Posterior inference on the degrees of freedom parameter in multivariate-t regression models

Chib, S.; Osiewalski, J.; Steel, M.F.J.

Publication date:
1990

CentER
for
omic Research

Discussion paper

[redacted]

TILBURG UNIVERSITY

No. 9043

POSTERIOR INFERENCE ON THE DEGREES OF FREEDOM PARAMETER IN MULTIVARIATE-t REGRESSION MODELS

by Siddharta Chib,
Jacek Osiewalski and Mark Steel

August 1990

ISSN 0924-7815

POSTERIOR INFERENCE ON THE DEGREES OF FREEDOM

PARAMETER IN MULTIVARIATE-t REGRESSION MODELS

Siddhartha Chib*
235 Professional Bldg
Department of Economics
University of Missouri
Columbia, MO 65211

Jacek Osiewalski
Academy of Economics
Kraków, Poland, and
CORE, Université Catholique
de Louvain, Belgium

Mark F. J. Steel
Department of Econometrics
Tilburg University
Tilburg, Netherlands

Abstract: This paper considers the nonlinear regression model with errors that are distributed as scale mixtures of Normals. Specifically, we focus attention on the multivariate Student-t distribution with ν degrees of freedom. We provide general conditions on the overall prior structure (either proper or improper) under which the prior of ν is not updated by the sample information. In addition, we propose several prior families that allow revision of our prior opinions on ν. A number of examples are used to illustrate our findings.

Keywords: Bayesian updating; degrees of freedom parameter; elliptical models.

JEL Number: 211.

Headnote: See Abstract.

April 1990

---

* Please direct all correspondence to this author.

1. INTRODUCTION

In regression analysis it is usually important to study the consequences of non-Normal error distributions. The early work of Zellner (1976) shows that the multivariate-t (MVt) distribution provides a useful alternative to the multivariate Normal (MVN). The fact that several distributions, including the MVt, can be expressed as scale mixtures of the MVN has been exploited in some recent studies. See Phillips (1988), Chib, Tiwari and Jammalamadaka (1988), Osiewalski (1990), Osiewalski and Steel (1990), and the references therein. As a result much is now known about the robustness of inferences to departures from the MVN error distribution. In the Bayesian context, for example, it has been shown that under some restrictions on the prior, the marginal posterior of the regression parameter, β, is unaffected by the MVt assumption. This is the conclusion of Zellner (1976), generalized further by Chib et al. (1988) and Osiewalski (1990). Other invariance results are also obtained in the papers cited above.

Most of the papers that have adopted the MVt framework have made one important assumption, i.e., the degrees of freedom, ν, of the MVt error distribution is known. If this assumption is relaxed, the maximum likelihood method cannot be used to estimate ν (cf. Zellner, 1976). However, a method of moments estimator does exist, and has been provided by Singh (1988).

In this paper we pursue the unknown ν from an entirely Bayesian standpoint and show that the prior-posterior analysis for ν must proceed with some care. For example, if we adopt the usual Jeffreys' prior for the error precision, τ^2, and exclude prior links between ν and the other parameters of the model, the posterior of ν is the same as the prior, for any data. In fact, the same outcome is obtained for some other priors. Thus, it becomes necessary to define classes of priors for which the posterior is different from the prior. Three useful families of priors are provided and it is shown that the updated moments of ν can usually be extracted by bivariate numerical integration.

A few important points should be noted. For the most part, the analysis focuses on ν, deferring the implications of our study for the posteriors of β and τ^2 to a companion paper. Next, the results are derived under fairly weak restrictions for a general nonlinear model with an unknown covariance matrix.

Finally, some comments about the notation that is used. The symbol p is used generically to denote density functions whether they be marginal, joint or conditional. Next, if $w \sim N_n(\mu,\Sigma)$, a n-variate Normally distributed random variable with mean μ and covariance Σ, we denote its pdf by $f_N(w|\mu,\Sigma)$ where the dimension of w indicates that this is a n-variate pdf. Similarly, if $w \sim MVt_n(\nu,\mu,\Sigma)$, a n-variate t distribution with degrees of freedom ν, location μ, and dispersion Σ, its pdf is $f_T(w|\nu,\mu,\Sigma)$. Conditional independence is denoted by $a \perp\!\!\!\perp b|c$, and is read "a is independent of b given c."

2. MAIN RESULTS

2.1 MODEL AND PRELIMINARIES

Consider the (linear or nonlinear) scale mixture of Normals regression model in which a n-vector of observations y satisfies

$$y = h(X;\beta) + \xi,$$

$$\xi = \psi(z,\nu)u, \quad u|X,\beta,\eta,\tau^2 z,\nu \sim N_n(u|0,\tau^{-2}V(X,\eta)) , \tag{2.1}$$

where X: nxr is a (possibly stochastic) set of regressors, β is the regression coefficient vector, $h(X,\beta)$ is a vector function of (X,β), and ξ is an elliptically distributed error vector which, given $(X,\beta,\eta,\tau^2,\nu,z)$, is distributed as a n-variate normal. More specifically, we make the following assumptions about the quantities in (2.1):

A1) The parameter space of $\omega = (\beta, \eta, \tau^2, \nu) = (\theta, \tau^2, \nu)$ is $\Omega = B \times H \times R_+ \times N$, where $B \subseteq R^k$, $H \subseteq R^q$, and $N \subseteq R^\ell$.

A2) $h(X,\beta)$: $n \times 1$ is a known function of X and $\beta \in B$.

A3) $V(X,\eta)$: $n \times n$ is a positive definite matrix, and a known matrix function of X and $\eta \in H$.

A4) z is a positive random variable with conditional distribution, $G(z|\omega)$ with density $p(z|\omega)$.

A5) $\psi(z,\nu)$ is positive and continuously differentiable wrt z and ν.

A6) The prior density $p(\omega)$ is such that $p(y|X) = \int p(y|X,\omega,z)p(z|\omega)p(\omega)dzd\omega < \infty$, where $p(y|X,\omega,z)$, the conditional pdf of y, is given in (2.3).

A7) X is a random matrix such that the joint density $p(X,z,\omega)$ factorizes as $p(X)p(z,\omega)$.

It is important to emphasize the general nature of assumptions A1-A7. For example, we can obtain the model considered in Jammalamadaka, Tiwari and Chib (1987), which we will refer to as the *linear spherical model*, simply by letting

A2') $h(X,\beta) = X\beta$, and

A3') $V(X,\eta) = I_n$, (the identity matrix of order n), so that $\theta = \beta$.

On the other hand, assumption A4 is satisfied for a variety of distributions including those with point masses. Of course, the implied sampling distribution of y is a *scale mixture of Normals* with pdf derived as

$$p(y|X,\omega) = \int p(y|X,\omega,z)dG(z|\omega), \tag{2.2}$$

where

$$p(y|X,\omega,z) = f_N(y|h(X,\beta),\tau^{-2}\psi(z,\nu)^2V(X,\eta)). \tag{2.3}$$

From (2.2) and (2.3) we can secure the multivariate-t (MVt) distribution for y if we specialize assumptions A4 and A5 to

A4') $z|\omega \sim \chi^2_\nu$, the chi-squared distribution with ν degrees of freedom, and $\nu \in R_+$ (ie., $\ell = 1$ and $N = R_+$).
A5') $\psi(z,\nu) = (z/\nu)^{-1/2}$, $\nu > 0$.

It can be deduced that under A4' and A5' the sampling distribution is given by

$$p(y|X,\omega) = f_T(y|\nu,h(X,\beta),\tau^{-2}V(X,\eta)), \tag{2.4}$$

which is a non-linear, elliptical version of the model considered in Zellner (1976). Next, the differentiability requirement in A5 is imposed to ensure that the reparameterization[1]

$$(\tau^2,z,\nu) \rightarrow (\phi^2,z,\nu), \quad \phi^2 \equiv \tau^2\psi(z,\nu)^{-2}, \tag{2.5}$$

has a well defined Jacobian given by $\psi(z,\nu)^2$. It follows that if the joint density of (τ^2,z,ν) given θ is $p(\tau^2,z,\nu|\theta)$, then the joint density of (ϕ^2,z,ν) given θ is,

$$p(\phi^2,z,\nu|\theta) = \psi(z,\nu)^2\, p(\tau^2,z,\nu|\theta), \tag{2.6}$$

where the density at the r.h.s. is evaluated at $\tau^2 = \psi(z,\nu)^2\phi^2$. The role of A6 is to ensure that the posterior of ω is proper. Finally, Assumption A7, which allows for random regressors, serves to make posterior inference on ω independent of the distribution of X.

---

[1] For ν known, the transformation $(\tau^2,z) \rightarrow (\phi^2,z)$ appeared in Jammalamadaka et al. (1987), and it proved very clarifying for both posterior and predictive analyses in Osiewalski (1990).

2.2 PRIOR-POSTERIOR ANALYSIS: PROPER PRIORS

The general problem is concerned with the prior-posterior analysis for $\nu \epsilon N$ in model (2.1) when assumptions A1-A7 hold. By Bayes theorem, the posterior of ν has density given by

$$p(\nu|y,X) = \int p(\omega|y,X)d\theta d\tau^2 ,$$

where

$$p(\omega|y,X) = \int p(y,z,\omega|X)dz/\int p(y,z,\omega|X)d\omega dz ,$$

is the posterior of $\omega \equiv (\theta,\tau^2,\nu)$, and

$$p(y,z,\omega|X) = p(y|z,\omega,X)\ p(z|\omega)\ p(\omega) , \tag{2.7}$$

is the joint density of y, z, ω, given X. We will refer to (2.7) as the *complete Bayesian model*. In this section, our plan is to provide conditions under which the posterior of ν, $p(\nu|y,X)$, is the same as the prior, $p(\nu)$. We restrict attention to priors of ω that are *proper*.

We begin by noting that (2.3) and (2.5) allow us to conclude that y and (z,ν) are conditionally independent given X, θ and ϕ^2:

$$y \perp\!\!\!\perp (z,\nu)|X,\theta,\phi^2 . \tag{2.8}$$

It also ensues from A7 and (2.5) that

$$X \perp\!\!\!\perp (\theta,\phi^2,z,\nu). \tag{2.9}$$

The fundamental properties of conditional independence [see for e.g. Dawid (1979), and Mouchart and Rolin (1984)] enable us to infer that the pair of independence conditions (2.8) and (2.9) imply that $y \perp\!\!\!\perp \nu|X,\theta,\phi^2$ and $X \perp\!\!\!\perp \nu|\theta,\phi^2$. This in turn implies that ν is independent of the data given (θ,ϕ^2), i.e.,

$$\nu \perp\!\!\!\perp (y, X) \mid \theta, \phi^2 \ . \tag{2.10}$$

From (2.10) we can observe the fundamental point that if ν and (θ, ϕ^2) are independent, then (X,y) and ν are unconditionally independent, and the marginal posterior of ν is identical to its marginal prior. Thus, in this case the data cannot modify our prior opinions about ν.[2] We summarize this fact in the following result which provides the principal sufficient condition for the impossibility of updating the prior of ν.

THEOREM 1. Consider (2.1) under Assumptions A1-A7. Suppose that

$$\nu \perp\!\!\!\perp (\theta, \phi^2) \ , \tag{2.11}$$

where ϕ^2 is defined in (2.3); then,

$$\nu \perp\!\!\!\perp (y, X) \ . \tag{2.12}$$

PROOF: As we have argued, (2.10) follows from the hypothesis. However, (2.10) and (2.11) are (jointly) equivalent to $\nu \perp\!\!\!\perp (y, X, \theta, \phi^2)$. This leads to the conclusion stated in (2.12). The proof is completed. Δ

From Theorem 1 we can obtain a simple condition that is stronger than (2.11) but perhaps easier to check.

COROLLARY 1. If $(z, \nu) \perp\!\!\!\perp (\theta, \phi^2)$, (2.13)

then $\nu \perp\!\!\!\perp (y, X)$.

It is important to remark that Theorem 1 holds under very weak restrictions

---

[2] A general analysis, though in terms of densities, can be found in Dreze and Richard (1983).

on the distribution of z (cf., assumption A4). In the rest of the paper, however, we confine our attention to pdfs of z, $p(z|\theta,\tau^2,\nu)$, that are defined wrt the Lebesgue measure. This is because our leading interest is in multivariate t distributions, and these are continuous scale mixtures of Normals.

In order to illustrate situations where Theorem 1 holds, i.e., where the prior of ν is not updated (under proper pdf's) due to independence between ν and (θ,ϕ^2), we consider two regression models with MVt errors. The first example, which is presented in Corollary 2, is based on a generalization of the model and prior used in Jammalamadaka et al (1987); see also Osiewalski (1990) and Osiewalski and Steel (1990, Section 3). The reader can check (by applying (2.2)), that under the hypotheses of Corollary 2, the distribution of y is MVt although the precision matrix of y depends on τ^2 and θ in a very complicated way, and the degrees-of-freedom parameter $(m+\nu)$ is functionally related to the hyperparameter m of the conditional prior $p(\tau^2|\theta,\nu)$. In the second example, which forms Corollary 3, we appraise the more familiar MVt sampling model given in (2.4).

COROLLARY 2. In (2.1), suppose A1-A3, A5′, A6 and A7 hold. If

(i) $p(z|\theta,\tau^2,\nu) = f_G(z|(m+\nu)/2,\{\nu+\tau^2 d(\theta)/2\nu\})$,

(ii) $p(\tau^2|\theta,\nu) = f_{IB}(\tau^2|\ \frac{\nu}{2},\frac{m}{2},\frac{\nu}{d(\theta)})$;

(iii) $p(\theta,\nu) = p(\theta)\ p(\nu)$;

where m is a positive constant, $d(\theta)$ is some known positive function of θ, and f_G and $f_{IB}(.|.,.,.)$ denote the gamma pdf, and the three-parameter inverted beta pdf, repectively.[3] Then, $\nu \perp\!\!\!\perp (y,X)$.

PROOF: With the change of variable $\phi^2 = \tau^2\{\psi(z,\nu)\}^{-2} = \tau^2(\frac{z}{\nu})$, standard calculations show that $p(z|\theta,\phi^2,\nu) = p(z|\nu) = f_G(z|\frac{\nu}{2}, \frac{1}{2})$, and $p(\phi^2|\theta,\nu) = p(\phi^2|\theta) = f_G(\phi^2|\frac{m}{2},\frac{d(\theta)}{2})$. Thus, $p(\theta,\phi^2,\nu) = p(\theta)\ f_G(\phi^2|\frac{m}{2},\frac{d(\theta)}{2})\ p(\nu)$, and the condition (2.11) of Theorem 1 is fulfilled. (Even the stronger condition (2.13) is met here.)

Δ

COROLLARY 3. In (2.1), under A1-A3, A4', A5', A6 and A7 (these lead to (2.4)), suppose that

(i) $\quad p(\tau^2|\theta,\nu) = f_B(\tau^2|\frac{m}{2},\frac{\nu-m}{2},\frac{\nu}{d(\theta)})$,

(ii) $\quad p(\theta,\nu) = p(\theta)\ p(\nu)$,

(iii) $\quad p(\nu) = 0$ if $\nu \leq m$,

where m is a positive constant, $d(\theta)$ is some known positive function of θ, and $f_B(\tau^2|\frac{m}{2},\frac{\nu-m}{2},\frac{\nu}{d(\theta)})$ denotes the beta pdf with the parameters $(\frac{m}{2},\frac{\nu-m}{2})$ and nonzero over the interval $(0,\frac{\nu}{d(\theta)})$.[4] Then $\nu \perp\!\!\!\perp (y,X)$.

PROOF: From the parametrization, $\phi^2 = \tau^2\psi(z,\nu)^{-2} = \tau^2\frac{z}{\nu}$, it follows that $p(\phi^2|\theta,\nu) = p(\phi^2|\theta) = f_G(\phi^2|\frac{m}{2},\frac{d(\theta)}{2})$, and $p(z|\theta,\phi^2,\nu) = f_G(z-\phi^2 d(\theta)|\frac{\nu-m}{2},\frac{1}{2})$.

---

[3] The three-parameter inverted beta (or beta prime) density on w>0 with a,b,c>0 [see Zellner (1971, p. 376)] is

$$f_{IB}(w|a,b,c) = c^{-1}B(a,b)\ (w/c)^{b-1}\ (1+\frac{w}{c})^{-(a+b)} ,$$

where $B(a,b) = \Gamma(a+b)/(\Gamma(a)\Gamma(b))$ is the Beta function. For $a=\nu_2/2$, $b=\nu_1/2$, $c=\nu_2/\nu_1$, we obtain the F density with ν_1,ν_2 degrees of freedom.

[4] A beta density on $w\epsilon(0,c)$ with a,b>0 is given by

$$f_B(w|a,b,c) = c^{-1}B(a,b)\ (\frac{w}{c})^{a-1}(1-\frac{w}{c})^{b-1} .$$

The last density is only nonzero when $z > \phi^2 d(\theta)$, i.e., in the support of τ^2, and it also implicitly imposes the other prior constraint that $\nu > m$. Hence, $p(\theta,\phi^2,\nu) = p(\theta)\; f_G(\phi^2|\frac{m}{2},\frac{d(\theta)}{2})\; p(\nu)$, and condition (2.11) of Theorem 1 is satisfied (although the stronger condition (2.13) of Corollary 1 is not). Δ

In the case of a linear model, ie., when Assumption 2′ holds and r=k, the prior structures appearing in Corollaries 2 and 3 are closely related to the semi-conjugate priors that are introduced by Osiewalski and Steel (1990). Semi-conjugate priors of β and τ^2 (given η and ν) are defined as those priors which correspond to Normal-gamma distributions of β and ϕ^2 (given η and ν).[5] Let $m = k+\mu$ ($\mu>0$) and $d(\theta) = f_\eta + (\beta-\bar{\beta}_\eta)'A_\eta(\beta-\bar{\beta}_\eta)$, where f_η, $\bar{\beta}_\eta$ and A_η are, respectively, a positive scalar, a k-dimensional vector, and a kxk PDS matrix (all possibly depending on η, but not on ν). Assume that

$$p(\beta|\eta,\nu) = p(\beta|\eta) = f_T(\beta|\mu,\bar{\beta}_\eta,(\tfrac{\mu}{f_\eta}A_\eta)^{-1}),$$

then in the cases considered in Corollary 2 and 3 we have respectively

$$\begin{aligned} p(\beta,\tau^2|\eta,\nu) &= f_T(\beta|\mu,\bar{\beta}_\eta,(\tfrac{\mu}{f_\eta}A_\eta)^{-1})\; f_{IB}(\tau^2|\tfrac{\nu}{2},\tfrac{k+\mu}{2},\tfrac{\nu}{d(\theta)}), \\ p(\beta,\tau^2|\eta,\nu) &= f_T(\beta|\mu,\bar{\beta}_\eta,(\tfrac{\mu}{f_\eta}A_\eta)^{-1})\; f_B(\tau^2|\tfrac{k+\mu}{2},\tfrac{\nu-k-\mu}{2},\tfrac{\nu}{d(\theta)}). \end{aligned} \tag{2.14}$$

In both situations, the priors in (2.14) lead to a Normal-gamma distribution of β and ϕ^2 (given η), independent of ν, given by

$$p(\beta,\phi^2|\eta,\nu) = p(\beta,\phi^2|\eta) = f_N(\beta|\bar{\beta}_\eta,\phi^{-2}A_\eta^{-1})\; f_G(\phi^2|\tfrac{\mu}{2},\; f_\eta/2)\ .$$

---

[5] Osiewalski and Steel (1990) assume that ν and V are fully known. In our framework, where ν is unknown and $V = V(X,\eta)$, their considerations can be interpreted as conditional on ν,η and X.

Due to the fact that μ, f_η, $\bar{\beta}_\eta$ and A_η are not functionally related to ν, the semi-conjugate priors above preclude learning from the data about ν, provided that η and ν are a priori independent.

2.3 PRIOR-POSTERIOR ANALYSIS: IMPROPER PRIORS

We now turn to the specially interesting case of the usual improper prior of τ^2, which is not covered by Theorem 1, since the arguments made there are not necessarily valid for distributions that are not proper.

THEOREM 2: Consider (2.1) under A1-A5, and A7. Suppose

$$p(\omega,z) = p(\theta)p(\tau^2,z,\nu|\theta)$$

and

$$p(\tau^2,z,\nu|\theta) = c\tau^{-2}p(z|\theta,\nu)p(\nu), \tag{2.15}$$

where $c > 0$ is any constant, the pdf of $z|\theta,\nu$ is proper and functionally independent of τ^2, and that of ν is proper and functionally independent of (τ^2,θ). If

$$\int_{BxH} p(\theta)|V(X,\eta)|^{-\frac{1}{2}}[(y - h(X,\beta))'V(X,\eta)^{-1}(y - h(X,\beta))]^{-\frac{n}{2}}d\theta < \infty, \tag{2.16}$$

then $p(\nu|y,X) = p(\nu)$.

PROOF: From (2.15), and the ϕ^2 parametrization, $p(\phi^2,z,\nu|\theta) = c\phi^{-2}p(z|\theta,\nu)p(\nu)$. Applying (2.7), the marginal posterior of ν is proportional to

$$p(\nu|y,X) \propto p(\nu)\int_{BxH}\int_0^\infty\int_0^\infty f_N(y|h(X,\beta),\phi^{-2}V(X,\eta))\phi^{-2}p(z|\theta,\nu)p(\theta)dzd\phi^2d\theta$$

$$\propto p(\nu)\int p(\theta)|V(X,\eta)|^{-\frac{1}{2}}[(y - h(X,\beta))'V(X,\eta)^{-1}(y - h(X,\beta))]^{-\frac{n}{2}}d\theta,$$

and the result is immediate. Δ

We now point out an implication of Theorem 2 for the case when the observation vector y is distributed as MVt (given X, θ,τ^2,ν).

COROLLARY 4. In (2.1), suppose A1-A3, A4′, A5′ and A7 hold. Also suppose that in (2.15) we have

$$p(\tau^2,z,\nu|\theta) = c\tau^{-2}p(\nu)\ f_G(z|\tfrac{\nu}{2},\tfrac{1}{2}),$$

where $p(\nu)$ is proper and assume that the integral in (2.16) converges. Then, $p(\nu|y,X) = p(\nu)$.

Of course, Theorem 2 can be applied to situations other than that described in Corollary 4 as long as the distribution of $z|\omega$ does not functionally depend on τ^2. If the latter condition holds, then the improper prior of the precision parameter τ^2 together with the prior independence between θ and ν are sufficient to prevent an updating of the prior of ν.

3. USEFUL PRIOR FAMILIES

We now examine classes of prior distributions that may allow us to update the prior of ν. Essentially, the idea is to propose families of priors that do not satisfy the sufficient conditions of Theorems 1 and 2 and Corollaries 1-4. We define three such families. We limit our attention to the model with the observation vector distributed as n-variate Student-t with ν degrees of freedom, the location vector $h(X,\beta)$ and the precision matrix $\tau^2(V(X,\eta))^{-1}$(cf. (2.4)). Therefore, for the remainder of the discussion, $y|X,\theta,\tau^2,z,\nu$ and $z|\theta,\tau^2,\nu$ have distributions with pdfs given by

$$p(y|X,\theta,\tau^2,z,\nu) = f_N(y|h(X;\beta),\nu\tau^{-2}z^{-1}V(X,\eta)) ,$$

$$p(z|\theta,\tau^2,\nu) = p(z|\nu) = f_G(z|\frac{\nu}{2},\frac{1}{2}). \qquad (3.1)$$

3.1 PROPER PRIOR FAMILIES

A simple and useful proper prior family that does not satisfy the condition of Corollary 1 can be based on an informative prior on τ^2 as follows. Consider the following general prior structure

$$\text{Prior 1:} \qquad p(\theta,\tau^2,\nu) = p(\theta)\ f_G(\tau^2|\frac{m}{2},\frac{d(\theta)}{2})\ p(\nu) \qquad (3.2)$$

where $p(\tau^2|\theta)$ is the gamma pdf, m is a positive constant, and $d(\theta)$ is a known positive function of θ, which may be the constant function. In (3.2), (θ,τ^2) and ν are independent. Simple calculations show that, for $\phi^2 = \tau^2 z/\nu$,

$$p(\phi^2|\theta,z,\nu) = f_G(\phi^2|\frac{m}{2},\frac{\nu}{2z}\ d(\theta)) .$$

Thus ϕ^2 and (z,ν) are dependent, given θ, and condition (2.13) of Corollary 1 is not met, in spite of the independence between θ and (z,ν) from (3.1) and (3.2). The next example describes some implications of adopting Prior 1.

Example 1. Consider (2.4) and suppose A2' and A3' hold. Assume the following Normal-gamma prior structure for (θ,τ^2) (independent of ν):

$$p(\theta,\tau^2) = f_N(\beta|b,\tau^{-2}C)\ f_G(\tau^2|\frac{a}{2},\frac{f}{2}).$$

Since $p(\theta,\tau^2)$ can be also expressed as

$$p(\theta,\tau^2) = f_T(\beta|a,b,\frac{f}{a}C)f_G(\tau^2|\frac{a+k}{2},\frac{f+(\beta-b)'C^{-1}(\beta-b)}{2}),$$

we have a special case of Prior 1, with $\theta = \beta$, $m = a+k$, $d(\theta) = f+(\beta-b)'C^{-1}(\beta-b)$, and $p(\beta) = f_T(\beta|a,b,(f/a)C)$. Under these prior assumptions we obtain the following marginalized likelihood:

$$p(y|X,\nu) = \int_0^\infty f_T(y|a,Xb,\frac{f}{a}(\frac{\nu}{z}I_n+XCX'))f_G(z|\frac{\nu}{2},\frac{1}{2})\ dz,$$

which is, generally, not constant in ν. However, the calculation of posterior moments of ν will require bivariate numerical integrations (w.r.t. z and ν).

Δ

Another prior family that allows for the updating of $p(\nu)$ can be proposed. Consider a three-parameter inverted beta prior and assume that

Prior 2: $$p(\theta,\nu) = p(\theta)\ p(\nu)$$

$$p(\tau^2|\theta,\nu) = f_{IB}(\tau^2|\frac{\nu}{2},\frac{m}{2},\frac{\nu}{d(\theta)}). \tag{3.3}$$

In this case, after some calculations we find that the pdf of $\phi^2 = \tau^2 z/\nu$ is given by

$$p(\phi^2|\theta,z,\nu) = f_{IB}(\phi^2|\frac{\nu}{2},\frac{m}{2},\frac{z}{d(\theta)}),$$

which shows that ϕ^2 and (z,ν) are dependent, given θ, and that θ and (z,ν) are independent (which can be verified using (3.1) and (3.3)). However, since the condition (2.13) of Corollary 1 is not met, updating ν's prior may be possible. Prior 2 generalizes the joint informative prior specified by Zellner (1976) for the linear model with diagonal error covariance matrix. This fact is brought out in the following example.

Example 2. Consider the linear spherical model used in Example 1. Assume the following conditional prior of (β,τ^2) given ν:

$$p(\theta,\tau^2|\nu) = p(\beta|\tau^2,\nu)\ p(\tau^2|\nu) \tag{3.4}$$
$$= f_T(\beta|a+\nu,b,\frac{\nu+f\tau^2}{a+\nu}\tau^{-2}C)f_{IB}(\tau^2|\frac{\nu}{2},\frac{a}{2},\frac{\nu}{f}),$$

where the conditional prior of $\theta = \beta$ (given τ^2 and ν) is k-variate MVt with $a+\nu$ degrees of freedom, mean vector b and precision matrix $\{(a+\nu)/(\nu+f\tau^2)\}\tau^2C^{-1}$. The conditional prior of τ^2 given ν is a three-parameter inverted beta distribution such that $f\tau^2/a$ has an F distribution with (a,ν) degrees of freedom. This is exactly the informative prior proposed by Zellner (1976). Note that (3.4) can be equivalently written as

$$p(\theta,\tau^2|\nu) = p(\beta|\nu)\ p(\tau^2|\beta,\nu) = p(\beta)\ p(\tau^2|\beta,\nu)$$
$$= f_T(\beta|a,b,\frac{f}{a}C)\ f_{IB}(\tau^2|\frac{\nu}{2},\frac{a+k}{2},\frac{\nu}{g(\beta)}), \tag{3.5}$$

where $g(\beta) = f+(\beta-b)'C^{-1}(\beta-b)$. The form of $p(\tau^2|\beta,\nu)$ appearing in (3.5) is a special case of Prior 2 with $m=a+k > k$ and $d(\theta) = g(\beta)$. If the prior of (3.5) is used, the calculations are slightly heavier than in the Normal-gamma case in Example 1. The marginalized likelihood is given by

$$p(y|X,\nu) = \int_0^\infty f_T(y|a,Xb,\frac{f}{a}(sI_n+XCX'))f_F(s|\nu,\nu)\ ds,$$

where $f_F(s|\nu,\nu) = f_{IB}(s|\nu/2,\nu/2,1)$ is an F density with (ν,ν) degrees of freedom. Again, $p(y|X,\nu)$ is (generally) not constant in ν, so $p(\nu|y,X) \neq p(\nu)$. Deriving the posterior moments of ν will require bivariate numerical integration (w.r.t. s and ν).

Δ

3.2 IMPROPER PRIOR FAMILIES

In this section we focus on prior families that are improper and which permit an update of prior beliefs. Note that the main condition of Theorem 2 (i.e. (2.15)) could be satisfied if prior beliefs about the precision parameter τ^2 are vague and the joint prior of ω is

$$\text{Prior 3.} \qquad p(\theta,\tau^2,\nu) \propto p(\theta,\nu)\ \tau^{-2}. \tag{3.6}$$

However, (3.6) will only imply (2.15) if we assume that $p(\theta,\nu) = p(\theta)\ p(\nu)$, in addition to (3.1). Therefore, in order to update our prior beliefs about ν, we cannot allow such a factorization. Thus, only those priors of (θ,ν) which make θ dependent on ν are worth considering. The prior-posterior analysis for ν with Prior 3 is examined in the following example.

Example 3. Consider the linear spherical model from Example 1. Assume the following prior structure on (β,ν):

$$p(\beta,\nu) = f_N(\beta|b_\nu,C_\nu)p(\nu). \tag{3.7}$$

Combining (3.1) with (3.6) and (3.7), the complete Bayesian model (cf. (2.7)), in the ϕ^2 parametrization is given by

$$p(y,\beta,\phi^2,z,\nu|X) \propto f_N(y|X\beta,\ \phi^{-2}\ I_n)\phi^{-2}\ f_G(z|\tfrac{\nu}{2},\tfrac{1}{2})\ p(\beta,\nu). \tag{3.8}$$

Integrating out z and β from (3.8) yields the following posterior of ν:

$$p(\nu|y,X) \propto p(\nu)\cdot \int_0^\infty \phi^{-2}\ f_N(y|Xb_\nu,(\phi^{-2}I_n + XC_\nu X'))\ d\phi^2.$$

This posterior is not equal to the prior, $p(\nu)$, as long as b_ν or C_ν are not constant in ν. Δ

Some remarks about these examples are important. Note that in Examples 1 and 2, where the priors on τ^2 are informative, we have also assumed an informative prior on β, and prior dependence between β and τ^2. First of all, this prior dependence is not crucial. We could assume in (3.2) or (3.3) that $d(\theta) = h$, where h is a positive constant, and in this way make τ^2 independent of θ given ν. On the other hand, prior dependence makes calculations easier. Without this we could be faced with the necessity of more than bivariate numerical integration in the informative case. The second point worth mentioning is that, when the prior of τ^2 is informative, the prior of β can be diffuse and still the prior of ν can be updated, as illustrated in the following example.

Example 4. Consider the linear spherical model of Example 1. Assume the following improper prior (which, loosely speaking, can be treated as a special case of Prior 1):

$$p(\theta,\tau^2,\nu) = p(\beta)\ p(\tau^2)\ p(\nu) \ ,$$

$$p(\beta) = c \ , \ \beta\varepsilon R^k \ ,$$

$$p(\tau^2) = f_G(\tau^2|\tfrac{m}{2},\tfrac{h}{2}) \ ,$$

where c,m,h are positive constants. In this example, X is of full column rank k, and the complete Bayesian model (2.7) is given by

$$p(y,\beta,\tau^2,z,\nu|X) = c\ f_N(y|X\beta,\tfrac{\nu}{z}\tau^{-2}I_n)\ f_G(\tau^2|\tfrac{m}{2},\tfrac{h}{2})\ f_G(z|\tfrac{\nu}{2},\tfrac{1}{2})\ p(\nu)$$

or, in terms of $\phi^2 = (z\tau^2)/\nu$,

$$p(y,\beta,\phi^2,z,\nu|X) = c\ f_N(y|X\beta,\phi^{-2}\ I_n)\ f_G(\phi^2|\tfrac{m}{2},\tfrac{h\nu}{2z})\ f_G(z|\tfrac{\nu}{2},\tfrac{1}{2})\ p(\nu) \ .$$

Letting $\hat{\beta} = (X'X)^{-1}X'y$, $s^2 = (y-X\hat{\beta})'(y-X\hat{\beta})$, and performing analytical integration w.r.t. β and ϕ^2 we obtain the following formula for the posterior pdf of ν:

$$p(\nu|y,X) \propto p(\nu) \int_0^\infty \{1 + zs^2/(\nu h)\}^{-(m+n-k)/2} \{z/\nu\}^{(n-k)/2} f_G(z|\tfrac{\nu}{2},\tfrac{1}{2})\, dz. \quad \Delta$$

It is important to note that in all four examples numerical integration w.r.t. some auxiliary variables is necessary in order to obtain the marginal posterior density of the degrees-of-freedom parameter. Furthermore, we have explicitly treated only the simplest case: the linear model with scalar precision matrix. Our results, however, can be applied to more complicated situations, such as nonlinear models and non-constant $V(X,\eta)$, from the general theoretical basis for inference on ν developed in Sections 2 and 3.

4. CONCLUSION

This paper has focused on a general class of nonlinear, elliptical error regression models and discussed Bayesian inference on the degrees of freedom parameter of the error distribution. We have provided conditions under which the prior of ν is not updated by the sample data. Three classes of priors that can be updated are specified, and the prior-posterior analysis with these priors is illustrated within the context of linear models. We feel that the results obtained here would be quite useful in applications involving heavy-tailed error distributions such as the MVt.

References

Chib, S., R.C. Tiwari, and S.R. Jammalamadaka, 1988, Bayes Prediction in Regressions with Elliptical Errors, Journal of Econometrics, 38, 349-360.

David, A.P., 1979, Conditional Independence in Statistical Theory, Journal of the Royal Statistical Society, B, 41, 1-31.

Dreze, J. and J.F. Richard, 1983, Bayesian Analysis of Simultaneous Equation Systems, in Z. Griliches and M.D. Intriligator, eds, Handbook of Econometrics, North Holland, Amsterdam.

Jammalamadaka S.R., R.C. Tiwari, S. Chib, 1987, Bayes Prediction in the Linear Model with Spherically Symmetric Errors, Economics Letters, 24, 39-44.

Mouchart, M. and J.M. Rolin, 1984, A Note on Conditional Independence with Statistical Applications, Statistica, 44, 557-584.

Osiewalski, J., 1990, A Note on Bayesian Inference in a Regression Model with Elliptical Errors, Journal of Econometrics, forthcoming.

Osiewalski, J. and M.F.J. Steel, 1990, Semi-Conjugate Prior Densities in Multivariate-t Regression Models, CentER Discussion Paper 9007, Tilburg University, Tilburg.

Phillips, P.C.B., 1988, Conditional and Unconditional Statistical Independence, Journal of Econometrics, 38, 341-348.

Singh, R.S., 1988, Estimation of Error Variance in Linear Regression Models with Errors Having Multivariate Student-t Distributions with Unknown Degrees of Freedom, Economics Letters, 27, 47-53.

Zellner, A., 1971, An Introduction to Bayesian Inference in Econometrics, Wiley, New York.

Zellner, A., 1976, Bayesian and non-Bayesian Analysis of the Regression Model with Multivariate Student-t Error Terms, Journal of the American Statistical Association, 71, 400-405.

Discussion Paper Series, CentER, Tilburg University, The Netherlands:

(For previous papers please consult previous discussion papers.)

No.	Author(s)	Title
8932	E. van Damme, R. Selten and E. Winter	Alternating Bid Bargaining with a Smallest Money Unit
8933	H. Carlsson and E. van Damme	Global Payoff Uncertainty and Risk Dominance
8934	H. Huizinga	National Tax Policies towards Product-Innovating Multinational Enterprises
8935	C. Dang and D. Talman	A New Triangulation of the Unit Simplex for Computing Economic Equilibria
8936	Th. Nijman and M. Verbeek	The Nonresponse Bias in the Analysis of the Determinants of Total Annual Expenditures of Households Based on Panel Data
8937	A.P. Barten	The Estimation of Mixed Demand Systems
8938	G. Marini	Monetary Shocks and the Nominal Interest Rate
8939	W. Güth and E. van Damme	Equilibrium Selection in the Spence Signaling Game
8940	G. Marini and P. Scaramozzino	Monopolistic Competition, Expected Inflation and Contract Length
8941	J.K. Dagsvik	The Generalized Extreme Value Random Utility Model for Continuous Choice
8942	M.F.J. Steel	Weak Exogenity in Misspecified Sequential Models
8943	A. Roell	Dual Capacity Trading and the Quality of the Market
8944	C. Hsiao	Identification and Estimation of Dichotomous Latent Variables Models Using Panel Data
8945	R.P. Gilles	Equilibrium in a Pure Exchange Economy with an Arbitrary Communication Structure
8946	W.B. MacLeod and J.M. Malcomson	Efficient Specific Investments, Incomplete Contracts, and the Role of Market Alternatives
8947	A. van Soest and A. Kapteyn	The Impact of Minimum Wage Regulations on Employment and the Wage Rate Distribution
8948	P. Kooreman and B. Melenberg	Maximum Score Estimation in the Ordered Response Model

No.	Author(s)	Title
8949	C. Dang	The D_3-Triangulation for Simplicial Deformation Algorithms for Computing Solutions of Nonlinear Equations
8950	M. Cripps	Dealer Behaviour and Price Volatility in Asset Markets
8951	T. Wansbeek and A. Kapteyn	Simple Estimators for Dynamic Panel Data Models with Errors in Variables
8952	Y. Dai, G. van der Laan, D. Talman and Y. Yamamoto	A Simplicial Algorithm for the Nonlinear Stationary Point Problem on an Unbounded Polyhedron
8953	F. van der Ploeg	Risk Aversion, Intertemporal Substitution and Consumption: The CARA-LQ Problem
8954	A. Kapteyn, S. van de Geer, H. van de Stadt and T. Wansbeek	Interdependent Preferences: An Econometric Analysis
8955	L. Zou	Ownership Structure and Efficiency: An Incentive Mechanism Approach
8956	P.Kooreman and A. Kapteyn	On the Empirical Implementation of Some Game Theoretic Models of Household Labor Supply
8957	E. van Damme	Signaling and Forward Induction in a Market Entry Context
9001	A. van Soest, P. Kooreman and A. Kapteyn	Coherency and Regularity of Demand Systems with Equality and Inequality Constraints
9002	J.R. Magnus and B. Pesaran	Forecasting, Misspecification and Unit Roots: The Case of AR(1) Versus ARMA(1,1)
9003	J. Driffill and C. Schultz	Wage Setting and Stabilization Policy in a Game with Renegotiation
9004	M. McAleer, M.H. Pesaran and A. Bera	Alternative Approaches to Testing Non-Nested Models with Autocorrelated Disturbances: An Application to Models of U.S. Unemployment
9005	Th. ten Raa and M.F.J. Steel	A Stochastic Analysis of an Input-Output Model: Comment
9006	M. McAleer and C.R. McKenzie	Keynesian and New Classical Models of Unemployment Revisited
9007	J. Osiewalski and M.F.J. Steel	Semi-Conjugate Prior Densities in Multivariate t Regression Models

No.	Author(s)	Title
9007	J. Osiewalski and M.F.J. Steel	Semi-Conjugate Prior Densities in Multivariate t Regression Models
9008	G.W. Imbens	Duration Models with Time-Varying Coefficients
9009	G.W. Imbens	An Efficient Method of Moments Estimator for Discrete Choice Models with Choice-Based Sampling
9010	P. Deschamps	Expectations and Intertemporal Separability in an Empirical Model of Consumption and Investment under Uncertainty
9011	W. Güth and E. van Damme	Gorby Games - A Game Theoretic Analysis of Disarmament Campaigns and the Defense Efficiency-Hypothesis
9012	A. Horsley and A. Wrobel	The Existence of an Equilibrium Density for Marginal Cost Prices, and the Solution to the Shifting-Peak Problem
9013	A. Horsley and A. Wrobel	The Closedness of the Free-Disposal Hull of a Production Set
9014	A. Horsley and A. Wrobel	The Continuity of the Equilibrium Price Density: The Case of Symmetric Joint Costs, and a Solution to the Shifting-Pattern Problem
9015	A. van den Elzen, G. van der Laan and D. Talman	An Adjustment Process for an Exchange Economy with Linear Production Technologies
9016	P. Deschamps	On Fractional Demand Systems and Budget Share Positivity
9017	B.J. Christensen and N.M. Kiefer	The Exact Likelihood Function for an Empirical Job Search Model
9018	M. Verbeek and Th. Nijman	Testing for Selectivity Bias in Panel Data Models
9019	J.R. Magnus and B. Pesaran	Evaluation of Moments of Ratios of Quadratic Forms in Normal Variables and Related Statistics
9020	A. Robson	Status, the Distribution of Wealth, Social and Private Attitudes to Risk
9021	J.R. Magnus and B. Pesaran	Evaluation of Moments of Quadratic Forms in Normal Variables

No.	Author(s)	Title
9022	K. Kamiya and D. Talman	Linear Stationary Point Problems
9023	W. Emons	Good Times, Bad Times, and Vertical Upstream Integration
9024	C. Dang	The D_2-Triangulation for Simplicial Homotopy Algorithms for Computing Solutions of Nonlinear Equations
9025	K. Kamiya and D. Talman	Variable Dimension Simplicial Algorithm for Balanced Games
9026	P. Skott	Efficiency Wages, Mark-Up Pricing and Effective Demand
9027	C. Dang and D. Talman	The D_1-Triangulation in Simplicial Variable Dimension Algorithms for Computing Solutions of Nonlinear Equations
9028	J. Bai, A.J. Jakeman and M. McAleer	Discrimination Between Nested Two- and Three-Parameter Distributions: An Application to Models of Air Pollution
9029	Th. van de Klundert	Crowding out and the Wealth of Nations
9030	Th. van de Klundert and R. Gradus	Optimal Government Debt under Distortionary Taxation
9031	A. Weber	The Credibility of Monetary Target Announcements: An Empirical Evaluation
9032	J. Osiewalski and M. Steel	Robust Bayesian Inference in Elliptical Regression Models
9033	C. R. Wichers	The Linear-Algebraic Structure of Least Squares
9034	C. de Vries	On the Relation between GARCH and Stable Processes
9035	M.R. Baye, D.W. Jansen and Q. Li	Aggregation and the "Random Objective" Justification for Disturbances in Complete Demand Systems
9036	J. Driffill	The Term Structure of Interest Rates: Structural Stability and Macroeconomic Policy Changes in the UK
9037	F. van der Ploeg	Budgetary Aspects of Economic and Monetary Integration in Europe

No.	Author(s)	Title
9038	A. Robson	Existence of Nash Equilibrium in Mixed Strategies for Games where Payoffs Need not Be Continuous in Pure Strategies
9039	A. Robson	An "Informationally Robust Equilibrium" for Two-Person Nonzero-Sum Games
9040	M.R. Baye, G. Tian and J. Zhou	The Existence of Pure-Strategy Nash Equilibrium in Games with Payoffs that are not Quasiconcave
9041	M. Burnovsky and I. Zang	"Costless" Indirect Regulation of Monopolies with Substantial Entry Cost
9042	P.J. Deschamps	Joint Tests for Regularity and Autocorrelation in Allocation Systems
9043	S. Chib, J. Osiewalski and M. Steel	Posterior Inference on the Degrees of Freedom Parameter in Multivariate-t Regression Models

PO. BOX 90153 5000 LE TILBURG, THE NETHERLANDS