parameters from the posterior of the SEM (36), we use the decomposition of the posterior of the SEM (40). This allows us to split the simulation in two different steps. Furthermore, we add in the two different steps parameters to the model which we, similar to section 5, assume to be generated from some conditional density g, which we specify ourselves. In case of diffuse priors, the following choices of these functions are the most natural ones,

$$
\begin{aligned}
&g_1(\lambda_3|\beta_1, \Pi_{32}, \Omega) = (2\pi)^{-\frac{1}{2}l_3}|B_{1\perp}\Omega^{-1}B_{1\perp}'|^{\frac{1}{2}l_3}|\Pi_{32\perp}'Z_3'M_{(Z_1\ Z_2)}Z_3\Pi_{32\perp}|^{\frac{1}{2}m_1} \\
&\exp[-\frac{1}{2}tr(B_{1\perp}\Omega^{-1}B_{1\perp}'(\lambda_3-\hat{\lambda}_3)'\Pi_{32\perp}'Z_3'M_{(Z_1\ Z_2)}Z_3\Pi_{32\perp}(\lambda_3-\hat{\lambda}_3))], \qquad (41)
\end{aligned}
$$

$$
\begin{aligned}
&g_2(\lambda_2|\beta_2, \Pi_{21}, \Phi_3, \Omega) = (2\pi)^{-\frac{1}{2}l_2}|B_{2\perp}\Omega^{-1}B_{2\perp}'|^{\frac{1}{2}l_2}|\Pi_{21\perp}'Z_2'M_{Z_1}Z_2\Pi_{21\perp}|^{\frac{1}{2}m_2} \\
&\exp[-\frac{1}{2}tr(B_{2\perp}\Omega^{-1}B_{2\perp}'(\lambda_2-\hat{\lambda}_2)'\Pi_{32\perp}'Z_2'M_{Z_1}Z_3\Pi_{32\perp}(\lambda_2-\hat{\lambda}_2))], \qquad (42)
\end{aligned}
$$

where $l_2 = k_2 - m_1$, $l_3 = k_3 - m_2$, $B_1 = \begin{pmatrix} \beta_1 & I_{m_2} \end{pmatrix}$, $B_2 = \begin{pmatrix} I_{m_1} & \beta_2 \end{pmatrix}$,

$$
\begin{aligned}
\hat{\lambda}_3 &= (\Pi_{32\perp}'Z_3'M_{(Z_1\ Z_2)}Z_3\Pi_{32\perp})^{-1}\Pi_{32\perp}'Z_3'M_{(Z_1\ Z_2)}Y\Omega^{-1}B_{1\perp}'(B_{1\perp}\Omega^{-1}B_{1\perp}')^{-1}, \\
\hat{\lambda}_2 &= (\Pi_{21\perp}'Z_2'M_{Z_1}Z_2\Pi_{21\perp})^{-1}\Pi_{21\perp}'Z_2'M_{Z_1}(Y - Z_3\Phi_3)\Omega^{-1}B_{2\perp}'(B_{2\perp}\Omega^{-1}B_{2\perp}')^{-1}.
\end{aligned}
$$

The weight functions of the two different steps of the simulation algorithm, involving both (41) and (42), then become,

$$
\begin{aligned}
w_1(\beta_1, \lambda_3, \Pi_{32}, \Omega) &= \frac{|J(\Phi_3, (\beta_1, \lambda_3, \Pi_{32}))|_{\lambda_3=0}|}{|J(\Phi_3, (\beta_1, \lambda_3, \Pi_{32}))|} g_1(\lambda_3|\beta_1, \Pi_{32}, \Omega)|_{\lambda_3=0}, \qquad (43) \\
w_2(\beta_2, \lambda_2, \Pi_{21}, \Omega|\Phi_3) &= \frac{|J(\Phi_2, (\beta_2, \lambda_2, \Pi_{21}))|_{\lambda_2=0}|}{|J(\Phi_2, (\beta_2, \lambda_2, \Pi_{21}))|} g_2(\lambda_2|\beta_2, \Pi_{21}, \Phi_3, \Omega)|_{\lambda_2=0} \qquad (44)
\end{aligned}
$$

The different steps involved in obtaining the weight attached to a certain drawing i, $i = 1, ..., N$, of the parameters of the SEM, can then be summarized as follows,

1. Draw Ω^i from $p_{lin}(\Omega|Y,Z)$

   Draw Φ_3^i from $p_{lin}(\Phi_3|\Omega^i,Y,Z)$.

2. Compute β_1^i, λ_3^i, Π_{32}^i from Φ_3^i using a SVD

3. Compute $w_1(\beta_1^i,\lambda_3^i,\Pi_{32}^i,\Omega^i)$ according to (43)

4. Draw Φ_2^i from $p_{lin}(\Phi_2|\Phi_3(\beta_1^i,\lambda_3,\Pi_{32}^i),\Omega^i,Y,Z)|_{\lambda_3=0}$

5. Compute β_2^i, λ_2^i, Π_{21}^i from Φ_2^i using a SVD

6. Compute $w_2(\beta_2^i,\lambda_2^i,\Pi_{21}^i,\Omega^i|\Phi_3(\beta_1^i,\lambda_3,\Pi_{32}^i))|_{\lambda_3=0}$ according to (44)

7. Compute total weight $i-$th drawing :

   $w(\beta_1^i,\lambda_3^i,\Pi_{32}^i,\beta_2^i,\lambda_2^i,\Pi_{21}^i,\Omega^i)=w_1\times w_2$

8. Draw Φ_1^i from $p_{lin}(\Phi_1|\Phi_2(\beta_2^i,\lambda_2,\Pi_{21}^i),\Phi_3(\beta_1^i,\lambda_3,\Pi_{32}^i),\Omega^i,Y,Z)|_{\lambda_2=0,\lambda_3=0}$

The posteriors from which we simulate are all standard, in case of diffuse or natural conjugate priors, and are similar to the ones used in the algorithm in section 5. The values of other structural form parameters can directly be calculated using the equations directly below (36) given the drawings from the algorithm above. The resulting total weights, w, can then be used in an Importance or M-H sampler as discussed in section 5 to obtain a posterior simulator of the posterior of the parameters of the SEM (36).

6.2 Three (sets of) Equations

As an example of a three (sets of) equation(s) model, we use (Note that contrary to the two equation model, the specification of a three equation model is not unique),

$$\begin{aligned} Y_1 &= Y_2\beta_{21}+Z_1\Gamma_{11}+\varepsilon_1, \\ Y_2 &= Y_3\beta_{32}+Z_1\Gamma_{12}+Z_2\Gamma_{22}+\varepsilon_2, \\ Y_3 &= Y_1\beta_{13}+Y_2\beta_{23}+Z_2\Gamma_{23}+Z_3\Gamma_{33}+\varepsilon_3, \end{aligned} \tag{45}$$

where $Y_1\,:\,T\times m_1$, $Y_2\,:\,T\times m_2$, and $Y_3\,:\,T\times m_3$, contain the endogenous variables and $Z_1:T\times k_1$, $Z_2:T\times k_2$, and $Z_3:T\times k_3$, contain lagged endogenous and weakly exogenous variables, $\beta_{21}:m_2\times m_1$, $\beta_{32}:m_3\times m_2$, $\beta_{13}:m_1\times m_3$, $\beta_{23}:m_2\times m_3$, $\Gamma_{11}:k_1\times m_1$, $\Gamma_{12}:k_1\times m_2$, $\Gamma_{22}:k_2\times m_2$, $\Gamma_{23}:k_2\times m_3$, $\Gamma_{33}:k_3\times m_3$, $m=m_1+m_2+m_3$. $(\varepsilon_1\ \varepsilon_2\ \varepsilon_3)\sim n(0,\Sigma\otimes I_T)$. For the model in equation (45) to be properly identified the following (INSEM) order conditions need to be fulfilled, $k_2+k_3\geq m_2$, $k_3\geq m_3$, $k_1\geq m_1+m_2$. Using the proof of theorem 1, the reduced form of the model in equation (45) is constructed and reads,

$$\begin{aligned} Y_1 &= Z_1\Pi_{11} + \begin{pmatrix} Z_1 & Z_2\Pi_{33} \end{pmatrix} \begin{pmatrix} \Pi_{22} \\ \beta_{32} \end{pmatrix} \beta_{21} + \xi_1, \\ Y_2 &= Z_1\Pi_{12} + Z_2\Pi_{22} + Z_3\Pi_{33}\beta_{32} + \xi_2, \\ Y_3 &= Z_1(\Pi_{11}\ \Pi_{12}) \begin{pmatrix} \beta_{13} \\ \beta_{23} \end{pmatrix} + Z_2\Pi_{23} + Z_3\Pi_{33} + \xi_3, \end{aligned} \tag{46}$$

where $(\Gamma_{11}\ \Gamma_{12}) = (\Pi_{11}\ \Pi_{12}) \begin{pmatrix} I_{m_1} & -\beta_{13}\beta_{23} \\ -\beta_{21} & I_{m_2} - \beta_{23}\beta_{13} \end{pmatrix}$, $\Gamma_{33} = \Pi_{33}(I_{m_3} - \beta_2(\beta_1\beta_3 - \beta_4))$, $(\Gamma_{22}\ \Gamma_{23}) = (\Pi_{22}\ \Pi_{23}) \begin{pmatrix} I_{m_2} & -(\beta_{23} + \beta_{21}\beta_{13}) \\ -\beta_{32} & I_{m_3} \end{pmatrix}$, $(\xi_1\ \xi_2\ \xi_3)B, = (\varepsilon_1\ \varepsilon_2\ \varepsilon_3)$, $\Sigma = B'\Omega B$, $B = \begin{pmatrix} I_{m_1} & 0 & -\beta_{13} \\ -\beta_{21} & I_{m_2} & -\beta_{23} \\ 0 & -\beta_{32} & I_{m_3} \end{pmatrix}$. The reduced form in equation (46) is again a system of reduced rank matrices like the reduced forms of the one equation (2) and two equation (36) models. An important difference with these models is that its reduced rank matrices depend on another which is a.o. reflected in the identification of β_{21} which depends on one of the other structural form parameters, β_{32}. This difference also leads to a change in the order condition compared to the INSEM. According to the INSEM order condition, β_{21} is identified when $k_2 + k_3 \geq m_2$, i.e. the number of excluded exogenous variables is at least equal to the number of included endogenous variables, see [18]. The model in equation (45) shows, however, that β_{21} is identified when $\begin{pmatrix} \Pi_{22} \\ \Pi_{33}\beta_{32} \end{pmatrix}$ has full rank. Although this matrix has $k_2 + k_3$ rows, which accords with the standard order condition, it's row rank can never exceed $k_2 + m_3$ ($\leq k_2 + k_3$) as it can be specified as $\begin{pmatrix} I_{k_2} & 0 \\ 0 & \Pi_{33} \end{pmatrix} \begin{pmatrix} \Pi_{22} \\ \beta_{32} \end{pmatrix}$ and the last matrix in this product has $k_2 + m_3$ rows. It is, therefore, important that the identification of the different parameters of a SEM in a full system analysis is conducted using the restricted reduced form parameter matrix instead of the unrestricted one as this can lead to different rank and order conditions, see also [25]. This different order condition results from the dependence of the, by the SEM (46) imposed, reduced rank structures on one another, see also proof of theorem 1. The reduced rank structures appearing in the two equation model do not depend on one another, as can be concluded from (37), and therefore the INSEM order conditions still apply there.

As a consequence of the sequential dependence between the reduced rank structures, not only the order conditions of the INSEM and the SEM (45) differ, as indicated above, but also the parameters which we add to the model (46) to make it observationally equivalent to a linear model are different from the ones we used before, see also the proof of theorem 1. In the cases of the INSEM (6) and the two equation SEM (37), the parameters added to the reduced form, to make it observationally equivalent to a linear model, do not depend on one another in a sequential way. The parameters added to (46) do, however, depend on each other sequentially. To show this consider the linear model,

$$\begin{pmatrix} Y_1 & Y_2 & Y_3 \end{pmatrix} = \begin{pmatrix} Z_1 & Z_2 & Z_3 \end{pmatrix} \begin{pmatrix} \Phi_1 \\ \Phi_2 \\ \Phi_3 \end{pmatrix} + \begin{pmatrix} \xi_1 & \xi_2 & \xi_3 \end{pmatrix}, \tag{47}$$

The model (46) can be obtained by using a, what we call, unrestricted SEM specification of the parameters of (47),

$$\begin{aligned}
\Phi_1 &= \begin{pmatrix} \Pi_{11} & \Pi_{12} \end{pmatrix} \begin{pmatrix} I_{m_1} & 0 & \beta_{13} \\ 0 & I_{m_2} & \beta_{23} \end{pmatrix} \qquad (48) \\
&\quad + \begin{pmatrix} \Pi_{11} & \Pi_{12} \end{pmatrix}_{\perp} \lambda_1 \begin{pmatrix} I_{m_1} & 0 & \beta_{13} \\ 0 & I_{m_2} & \beta_{23} \end{pmatrix}_{\perp}, \\
\begin{pmatrix} \Phi_2 \\ \Phi_3 \end{pmatrix} &= \Theta \begin{pmatrix} \beta_{21} & I_{m_2} & 0 \\ 0 & 0 & I_{m_3} \end{pmatrix} + \Theta_{\perp} \lambda_2 \begin{pmatrix} \beta_{21} & I_{m_2} & 0 \\ 0 & 0 & I_{m_3} \end{pmatrix}_{\perp}, \qquad (49) \\
\Theta &= \begin{pmatrix} \Pi_{22} \;\; \Pi_{23} \\ \Theta_2 \end{pmatrix}, \qquad (50) \\
\Theta_2 &= \Pi_{33} \begin{pmatrix} \beta_{32} & I_{m_3} \end{pmatrix} + \Pi_{33\perp} \lambda_3 \begin{pmatrix} \beta_{32} & I_{m_3} \end{pmatrix}_{\perp} \qquad (51)
\end{aligned}$$

where the orthogonal complements are defined similar to the ones used in (6), see also appendix d, $\lambda_1 : (k_1 - m_1 - m_2) \times m_3$, $\lambda_2 : (k_2 + k_3 - m_2 - m_3) \times m_1$, $\lambda_3 : (k_3 - m_3) \times m_2$. To get a better impression of the implications of the different orthogonality conditions in (48)-(51), we substitute the expression of Θ in $(\Phi_2' \; \Phi_3')'$,

$$\begin{aligned}
\begin{pmatrix} \Phi_2 \\ \Phi_3 \end{pmatrix} &= \begin{pmatrix} \Pi_{22} & \Pi_{23} \\ \Pi_{33} \begin{pmatrix} \beta_{32} & I_{m_3} \end{pmatrix} + \Pi_{33\perp} \lambda_3 \begin{pmatrix} \beta_{32} & I_{m_3} \end{pmatrix}_{\perp} \end{pmatrix} \begin{pmatrix} \beta_{21} & I_{m_2} & 0 \\ 0 & 0 & I_{m_3} \end{pmatrix} \\
&+ \begin{pmatrix} \Pi_{22} & \Pi_{23} \\ \Pi_{33} \begin{pmatrix} \beta_{32} & I_{m_3} \end{pmatrix} + \Pi_{33\perp} \lambda_3 \begin{pmatrix} \beta_{32} & I_{m_3} \end{pmatrix}_{\perp} \end{pmatrix}_{\perp} \lambda_2 \begin{pmatrix} \beta_{21} & I_{m_2} & 0 \\ 0 & 0 & I_{m_3} \end{pmatrix}_{\perp} \qquad (52)
\end{aligned}$$

It is clear from (48)-(51) that when $\lambda_1 = 0$, $\lambda_2 = 0$, $\lambda_3 = 0$, the model (46) results. Furthermore, when $\lambda_1 = 0$, $\lambda_2 = 0$, $\lambda_3 = 0$, λ_1 is locally uncorrelated with $(\Pi_{11}, \Pi_{12}, \beta_{13}, \beta_{23})$, λ_3 with (Π_{33}, β_{32}), and λ_2 with β_{21} and all parameters contained in Θ, i.e. $\Pi_{22}, \Pi_{23}, \Pi_{33}, \lambda_3, \beta_{32}$. SVDs are needed to obtain $(\Pi_{11}, \Pi_{12}, \lambda_1, \beta_{13}, \beta_{23})$ from Φ_1, $(\Theta, \lambda_2, \beta_{21})$ from (Φ_2, Φ_3) and $(\Pi_{22}, \Pi_{23}, \Pi_{33}, \lambda_3, \beta_{32})$ from Θ, and to show the observational equivalence between the model imposed by (48)-(51) and (47) when $\lambda_1 \neq 0$, $\lambda_2 \neq 0$, $\lambda_3 \neq 0$. These SVDs are stated in appendix d. The sequential dependence between the structural form parameters now reflects itselves in SVDs which have to be applied recursively, a.o. of Θ which already results from a SVD as it is the reduced form of an INSEM,

$$\begin{aligned} \tilde{Y}_1 &= \tilde{Y}_2\beta_{32} + \tilde{Z}_1\Delta_{11} + v_1, \\ \tilde{Y}_2 &= \tilde{Z}_1\Delta_{21} + \tilde{Z}_2\Delta_{22} + v_2, \end{aligned} \tag{53}$$

where $\tilde{Y}_1$, $\tilde{Y}_2$, $\tilde{Z}_1$ and $\tilde{Z}_2$ are datamatrices, $\Delta_{21} = \Pi_{23}$, $\Delta_{22} = \Pi_{33}$, $\Delta_{11} = \Pi_{22} - \Delta_{21}\beta_{32}$. Θ is therefore similar to the $(\Pi'_{\bar{m}\bar{m}}\ \Pi'_{m\bar{m}})'$ parameter matrix used in the proof of theorem 1.

So, the SEM (46) is again a linear model with restrictions on its parameters. We can, therefore, again apply the framework for prior/posterior analysis constructed in the previous sections, i.e. we take the prior/posterior of the parameters of (46) as proportional to the prior/posterior of the parameters of the linear model under the condition that the restrictions hold,

$$\begin{aligned} & p_{sem}(\beta_{21}, \beta_{32}, \beta_{13}, \beta_{23}, \Pi_{11}, \Pi_{12}, \Pi_{22}, \Pi_{23}, \Pi_{33}, \Omega) \\ \propto\ & p_{unsem}(\beta_{21}, \beta_{32}, \beta_{13}, \beta_{23}, \lambda_1, \lambda_2, \lambda_3, \Pi_{11}, \Pi_{12}, \Pi_{22}, \Pi_{23}, \Pi_{33}, \Omega)|_{(\lambda_1,\lambda_2,\lambda_3)=0} \\ \propto\ & p_{lin}(\Phi_1(\beta_{13}, \beta_{23}, \lambda_1, \Pi_{11}, \Pi_{12}), \\ & (\Phi_2, \Phi_3)(\beta_{21}, \lambda_2, \Theta(\Pi_{22}, \Pi_{23}, \Pi_{33}, \lambda_3, \beta_{32}), \Omega)|_{(\lambda_1,\lambda_2,\lambda_3)=0} \\ & |J(\Phi_1, (\beta_{13}, \beta_{23}, \lambda_1, \Pi_{11}, \Pi_{12}))|_{\lambda_1=0}||J(\Theta, (\Pi_{22}, \Pi_{23}, \Pi_{33}, \lambda_3, \beta_{32}))|_{\lambda_3=0}| \\ & |J((\Phi_2, \Phi_3), (\beta_{21}, \lambda_2, \Theta(\Pi_{22}, \Pi_{23}, \Pi_{33}, \lambda_3, \beta_{32})|_{\lambda_3=0}))|_{\lambda_2=0}|, \end{aligned} \tag{54}$$

where $J(\Phi_1, (\beta_{13}, \beta_{23}, \lambda_1, \Pi_{11}, \Pi_{12}))$, $J(\Theta, (\Pi_{22}, \Pi_{23}, \Pi_{33}, \lambda_3, \beta_{32}))$, $J((\Phi_2, \Phi_3), (\beta_{21}, \lambda_2, \Theta))$ are the jacobians of the transformation from Φ_1 to $(\beta_{13}, \beta_{23}, \lambda_1, \Pi_{11}, \Pi_{12})$, (Φ_2, Φ_3) to $(\beta_{21}, \lambda_2, \Theta)$ and Θ to $(\Pi_{22}, \Pi_{23}, \Pi_{33}, \lambda_3, \beta_{32})$ and these jacobians are stated in appendix d. When we specify a diffuse (14) or natural conjugate prior (16) for the linear model, (54) shows the implied prior for the SEM. We do not give the exact functional expressions as they can be constructed along the lines of section 3.

Also for the posterior, we can use the framework described previously. Furthermore, we can use the decomposition of the posterior of the linear model into a product of conditional and marginal densities,

$$\begin{aligned} & p_{sem}(\beta_{21}, \beta_{32}, \beta_{13}, \beta_{23}, \Pi_{11}, \Pi_{12}, \Pi_{22}, \Pi_{23}, \Pi_{33}, \Omega|Y, Z) \\ \propto\ & p_{unsem}(\beta_{21}, \beta_{32}, \beta_{13}, \beta_{23}, \lambda_1, \lambda_2, \lambda_3, \Pi_{11}, \Pi_{12}, \Pi_{22}, \Pi_{23}, \Pi_{33}, \Omega|Y, Z)|_{(\lambda_1,\lambda_2,\lambda_3)=0} \\ \propto\ & p_{lin}(\Phi_1(\beta_{13}, \beta_{23}, \lambda_1, \Pi_{11}, \Pi_{12}), \\ & (\Phi_2, \Phi_3)(\beta_{21}, \lambda_2, \Theta(\Pi_{22}, \Pi_{23}, \Pi_{33}, \lambda_3, \beta_{32}), \Omega|Y, Z)|_{(\lambda_1,\lambda_2,\lambda_3)=0} \\ & |J(\Phi_1, (\beta_{13}, \beta_{23}, \lambda_1, \Pi_{11}, \Pi_{12}))|_{\lambda_1=0}||J(\Theta, (\Pi_{22}, \Pi_{23}, \Pi_{33}, \lambda_3, \beta_{32}))|_{\lambda_3=0}| \\ & |J((\Phi_2, \Phi_3), (\beta_{21}, \lambda_2, \Theta(\Pi_{22}, \Pi_{23}, \Pi_{33}, \lambda_3, \beta_{32})|_{\lambda_3=0}))|_{\lambda_2=0}|, \\ \propto\ & p_{lin}(\Phi_1(\beta_{13}, \beta_{23}, \lambda_1, \Pi_{11}, \Pi_{12})|(\Phi_2, \Phi_3)(\beta_{21}, \lambda_2, \Theta), \Omega|Y, Z)|_{(\lambda_1,\lambda_2,\lambda_3)=0} \\ & |J(\Phi_1, (\beta_{13}, \beta_{23}, \lambda_1, \Pi_{11}, \Pi_{12}))|_{\lambda_1=0}| \end{aligned} \tag{55}$$

$$
\begin{aligned}
&p_{lin}((\Phi_2,\Phi_3)(\beta_{21},\lambda_2,\Theta(\Pi_{22},\Pi_{23},\Pi_{33},\lambda_3,\beta_{32})|\Omega,Y,Z)|_{(\lambda_2,\lambda_3)=0}\\
&|J((\Phi_2,\Phi_3),(\beta_{21},\lambda_2,\Theta(\Pi_{22},\Pi_{23},\Pi_{33},\lambda_3,\beta_{32})|_{\lambda_3=0}))|_{\lambda_2=0}|\\
&|J(\Theta,(\Pi_{22},\Pi_{23},\Pi_{33},\lambda_3,\beta_{32}))|_{\lambda_3=0}|\\
&p_{lin}(\Omega|Y,Z).
\end{aligned}
$$

Note that for this model only a few decompositions of the posterior into conditional and marginal posteriors are allowed for, i.e. (Φ_2,Φ_3) given Φ_1 and vice versa as because of the reduced rank structure imposed by the SEM, we cannot for example analyze Φ_2 given Φ_3 or vice versa. We use the decomposition of the posterior (55) to construct a posterior simulator. Again, similar to previous sections, to simulate from the posterior of (46) we add parameters to the model, i.e. $\lambda_1,\lambda_2,\lambda_3$, which we assume to be drawn from a specific conditional density, which we specify ourselves, see (25). In case of a diffuse prior for the linear model (14), natural choices for these conditional densities are,

$$
\begin{aligned}
&g_1(\lambda_1|\beta_{13},\beta_{23},\Pi_{11},\Pi_{12},\Phi_2,\Phi_3,\Omega) &(56)\\
=\ &(2\pi)^{-\frac{1}{2}l_1}|B_{1\perp}\Omega^{-1}B_{1\perp}'|^{\frac{1}{2}l_1}|(\Pi_{11}\ \Pi_{12})_\perp'Z_1'Z_1(\Pi_{11}\ \Pi_{12})_\perp|^{\frac{1}{2}m_3}\\
&\exp[-\frac{1}{2}tr(B_{1\perp}\Omega^{-1}B_{1\perp}'(\lambda_1-\hat{\lambda}_1)'(\Pi_{11}\ \Pi_{12})_\perp'Z_1'Z_1(\Pi_{11}\ \Pi_{12})_\perp(\lambda_1-\hat{\lambda}_1))],\\
&g_2(\lambda_2|\beta_{21},\Theta,\Omega) &(57)\\
=\ &(2\pi)^{-\frac{1}{2}l_2}|B_{2\perp}\Omega^{-1}B_{2\perp}'|^{\frac{1}{2}l_2}|\Theta_\perp'(Z_2\ Z_3)'M_{Z_1}(Z_2\ Z_3)\Theta_\perp|^{\frac{1}{2}m_1}\\
&\exp[-\frac{1}{2}tr(B_{2\perp}\Omega^{-1}B_{2\perp}'(\lambda_2-\hat{\lambda}_2)'\Theta_\perp'(Z_2\ Z_3)'M_{Z_1}(Z_2\ Z_3)\Theta_\perp(\lambda-\hat{\lambda}_2))],\\
&g_3(\lambda_3|\beta_{32},\beta_{21},\Pi_{33},\Omega) &(58)\\
=\ &(2\pi)^{-\frac{1}{2}l_3}|B_{3\perp}B_2\Omega^{-1}B_2'B_{3\perp}'|^{\frac{1}{2}l_3}|\Pi_{33\perp}'Z_3'M_{(Z_1\ Z_2)}Z_3)\Pi_{33\perp}|^{\frac{1}{2}m_2}\\
&\exp[-\frac{1}{2}tr(B_{3\perp}B_2\Omega^{-1}B_2'B_{3\perp}'(\lambda_3-\hat{\lambda}_3)'\Pi_{33\perp}'Z_3'M_{(Z_1\ Z_2)}Z_3)\Pi_{33\perp}(\lambda_3-\hat{\lambda}_3))],
\end{aligned}
$$

where $l_1=k_1-m_1-m_2$, $l_2=k_2+k_3-m_2-m_3$, $l_3=k_3-m_3$, $B_1=\begin{pmatrix} I_{m_1+m_2} & \begin{pmatrix}\beta_{13}\\ \beta_{23}\end{pmatrix}\end{pmatrix}$, $B_2=\begin{pmatrix}\begin{pmatrix}\beta_{21}\\ 0\end{pmatrix} & I_{m_2+m_3}\end{pmatrix}$, $B_3=\begin{pmatrix}\beta_{32} & I_{m_3}\end{pmatrix}$,

$$
\begin{aligned}
\hat{\lambda}_1 &= ((\Pi_{11}\ \Pi_{12})_\perp'Z_1'Z_1(\Pi_{11}\ \Pi_{12})_\perp)^{-1}(\Pi_{11}\ \Pi_{12})_\perp'Z_1'(Y-Z_2\Phi_2-Z_3\Phi_3)\\
&\quad\ \Omega^{-1}B_{1\perp}'(B_{1\perp}\Omega^{-1}B_{1\perp}')^{-1},\\
\hat{\lambda}_2 &= (\Theta_\perp'(Z_2\ Z_3)'M_{Z_1}(Z_2\ Z_3)\Theta_\perp)^{-1}\Theta_\perp'(Z_2\ Z_3)'M_{Z_1}Y\Omega^{-1}B_{2\perp}'(B_{2\perp}\Omega^{-1}B_{2\perp}')^{-1},\\
\hat{\lambda}_3 &= (\Pi_{33\perp}'Z_3'M_{(Z_1\ Z_2)}Z_3)\Pi_{33\perp})^{-1}\Pi_{33\perp}'Z_3'M_{(Z_1\ Z_2)}Y\Omega^{-1}B_2'B_{3\perp}'(B_{3\perp}B_2\Omega^{-1}B_2'B_{3\perp}')^{-1}.
\end{aligned}
$$

As we simulate from a density which approximates the posterior of (46), weight functions are involved in the different steps of the posterior simulator. As we simulate three different parameters, i.e. $\lambda_1,\lambda_2,\lambda_3$, which are not present in the original posterior we want to simulate from, three weight functions are involved,

$$
w_1(\beta_{13},\beta_{23},\lambda_1,\Pi_{11},\Pi_{12},\Omega|\Phi_2,\Phi_3) \qquad (59)
$$

$$= \frac{|J(\Phi_1, (\beta_{13}, \beta_{23}, \lambda_1, \Pi_{11}, \Pi_{12}))|_{\lambda_1=0}|}{|J(\Phi_1, (\beta_{13}, \beta_{23}, \lambda_1, \Pi_{11}, \Pi_{12}))|} g_1(\lambda_1|\beta_{13}, \beta_{23}, \Pi_{11}, \Pi_{12}, \Phi_2, \Phi_3, \Omega)|_{\lambda_1=0}$$

$$w_2(\beta_{21}, \lambda_2, \Theta, \Omega) \tag{60}$$

$$= \frac{|J((\Phi_2, \Phi_3), (\beta_{21}, \lambda_2, \Theta))|_{\lambda_2=0,\lambda_3=0}|}{|J((\Phi_2, \Phi_3), (\beta_{21}, \lambda_2, \Theta))|_{\lambda_3\neq 0}|} g_2(\lambda_2|\beta_{21}, \Theta, \Omega)|_{\lambda_2=0,\lambda_3=0}$$

$$w_3(\beta_{32}, \beta_{21}, \lambda_3, \Pi_{33}, \Omega) \tag{61}$$

$$= \frac{|J(\Theta_2, (\beta_{32}, \lambda_3, \Pi_{33}))|_{\lambda_2=0}|}{|J(\Theta_2, (\beta_{32}, \lambda_3, \Pi_{33}))|} g_3(\lambda_3|\beta_{32}, \beta_{21}, \Pi_{33}, \Omega)|_{\lambda_3=0},$$

where $J(\Phi_1, (\beta_{13}, \beta_{23}, \lambda_1, \Pi_{11}, \Pi_{12}))$, $J((\Phi_2, \Phi_3), (\beta_{21}, \lambda_2, \Theta))$ and $J(\Theta_2, (\beta_{32}, \lambda_3, \Pi_{33}))$ are the jacobians of the different parameter transformations, see appendix d.

The different steps involved in obtaining the weight attached to a certain drawing i, $i = 1, ..., N$, of the parameters of the SEM (46), can then be summarized as follows,

1. Draw Ω^i from $p_{lin}(\Omega|Y, Z)$
   Draw (Φ_2^i, Φ_3^i) from $p_{lin}(\Phi_2, \Phi_3|\Omega^i, Y, Z)$
2. Compute $\beta_{21}^i, \lambda_2^i, \Theta^i$ from (Φ_2^i, Φ_3^i) using SVD
3. Compute $\beta_{32}^i, \lambda_3^i, \Pi_{33}^i$ from Θ_2^i using SVD
4. Compute $w_3(\beta_{32}^i, \beta_{21}^i, \lambda_3^i, \Pi_{33}^i, \Omega^i)$
5. Compute $w_2(\beta_{21}^i, \lambda_2^i, \Theta^i, \Omega^i)$
6. Draw Φ_1^i from $p_{lin}(\Phi_1|\Phi_2^i, \Phi_3^i, \Omega^i, Y, Z)|_{\lambda_2=0,\lambda_3=0}$
7. Compute $\beta_{13}^i, \beta_{23}^i, \lambda_1^i, \Pi_{11}^i, \Pi_{12}^i$ from Φ_1^i
8. Compute $w_1(\beta_{13}^i, \beta_{23}^i, \lambda_1^i, \Pi_{11}^i, \Pi_{12}^i, \Omega^i|\Phi_2^i, \Phi_3^i)|_{\lambda_2=0,\lambda_3=0}$
9. Compute total weight i-th drawing: $w = w_1 \times w_2 \times w_3$

The total weights can be used in an Importance or M-H sampler, as indicated in section 5, to obtain a posterior simulator of the posterior of the parameters of (46).

Jointly with theorem 1, the examples of the two and three structural equations SEMs show how Bayesian analyses of generally specified SEMs are conducted.

7 Conclusions

The traditional Bayesian analysis of SEMs using diffuse priors, as proposed by e.g. [8], [10] and [11], suffers from local nonidentification problems which lead to an a posteriori favor for certain parameter values while it is not the result of information in the prior or data. We therefore constructed a framework in which the priors/posteriors of the parameters of the SEM are proportional to the priors/posteriors of the parameters of the linear model under the condition that the restrictions, imposed by the SEM on the parameters of the linear model, holds. We applied the resulting consistent framework to examples of one, two and three structural equation SEMs, for which expressions of the priors and posteriors are derived jointly with posterior simulators. Using a theorem, which states that the reduced form of any kind of SEM accords with a linear model with reduced rank restrictions of its parameters, the analysis of the examples can be generalized to other specifications of SEMs in a straightforward way.

Using results from [21], we can also construct tools for model comparison like Bayes Factors, Posterior Odds Ratios and Bayesian Lagrange Multiplier statistics. In future work we will construct and apply these procedures to analyze the support for (multiple structural equations) SEMs in practice. It is also interesting to analyze the theoretical properties of the derived posteriors, as for example in [5] where functional expressions are constructed for the marginal posterior of the structural form parameters of the INSEM using a Jeffreys' prior, to investigate the similarities/differences between small sample distributions of classical statistical estimators and the marginal posteriors of the structural form parameters. Both limited information maximum likelihood (LIML) estimators, see [2], and the posteriors of the parameters of the INSEM are namely constructed using SVDs, which correspond with canonical correlations in case of the LIML estimator. So, it is interesting to investigate to what extent these similarities hold further.

Appendix

A. Jacobian of transformation from linear model to INSEM

For the derivation of the Jacobian of the transformation from the linear model parameters to the parameters of the INSEM, it is notationally convenient to conduct this transformation in two steps, (*i.*) from Φ to $(\Pi_{221}, \theta_2, \beta, \lambda)$ where $\theta_2 = \Pi_{222}\Pi_{221}^{-1}$, and (*ii.*) from $(\Pi_{221}, \theta_2, \beta, \lambda)$ to $(\Pi_{221}, \Pi_{222}, \beta, \lambda)$. In the following we construct the jacobians of the two transformations.

We can denote Φ as,

$$\begin{aligned}\Phi &= \begin{pmatrix} \theta & \theta_\perp \end{pmatrix} \begin{pmatrix} \Pi_{221} & 0 \\ 0 & \lambda \end{pmatrix} \begin{pmatrix} B \\ B_\perp \end{pmatrix} \\ &= \theta\Pi_{221}B + \theta_\perp \lambda B_\perp,\end{aligned}$$

where $\theta = (I_{m-1}\ \theta_2')'$, $\theta_\perp = (-\theta_2\ I_{k_2-m+1})'(I_{k_2-m+1} + \theta_2\theta_2')^{-\frac{1}{2}}$, $B = (\beta\ I_{m-1})$, $B_\perp = (1+\beta'\beta)^{-\frac{1}{2}}(1\ -\beta')$. The jacobians of Φ with respect to Π_{221}, θ_2, β and λ then read,

$$\begin{aligned}
J_1 &= \frac{\partial vec(\Phi)}{\partial vec(\Pi_{221})'} = (B' \otimes \theta) \\
J_2 &= \frac{\partial vec(\Phi)}{\partial vec(\theta_2)'} = (B'\Pi_{221}' \otimes I_{k_2})\frac{\partial vec(\theta)}{\partial vec(\theta_2)'} + (B_\perp'\lambda' \otimes I_{k_2})\frac{\partial vec(\theta_\perp)}{\partial vec(\theta_2)'} \\
J_3 &= \frac{\partial vec(\Phi)}{\partial vec(\beta)'} = (I_m \otimes \theta\Pi_{221})\frac{\partial vec(B)}{\partial vec(\beta)'} + (I_m \otimes \theta_\perp\lambda)\frac{\partial vec(B_\perp)}{\partial vec(\beta)'} \\
J_4 &= \frac{\partial vec(\Phi)}{\partial vec(\lambda)'} = (B_\perp' \otimes \theta_\perp)
\end{aligned}$$

where

$$\begin{aligned}
\frac{\partial vec(\theta)}{\partial vec(\theta_2)'} &= \left(I_{m-1} \otimes \begin{pmatrix} 0 \\ I_{k_2-m+1} \end{pmatrix}\right), \\
\frac{\partial vec(\theta_\perp)}{\partial vec(\theta_2)'} &= -(H^{-\frac{1}{2}\prime} \otimes \begin{pmatrix} I_{m-1} \\ 0 \end{pmatrix})K_{k_2-m+1,m-1} + \\
&\quad (I_{k_2-m+1} \otimes \begin{pmatrix} -\theta_2' \\ I_{k_2-m+1} \end{pmatrix})\frac{\partial vec((H^{\frac{1}{2}})^{-1})}{\partial vec(H^{\frac{1}{2}})'}\frac{\partial vec(H^{\frac{1}{2}})}{\partial vec(H)'}\frac{\partial vec(H)}{\partial vec(\theta_2)'}, \\
\frac{\partial vec((H^{\frac{1}{2}})^{-1})}{\partial vec(H^{\frac{1}{2}})'} &= -(H^{-\frac{1}{2}\prime} \otimes H^{-\frac{1}{2}}), \\
\frac{\partial vec(H^{\frac{1}{2}})}{\partial vec(H)'} &= ((I_{k_2-m+1} \otimes H^{\frac{1}{2}})K_{k_2-m+1,k_2-m+1} + (H^{\frac{1}{2}} \otimes I_{k_2-m+1}))^{-1}, \\
\frac{\partial vec(H)}{\partial vec(\theta_2)'} &= (\theta_2 \otimes I_{k_2-m+1}) + (I_{k_2-m+1} \otimes \theta_2)K_{k_2-m+1,m-1},
\end{aligned}$$

$$
\begin{aligned}
\frac{\partial vec(B)}{\partial vec(\beta)'} &= (e_1 \otimes I_{m-1}), \\
\frac{\partial vec(B_\perp)}{\partial vec(\beta)'} &= -(\begin{pmatrix} 0 \\ I_{m-1} \end{pmatrix} \otimes \mathcal{B}^{-\frac{1}{2}})K_{m-1,1} + \\
&\quad (\begin{pmatrix} 1 & -\beta' \end{pmatrix}' \otimes 1)\frac{\partial vec(\mathcal{B}^{-\frac{1}{2}})}{\partial vec(\mathcal{B}^{\frac{1}{2}})'}\frac{\partial vec(\mathcal{B}^{\frac{1}{2}})}{\partial vec(\mathcal{B})'}\frac{\partial vec(\mathcal{B})}{\partial vec(\beta)'}, \\
\frac{\partial vec(\mathcal{B}^{-\frac{1}{2}})}{\partial vec(\mathcal{B}^{\frac{1}{2}})'} &= -(\mathcal{B}^{-\frac{1}{2}\prime} \otimes \mathcal{B}^{-\frac{1}{2}}) = -\mathcal{B}^{-1}, \\
\frac{\partial vec(\mathcal{B}^{\frac{1}{2}})}{\partial vec(\mathcal{B})'} &= ((1 \otimes \mathcal{B}^{\frac{1}{2}\prime}) + (\mathcal{B}^{\frac{1}{2}\prime} \otimes 1)K_{1,1})^{-1} = \frac{1}{2}\mathcal{B}^{-\frac{1}{2}}, \\
\frac{\partial vec(\mathcal{B}^{\frac{1}{2}})}{\partial vec(\beta)'} &= (\beta' \otimes 1)K_{m-1,1} + (1 \otimes \beta') = 2\beta',
\end{aligned}
$$

and $H = I_{k_2-m+1} + \theta_2\theta_2'$, $H^{\frac{1}{2}}H^{\frac{1}{2}\prime} = H$, $\mathcal{B} = (1+\beta'\beta)$, $\mathcal{B}^{\frac{1}{2}\prime}\mathcal{B}^{\frac{1}{2}} = \mathcal{B}$, e_1 is the first m dimensional unity vector, $K_{i,j} : ij \times ij$, are socalled commutation matrices such that for any $W : i \times j$, $vec(W') = K_{i,j}vec(W)$, $vec(W) = K_{j,i}vec(W')$, $K_{j,i} = K_{i,j}'$, see [26]. The jacobian of the transformation from Φ to $(\Pi_{221}, \theta_2, \beta, \lambda)$ then reads,

$$
\frac{\partial vec(\Phi)}{\partial(vec(\Pi_{221})'\ vec(\theta_2)'\ vec(\beta)'\ vec(\lambda)')} = \begin{pmatrix} J_1 & J_2 & J_3 & J_4 \end{pmatrix}.
$$

Since $\theta_2 = \Pi_{222}\Pi_{221}^{-1}$, the jacobians of the transformations from $(\Pi_{221}, \theta_2, \beta, \lambda)$ to Π_{221}, Π_{222}, β, and λ read,

$$
G_1 = \frac{\partial(vec(\Pi_{221})'\ vec(\theta_2)'\ vec(\beta)'\ vec(\lambda)')'}{\partial vec(\Pi_{221})'} = \begin{pmatrix} I_{m-1} \otimes I_{m-1} \\ -\Pi_{221}^{-1\prime} \otimes \Pi_{222}\Pi_{221}^{-1} \\ 0 \\ 0 \end{pmatrix}
$$

$$
G_2 = \frac{\partial(vec(\Pi_{221})'\ vec(\theta_2)'\ vec(\beta)'\ vec(\lambda)')'}{\partial vec(\Pi_{222})'} = \begin{pmatrix} 0 \\ \Pi_{221}^{-1\prime} \otimes I_{k_2-m+1} \\ 0 \\ 0 \end{pmatrix}
$$

$$
G_3 = \frac{\partial(vec(\Pi_{221})'\ vec(\theta_2)'\ vec(\beta)'\ vec(\lambda)')'}{\partial vec(\beta)'} = \begin{pmatrix} 0 \\ 0 \\ 1 \otimes I_{m-1} \\ 0 \end{pmatrix}
$$

$$
G_3 = \frac{\partial(vec(\Pi_{221})'\ vec(\theta_2)'\ vec(\beta)'\ vec(\lambda)')'}{\partial vec(\lambda)'} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \otimes I_{k_2-m+1} \end{pmatrix}
$$

The jacobian of the transformation from Φ to $(\Pi_{22}, \beta, \lambda)$ then becomes,

$$\begin{aligned}
& |J(\Phi, (\Pi_{22}, \beta, \lambda))| \\
= \; & \left|\frac{\partial vec(\Phi)}{\partial(vec(\Pi_{22})' \; vec(\beta)' \; vec(\lambda)')}\right| \\
= \; & \left|\frac{\partial vec(\Phi)}{\partial(vec(\Pi_{221})' \; vec(\theta_2)' \; vec(\beta)' \; vec(\lambda)')}\right| \\
& \left|\frac{\partial(vec(\Pi_{221})' \; vec(\theta_2)' \; vec(\beta)' \; vec(\lambda)')'}{\partial(vec(\Pi_{22})' \; vec(\beta)' \; vec(\lambda)')}\right| \\
= \; & \left|\begin{pmatrix} J_1 & J_2 & J_3 & J_4 \end{pmatrix}\right| \left|\begin{pmatrix} G_1 & G_2 & G_3 & G_4 \end{pmatrix}\right|.
\end{aligned}$$

So,

$$J(\Phi, (\Pi_{22}, \beta, \lambda))|_{\lambda=0} = \begin{pmatrix} B' \otimes I_{k_2} & e_1 \otimes \Pi_{22} & B'_{\perp} \otimes \Pi_{22\perp} \end{pmatrix}$$

B. Proof of theorem 1.

Assume that the reduced form of the SEM,

$$Y_{\bar{m}} B_{\bar{m}\bar{m}} = Z_{\bar{m}} \Gamma_{\bar{m}\bar{m}} + Z_{\bar{m}m} \Gamma_{m\bar{m}} + \varepsilon_{\bar{m}},$$

reads,

$$Y_{\bar{m}} = Z_{\bar{m}} \Pi_{\bar{m}\bar{m}} + Z_{\bar{m}m} \Pi_{m\bar{m}} + \xi_{\bar{m}},$$

where $\Pi_{\bar{m}\bar{m}} = \Gamma_{\bar{m}\bar{m}} B_{\bar{m}\bar{m}}^{-1}$, $\Pi_{m\bar{m}} = \Gamma_{m\bar{m}} B_{\bar{m}\bar{m}}^{-1}$, and this reduced form is equivalent to a set of nonlinear (reduced rank) restrictions on the parameters of a linear model and the (hyper) parameters of this linear model, which are restricted to zero to obtain the reduced form, are locally uncorrelated with specific other parameters.

The parameter matrix of the reduced form of the SEM from theorem 1 reads,

$$\begin{aligned}
& \begin{pmatrix} \Gamma_{\bar{m}\bar{m}} & 0 \\ \Gamma_{m\bar{m}} & \Gamma_{\bar{m}m} \\ 0 & \Gamma_{mm} \end{pmatrix} \begin{pmatrix} B_{\bar{m}\bar{m}} & B_{\bar{m}m} \\ B_{m\bar{m}} & B_{mm} \end{pmatrix}^{-1} \\
= \; & \begin{pmatrix} \Gamma_{\bar{m}\bar{m}} & 0 \\ \Gamma_{m\bar{m}} & \Gamma_{\bar{m}m} \\ 0 & \Gamma_{mm} \end{pmatrix} \begin{pmatrix} B_{\bar{m}\bar{m}}^{-1} + B_{\bar{m}\bar{m}}^{-1} B_{\bar{m}m} B_{mm.\bar{m}}^{-1} B_{m\bar{m}} B_{\bar{m}\bar{m}}^{-1} & -B_{\bar{m}\bar{m}}^{-1} B_{\bar{m}m} B_{mm.\bar{m}}^{-1} \\ -B_{mm.\bar{m}}^{-1} B_{m\bar{m}} B_{\bar{m}\bar{m}}^{-1} & B_{mm.\bar{m}}^{-1} \end{pmatrix} \\
= \; & \begin{pmatrix} \Pi_{\bar{m}\bar{m}}(I_{j_{\bar{m}}} + \beta_{\bar{m}m}\beta_{m\bar{m}}) & -\Pi_{\bar{m}\bar{m}}\beta_{\bar{m}m} \\ \Pi_{m\bar{m}}(I_{j_{\bar{m}}} + \beta_{\bar{m}m}\beta_{m\bar{m}}) - \Pi_{\bar{m}m}\beta_{m\bar{m}} & \Pi_{\bar{m}m} - \Pi_{m\bar{m}}\beta_{\bar{m}m} \\ -\Pi_{mm}\beta_{m\bar{m}} & \Pi_{mm} \end{pmatrix},
\end{aligned}$$

where $\Pi_{\bar{m}\bar{m}} = \Gamma_{\bar{m}\bar{m}} B_{\bar{m}\bar{m}}^{-1}$, $\Pi_{m\bar{m}} = \Gamma_{m\bar{m}} B_{\bar{m}\bar{m}}^{-1}$, $\Pi_{mm} = \Gamma_{mm} B_{mm.\bar{m}}^{-1}$, $\Pi_{\bar{m}m} = \Gamma_{\bar{m}m} B_{mm.\bar{m}}^{-1}$, $B_{\bar{m}\bar{m}.m} = B_{\bar{m}\bar{m}} - B_{\bar{m}m} B_{mm}^{-1} B_{m\bar{m}} = B_{\bar{m}\bar{m}} - B_{\bar{m}m} B_{m\bar{m}}$, $B_{mm.\bar{m}} = B_{mm} - B_{m\bar{m}} B_{\bar{m}\bar{m}}^{-1} B_{\bar{m}m}$, $\beta_{m\bar{m}} = B_{m\bar{m}} B_{\bar{m}\bar{m}}^{-1}$, $\beta_{\bar{m}m} = B_{\bar{m}m} B_{mm.\bar{m}}^{-1}$. This implies, as both Γ_{mm} and $\Gamma_{\bar{m}m}$ are unrestricted, that no restrictions are imposed on Π_{mm} and $\Pi_{\bar{m}m}$. The linear model of which the reduced form is a nonlinear restriction reads,

$$\begin{pmatrix} Y_{\bar{m}} & Y_m \end{pmatrix} = \begin{pmatrix} Z_{\bar{m}} & Z_{\bar{m}m} & Z_m \end{pmatrix} \Phi + \xi,$$

where $\Phi : (l_{\bar{m}} + l_m + i_m) \times (j_{\bar{m}} + j_{\bar{m}})$ and can be specified as,

$$\Phi = \begin{pmatrix} \Phi_{11} & \Phi_{12} \\ \Phi_{21} & \Phi_{22} \\ \Phi_{31} & \Phi_{23} \end{pmatrix},$$

$\Phi_{11} : l_{\bar{m}} \times j_{\bar{m}}$, $\Phi_{21} : l_m \times j_{\bar{m}}$, $\Phi_{31} : i_m \times j_{\bar{m}}$, $\Phi_{12} : l_{\bar{m}} \times j_m$, $\Phi_{22} : l_m \times j_m$, $\Phi_{23} : i_m \times j_m$. To obtain the restrictions on the linear model parameters which result in the reduced form, we specify Φ as,

$$\begin{aligned}
\begin{pmatrix} \Phi_{11} & \Phi_{12} \\ \Phi_{21} & \Phi_{22} \\ \Phi_{31} & \Phi_{23} \end{pmatrix} &= \begin{pmatrix} \Theta_{11} \\ \Theta_{21} \\ 0 \end{pmatrix} \begin{pmatrix} I_{j_{\bar{m}}} & 0 \end{pmatrix} + \begin{pmatrix} \Theta_{12} \\ \Theta_{22} \\ \Pi_{mm} \end{pmatrix} \begin{pmatrix} -\beta_{m\bar{m}} & I_{j_m} \end{pmatrix} \\
&\quad + \begin{pmatrix} 0 \\ 0 \\ \Pi_{mm\perp} \lambda_{mm} \begin{pmatrix} -\beta_{m\bar{m}} & I_{j_m} \end{pmatrix}_{\perp} \end{pmatrix}, \\
\begin{pmatrix} \Theta_{11} & \Theta_{12} \\ \Theta_{21} & \Theta_{22} \end{pmatrix} &= \begin{pmatrix} \Pi_{\bar{m}\bar{m}} \\ \Pi_{m\bar{m}} \end{pmatrix} \begin{pmatrix} I_{j_{\bar{m}}} & -\beta_{\bar{m}m} \end{pmatrix} + \begin{pmatrix} 0 \\ I_{j_m} \end{pmatrix} \Pi_{\bar{m}m} \begin{pmatrix} 0 & I_{j_m} \end{pmatrix} \\
&\quad + \begin{pmatrix} \Pi_{\bar{m}\bar{m}\perp} \lambda_{\bar{m}\bar{m}} \begin{pmatrix} I_{j_{\bar{m}}} & -\beta_{\bar{m}m} \end{pmatrix}_{\perp} \\ 0 \end{pmatrix},
\end{aligned}$$

where $\Theta_{11} : l_{\bar{m}} \times j_{\bar{m}}$, $\Theta_{21} : l_m \times j_{\bar{m}}$, $\Theta_{12} : l_{\bar{m}} \times j_m$, $\Theta_{22} : l_m \times j_m$. It is clear from the chosen specification that when $\lambda_{mm} = 0$, $\lambda_{\bar{m}\bar{m}} = 0$, the reduced form results and that λ_{mm} is locally uncorrelated (when it is zero) with the parameters contained in Π_{mm} and $\beta_{m\bar{m}}$, and $\lambda_{\bar{m}\bar{m}}$ is locally uncorrelated (when it is zero) with the parameters contained in $\Pi_{\bar{m}\bar{m}}$ and $\beta_{\bar{m}m}$. As we can apply the same kind of decomposition on Π_{mm} and $\Pi_{m\bar{m}}$, which we assumed to be possible, and since $\Pi_{\bar{m}m}$ and Π_{mm} are unrestricted, such that there is no need to decompose them further, we can recursively apply the above decomposition and thereby the theorem is proved.

C. Singular Value Decomposition and Jacobians two equation model

For the two equation model, reduced rank restrictions are imposed on the parameter matrices Φ_2 and Φ_3. In the following we state the SVDs and the jacobians involved with these two parameter matrices. We start with Φ_2.

$$\begin{aligned}
\Phi_2 &= \begin{pmatrix} \psi & \psi_{\perp} \end{pmatrix} \begin{pmatrix} \Pi_{211} & 0 \\ 0 & \lambda_2 \end{pmatrix} \begin{pmatrix} B_2 \\ B_{2\perp} \end{pmatrix} \\
&= \psi \Pi_{211} B_2 + \psi_{\perp} \lambda_2 B_{2\perp},
\end{aligned}$$

where $\Pi_{21}=(\Pi_{211}'\ \Pi_{212}')'$, $\Pi_{211}: m_1\times m_1$, $\Pi_{212}:(k_2-m_1)\times m_1$, $\psi_2=\Pi_{212}\Pi_{211}^{-1}$, $\psi=(I_{m_1}\ \psi_2')'$, $\psi_\perp=(-\psi_2\ I_{k_2-m_1})'(I_{k_2-m_1}+\psi_2\psi_2')^{-\frac{1}{2}}$, $B_2=(I_{m_1}\ \beta_2)$, $B_{2\perp}=(I_{m_2}+\beta_2'\beta_2)^{-\frac{1}{2}}(-\beta_2'\ I_{m_2})$. A SVD can be used to obtain these parameters from Φ_2,

$$\Phi_2=\begin{pmatrix} U_{11} & U_{12}\\ U_{21} & U_{22}\end{pmatrix}\begin{pmatrix} S_1 & 0\\ 0 & S_2\end{pmatrix}\begin{pmatrix} V_{11} & V_{12}\\ V_{21} & V_{22}\end{pmatrix}',$$

where $U'U=I_m$; $V'V=I_m$; $U_{11},\ S_1,\ V_{11}: m_1\times m_1$; $S_2,\ V_{22}: m_2\times m_2$; $U_{21}:(k_2-m_1)\times m_1$; $U_{12}: m_1\times m_2$; $U_{22}:(k_2-m_1)\times m_2$; $V_{21},\ V_{12}': m_2\times m_1$; and S_2 contains the smallest m_2 singular values of Φ_2. This leads to the relations,

$$\begin{aligned}\Pi_{211} &= U_{11}S_1V_{11}',\ \psi_2=U_{21}U_{11}^{-1},\\ \beta_2 &= (V_{21}V_{11}^{-1})',\ \lambda_2=(U_{22}'U_{22})^{-\frac{1}{2}}U_{22}S_2V_{22}'(V_{22}'V_{22})^{-\frac{1}{2}}.\end{aligned}$$

The jacobians of Φ_2 with respect to Π_{211}, ψ_2, β_2 and λ_2 read,

$$\begin{aligned}
J_1 &= \frac{\partial vec(\Phi_2)}{\partial vec(\Pi_{221})'}=(B_2'\otimes\psi)\\
J_2 &= \frac{\partial vec(\Phi_2)}{\partial vec(\psi_2)'}=(B_2'\Pi_{211}'\otimes I_{k_2})\frac{\partial vec(\psi)}{\partial vec(\psi_2)'}+(B_{2\perp}'\lambda_2'\otimes I_{k_2})\frac{\partial vec(\psi_\perp)}{\partial vec(\psi_2)'}\\
J_3 &= \frac{\partial vec(\Phi_2)}{\partial vec(\beta_2)'}=(I_m\otimes\psi\Pi_{211})\frac{\partial vec(B_2)}{\partial vec(\beta_2)'}+(I_m\otimes\psi_\perp\lambda_2)\frac{\partial vec(B_{2\perp})}{\partial vec(\beta_2)'}\\
J_4 &= \frac{\partial vec(\Phi_2)}{\partial vec(\lambda_2)'}=(B_{2\perp}'\otimes\psi_\perp)
\end{aligned}$$

where

$$\begin{aligned}
\frac{\partial vec(\psi)}{\partial vec(\psi_2)'} &= \left(I_{m_1}\otimes\begin{pmatrix}0\\ I_{k_2-m_1}\end{pmatrix}\right),\\
\frac{\partial vec(\psi_\perp)}{\partial vec(\psi_2)'} &= -\left(H^{-\frac{1}{2}\prime}\otimes\begin{pmatrix}I_{m_1}\\ 0\end{pmatrix}\right)K_{k_2-m_1,m_1}+\\
&\quad (I_{k_2-m_1}\otimes\begin{pmatrix}-\psi_2'\\ I_{k_2-m_1}\end{pmatrix})\frac{\partial vec(H^{-\frac{1}{2}})}{\partial vec(H^{\frac{1}{2}})'}\frac{\partial vec(H^{\frac{1}{2}})}{\partial vec(H)'}\frac{\partial vec(H)}{\partial vec(\psi_2)'},\\
\frac{\partial vec(H^{-\frac{1}{2}})}{\partial vec(H^{\frac{1}{2}})'} &= -(H^{-\frac{1}{2}\prime}\otimes H^{-\frac{1}{2}}),\\
\frac{\partial vec(H^{\frac{1}{2}})}{\partial vec(H)'} &= ((I_{k_2-m_1}\otimes H^{\frac{1}{2}})K_{k_2-m_1,k_2-m_1}+(H^{\frac{1}{2}}\otimes I_{k_2-m_1}))^{-1},\\
\frac{\partial vec(H)}{\partial vec(\psi_2)'} &= (\psi_2\otimes I_{k_2-m_1})+(I_{k_2-m_1}\otimes\psi_2)K_{k_2-m_1,m_1},\\
\frac{\partial vec(B_2)}{\partial vec(\beta_2)'} &= (\begin{pmatrix}0\\ I_{m_2}\end{pmatrix}\otimes I_{m_1}),
\end{aligned}$$

$$\frac{\partial vec(B_{2\perp})}{\partial vec(\beta_2)'} = -(\begin{pmatrix} I_{m_1} \\ 0 \end{pmatrix} \otimes \mathcal{B}^{-\frac{1}{2}})K_{m_1,m_2} + $$
$$((\begin{array}{cc} -\beta_2' & I_{m_2} \end{array})' \otimes I_{m_2})\frac{\partial vec(\mathcal{B}^{-\frac{1}{2}})}{\partial vec(\mathcal{B}^{\frac{1}{2}})'}\frac{\partial vec(\mathcal{B}^{\frac{1}{2}})}{\partial vec(\mathcal{B})'}\frac{\partial vec(\mathcal{B})}{\partial vec(\beta)'},$$
$$\frac{\partial vec(\mathcal{B}^{-\frac{1}{2}})}{\partial vec(\mathcal{B}^{\frac{1}{2}})'} = -(\mathcal{B}^{-\frac{1}{2}\prime} \otimes \mathcal{B}^{-\frac{1}{2}}),$$
$$\frac{\partial vec(\mathcal{B}^{\frac{1}{2}})}{\partial vec(\mathcal{B})'} = ((I_{m_2} \otimes \mathcal{B}^{\frac{1}{2}\prime}) + (\mathcal{B}^{\frac{1}{2}\prime} \otimes I_{m_2})K_{m_2,m_2})^{-1},$$
$$\frac{\partial vec(\mathcal{B}^{\frac{1}{2}})}{\partial vec(\beta_2)'} = (\beta_2' \otimes I_{m_2})K_{m_1,m_2} + (I_{m_2} \otimes \beta_2'),$$

and $H = I_{k_2-m_1} + \psi_2\psi_2'$, $H^{\frac{1}{2}}H^{\frac{1}{2}\prime} = H$, $\mathcal{B} = (I_{m_2} + \beta_2'\beta_2)$, $\mathcal{B}^{\frac{1}{2}\prime}\mathcal{B}^{\frac{1}{2}} = \mathcal{B}$. The jacobian of the transformation from Φ_2 to $(\Pi_{211}, \psi_2, \beta_2, \lambda_2)$ then reads,

$$\frac{\partial vec(\Phi_2)}{\partial(vec(\Pi_{211})'\ vec(\psi_2)'\ vec(\beta_2)'\ vec(\lambda_2)')} = \begin{pmatrix} J_1 & J_2 & J_3 & J_4 \end{pmatrix}.$$

Since $\psi_2 = \Pi_{212}\Pi_{211}^{-1}$, the jacobians of the transformations from $(\Pi_{211}, \psi_2, \beta_2, \lambda_2)$ to Π_{211}, Π_{212}, β_2, and λ_2 read,

$$G_1 = \frac{\partial(vec(\Pi_{211})'\ vec(\psi_2)'\ vec(\beta_2)'\ vec(\lambda_2)')'}{\partial vec(\Pi_{211})'} = \begin{pmatrix} I_{m_1} \otimes I_{m_1} \\ -\Pi_{211}^{-1\prime} \otimes \Pi_{212}\Pi_{211}^{-1} \\ 0 \\ 0 \end{pmatrix}$$

$$G_2 = \frac{\partial(vec(\Pi_{211})'\ vec(\psi_2)'\ vec(\beta_2)'\ vec(\lambda_2)')'}{\partial vec(\Pi_{212})'} = \begin{pmatrix} 0 \\ \Pi_{211}^{-1\prime} \otimes I_{k_2-m_1} \\ 0 \\ 0 \end{pmatrix}$$

$$G_3 = \frac{\partial(vec(\Pi_{211})'\ vec(\psi_2)'\ vec(\beta_2)'\ vec(\lambda_2)')'}{\partial vec(\beta_2)'} = \begin{pmatrix} 0 \\ 0 \\ I_{m_2} \otimes I_{m_1} \\ 0 \end{pmatrix}$$

$$G_3 = \frac{\partial(vec(\Pi_{211})'\ vec(\psi_2)'\ vec(\beta_2)'\ vec(\lambda_2)')'}{\partial vec(\lambda_2)'} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ I_{m_2} \otimes I_{k_2-m_1} \end{pmatrix}$$

The jacobian of the transformation from Φ_2 to $(\Pi_{21}, \beta_2, \lambda_2)$ then becomes,

$$|J(\Phi_2, (\Pi_{21}, \beta_2, \lambda_2))|$$
$$= |\frac{\partial vec(\Phi_2)}{\partial(vec(\Pi_{21})'\ vec(\beta_2)'\ vec(\lambda_2)')}|$$

$$
\begin{aligned}
&= \left|\frac{\partial vec(\Phi_2)}{\partial(vec(\Pi_{211})'\ vec(\psi_2)'\ vec(\beta_2)'\ vec(\lambda_2)')}\right| \\
&\quad \left|\frac{\partial(vec(\Pi_{211})'\ vec(\psi_2)'\ vec(\beta_2)'\ vec(\lambda_2)')'}{\partial(vec(\Pi_{21})'\ vec(\beta_2)'\ vec(\lambda_2)')}\right| \\
&= \left|\begin{pmatrix} J_1 & J_2 & J_3 & J_4 \end{pmatrix}\right|\left|\begin{pmatrix} G_1 & G_2 & G_3 & G_4 \end{pmatrix}\right|.
\end{aligned}
$$

The specification of Φ_3 reads,

$$
\Phi_3 = \begin{pmatrix} \theta & \theta_\perp \end{pmatrix}\begin{pmatrix} \Pi_{321} & 0 \\ 0 & \lambda_3 \end{pmatrix}\begin{pmatrix} B_1 \\ B_{1\perp} \end{pmatrix},
$$

where $\theta = (I_{m_2}\ \theta_2')'$, $B_1 = (\beta_1\ I_{m_2})$, $\Pi_{32} = (\Pi_{321}'\ \Pi_{322}')'$, $\Pi_{321} : m_2 \times m_2$, $\Pi_{322} : (k_3 - m_2) \times m_2$, $\theta_2 = \Pi_{322}\Pi_{321}^{-1}$. So, the specification of Φ_3 is identical to the specification of Φ for the INSEM. The parameters $(\Pi_{32}, \beta_1, \lambda_3)$ can therefore be obtained using the SVDs (8)-(10) and changing the sizes of the involved matrices, i.e. k_2 to k_3, $m-1$ to m_2, 1 to m_1. Also the jacobian involved in the parameter transformation of the INSEM is identical to the jacobian in case of Φ_3 when we change the sizes of the involved matrices in the outlined manner.

D. Singular Value Decomposition and Jacobians three equation model

For the three equation model, reduced rank restrictions are imposed on the parameter matrices $(\Phi_2'\ \Phi_3')'$, Θ and Φ_1. The important difference with the INSEM and the two equation model lies in Θ which itself already results from a reduced rank restriction. As we have to analyze Θ given $(\Phi_2'\ \Phi_3')'$, we start with the SVD and jacobian involved with $(\Phi_2'\ \Phi_3')'$. The specification of $(\Phi_2'\ \Phi_3')'$ reads,

$$
\begin{pmatrix} \Phi_2 \\ \Phi_3 \end{pmatrix} = \Theta\begin{pmatrix} \beta_{21} & I_{m_2} & 0 \\ 0 & 0 & I_{m_3} \end{pmatrix} + \Theta_\perp\lambda_2\begin{pmatrix} \beta_{21} & I_{m_2} & 0 \\ 0 & 0 & I_{m_3} \end{pmatrix}_\perp.
$$

This implies that when $\Phi_2 = (\Phi_{21}\ \Phi_{22})$, $\Phi_{21} : k_2 \times (m_1 + m_2)$, $\Phi_{22} : k_2 \times m_3$; $\Phi_3 = (\Phi_{31}\ \Phi_{32})$, $\Phi_{31} : k_3 \times (m_1 + m_2)$, $\Phi_{32} : k_3 \times m_3$; $\Theta = \begin{pmatrix} \Theta_{11} & \Theta_{12} \\ \Theta_{21} & \Theta_{22} \end{pmatrix}$, $\Theta_{11} : k_2 \times (m_1 + m_2)$, $\Theta_{12} : k_2 \times m_3$, $\Theta_{21} : k_3 \times (m_1 + m_2)$, $\Theta_{22} : k_3 \times m_3$; that the following equality holds,

$$
\begin{pmatrix} \Theta_{12} \\ \Theta_{22} \end{pmatrix} = \begin{pmatrix} \Phi_{22} \\ \Phi_{32} \end{pmatrix}.
$$

and we are left with,

$$
\begin{pmatrix} \Phi_{21} \\ \Phi_{31} \end{pmatrix} = \begin{pmatrix} \Theta_{11} \\ \Theta_{21} \end{pmatrix}\begin{pmatrix} \beta_{21} & I_{m_2} \end{pmatrix} + \begin{pmatrix} \Theta_{11} \\ \Theta_{21} \end{pmatrix}_\perp \lambda_2 \begin{pmatrix} \beta_{21} & I_{m_2} \end{pmatrix}_\perp,
$$

which is again identical to the specification of Φ for the INSEM such that when we change the sizes of the matrices in the appropriate manner, i.e. k_2 to k_2+k_3, $m-1$ to m_2 and 1 to m_3, we can directly use the SVDs and jacobians for Φ of the INSEM.

The SVDs and jacobians for Θ_2 are constructed using (50) and (51),

$$\begin{pmatrix} \Theta_{21} & \Theta_{22} \end{pmatrix} = \Pi_{33} \begin{pmatrix} \beta_{32} & I_{m_3} \end{pmatrix} + \Pi_{33\perp} \lambda_3 \begin{pmatrix} \beta_{32} & I_{m_3} \end{pmatrix}_\perp .$$

Again this specification is identical to the specification of Φ for the INSEM such that we can use the SVD and jacobians specified for the INSEM when we change the sizes of the matrices in the appropriate manner, i.e. k_2 to k_3, $m-1$ to m_3 and 1 to m_2.

The specification of Φ_1 reads,

$$\begin{aligned} \Phi_1 &= \begin{pmatrix} \Pi_{11} & \Pi_{12} \end{pmatrix} \begin{pmatrix} I_{m_1+m_2} & \begin{pmatrix} \beta_{13} \\ \beta_{23} \end{pmatrix} \end{pmatrix} + \\ & \begin{pmatrix} \Pi_{11} & \Pi_{12} \end{pmatrix}_\perp \lambda_1 \begin{pmatrix} I_{m_1+m_2} & \begin{pmatrix} \beta_{13} \\ \beta_{23} \end{pmatrix} \end{pmatrix}_\perp . \end{aligned}$$

This specification is identical to the specification of Φ_2 in the two equation model such that we can use the jacobians and the SVD listed there when we change the sizes of the matrices in the appropriate manner, i.e. k_2 to k_1, m_1 to m_1+m_2 and m_2 to m_3.

References

[1] Anderson, T.W., 1982, Some Recent Developments on the Distributions of Single Equation Estimators, in W. Hildebrand, ed., *Advances in Econometrics*, North-Holland Publishing Co., Amsterdam

[2] Anderson, T.W., and H. Rubin, 1949, Estimation of the Parameters of a Single Equation in a Complete Set of Stochastic Equations, *Annals of Mathematical Statistics,* 21, 570-582

[3] Bauwens, L., and H.K. van Dijk, 1989, Bayesian Limited Information Analysis Revisited, in. J.J. Gabszewicz et. al., eds., *Economic Decision-Making : Games, Econometrics and Optimisation,* North-Holland Publishing Co., Amsterdam

[4] Billingsley, P., 1986, *Probability and Measure,* Wiley, New York

[5] Chao, J.C., and P.C.B. Phillips, 1996, *Bayesian Posterior Distributions in Limited Information Analysis of the Simultaneous Equation Model using the Jeffreys' Prior,* Unpublished Manuscript, Cowles Foundation, Yale University

[6] Chen, M-H., 1994, Importance-Weighted Marginal Bayesian Posterior Density Estimation, *Journal of the American Statistical Association,* 89, 818-824

[7] Dickey, J., 1971, The Weighted Likelihood Ratio, Linear Hypothesis on Normal Location Parameters, *The Annals of Mathematical Statistics,* 42, 223-240

[8] Drèze, J.H., 1976, Bayesian Limited Information Analysis of the Simultaneous Equations Model, *Econometrica,* 44, 1045-1075

[9] Drèze, J.H., 1977, Bayesian Regression using poly t densities, *Journal of Econometrics,* 6, 329-354

[10] Drèze, J.H., and J.A. Morales, 1976, Bayesian Full Information Analysis of Simultaneous Equations, *Journal of the American Statistical Association,* 71, 329-354

[11] Drèze, J.H., and J.F. Richard, 1983, Bayesian Analysis of Simultaneous Equations Systems, in: Z. Griliches and M.D. Intrilligator, eds., *Handbook of Econometrics,* Vol.1., North-Holland Publishing Co., Amsterdam

[12] Engle, R.F., D.F. Hendry and J.F. Richard, 1983, Exogeneity, *Econometrica,* 51, 277-304

[13] Gelfand, A.E., and A.F.M. Smith, 1989, Sampling based Approaches to calculating Marginal Densities, *Journal of the American Statistical Association,* 85, 398-409

[14] Geweke, J., 1989, Bayesian Inference in Econometric Models using Monte-Carlo Integration, *Econometrica,* 57, 1317-1339

[15] Geweke, J., 1996, Bayesian Reduced Rank Regression in Econometrics, *Journal of Econometrics,* 75, 63-86

[16] Haavelmo, T., 1943, The Statistical Implications of a System of Simultaneous Equations, *Econometrica,* 11, 1-12

[17] Hastings, W.K., 1970, Monte Carlo Sampling using Markov Chains and their applications, *Biometrika,* 57, 97-109

[18] Hausman, J.A., 1983, Specification and Estimation of Simultaneous Equations Systems, in: Z. Griliches and M.D. Intrilligator, eds., *Handbook of Econometrics,* Vol.1., North-Holland Publishing Co., Amsterdam

[19] Kleibergen, F., 1996, Equality Restricted Random Variables: Densities and Sampling Algorithms, *Econometric Institute Report 9662/A,* Erasmus University Rotterdam, available at http://www.eur.nl/few/ei/faculty/kleibergen.

[20] Kleibergen, F. and H. Hoek, 1996, Bayesian Analysis of ARMA models using Noninformative Priors, *Econometric Institute Report 9562/A,* Erasmus University Rotterdam, available at http://www.eur.nl/few/ei/faculty/kleibergen.

[21] Kleibergen, F. and R. Paap, 1997, Priors, Posterior Odds and Lagrange Multiplier Statistics in Bayesian Analyses of Cointegration, *Econometric Institute Report 9668/A,* Erasmus University Rotterdam, available at http://www.eur.nl/few/ei/faculty/kleibergen.

[22] Kleibergen, F.R., and H.K. van Dijk, 1994, Bayesian Simultaneous Equation Model analysis: On the Existence of Structural Posterior Moments, *Tinbergen Institute Discussion Papers,* TI 93-25.

[23] Kleibergen, F.R., and H.K. van Dijk, 1994, On the Shape of the Likelihood/Posterior in Cointegration Models, *Econometric Theory,* 10, 514-552

[24] 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

[25] Maddala, G.S., 1976, Weak Priors and Sharp Posteriors in Simultaneous Equation Models, *Econometrica,* 44, 345-351

[26] Magnus, J.R. and H. Neudecker, 1988, *Matrix Differential Calculus with Applications in Statistics and Economics,* Wiley & Sons, Chichester

[27] Metropolis, N., A.W. Rosenbluth, M.N. Rosenbluth, A.H. Teller and E. Teller, 1953, Equations of State Calculations by fast computing Machines, *Journal of Chemical Physics,* 21, 1087-1092

[28] Phillips, P.C.B., 1983, Exact Small Sample Theory in the Simultaneous Equations Model, in: Z. Griliches and M.D. Intrilligator, eds., *Handbook of Econometrics,* Vol.1., North-Holland Publishing Co., Amsterdam

[29] Phillips, P.C.B., 1989, Partially Identified Econometric Models, *Econometric Theory,* 5, 181-240

[30] Poirier, D., 1994, *Prior Beliefs about Fit,* Unpublished Manuscript, University of Toronto

[31] Richard, J.-F. and H. Tompa, 1980, On the Evaluation of 1-1 Poly t density functions, *Journal of Econometrics,* 12, 335-351

[32] Roberts, G.O. and A.F.M. Smith, 1994, Simple Conditions for the convergence of the Gibbs Sampler and Metropolis-Hastings Algorithms, *Stochastic Processes and their Applications,* 49, 205-216

[33] Smith, A.F.M., and G.O. Roberts, 1993, Bayesian Computation via the Gibbs Sampler and related Markov Chain Methods, *Journal of the Royal Statistical Society B,* 55, 3-23

[34] Tierney, L., 1994, Markov Chains for exploring Posterior Distributions, *Annals of Statistics,* 22, 1701-1762

[35] Tintner, G., 1952, *Econometrics,* Wiley, New York

[36] Verdinelli, I. and L. Wasserman, 1995, Computing Bayes Factors using a Generalization of the Savage-Dickey Density Ratio, *Journal of the American Statistical Association,* 90, 614-618

[37] Zellner, A., 1971, *An Introduction to Bayesian Inference in Econometrics,* Wiley, New York

[38] Zellner, A., L. Bauwens, and H.K. van Dijk, 1988, Bayesian Specification Analysis and Estimation of Simultaneous Equation Models using Monte-Carlo Integration, *Journal of Econometrics,* 38, 39-72