CentER for Economic Research

Discussion paper

[redacted]

TILBURG UNIVERSITY

No. 9007

SEMI-CONJUGATE PRIOR DENSITIES IN MULTIVARIATE T REGRESSION MODELS

by Jacek Osiewalski
and Mark F.J. Steel

February 1990

ISSN 0924-7815

Semi-Conjugate Prior Densities
in Multivariate t Regression Models

Jacek Osiewalski
Academy of Economics
31-510 Kraków, Poland

Mark Steel
Tilburg University
5000 LE Tilburg, The Netherlands

Authors's footnote:

Jacek Osiewalski is Assistant Professor, Academy of Economics, 27 Rakowicka Str., 31-510 Kraków, Poland and Mark F.J. Steel is Research Fellow, Department of Econometrics, Tilburg University, P.O. Box 90153, 5000 LE Tilburg, The Netherlands. This paper was written while the first author was a visiting fellow at CORE, Louvain-la-Neuve, and the second author held a research fellowship of the Royal Netherlands Academy of Arts and Sciences (KNAW). The authors are indebted to Michel Mouchart and Hugo Keuzenkamp for helpful discussions, though the usual disclaimer applies.

Abstract

The meaning of conjugate prior densities for a linear regression model is examined when we venture outside the usual realm of exponential models. For a non-Normal elliptical family of data densities, we define a class of semi-conjugate prior densities, fully coherent with the uncontroversial conjugate prior in a Normal framework. We discuss results from the literature on the particular case of Student t errors, and derive a semi-conjugate prior for such models. Since the transformation used to obtain this prior does not affect the regression coefficient vector, any semi-conjugate prior leads to exactly the same marginal Student t prior and posterior densities for this vector as in the reference case of a Normal regression model with conjugate prior. It is shown that these semi-conjugate prior densities allow us to obtain most posterior results analytically under informative prior assumptions at the cost of putting a finite upper bound on the unknown error precision parameter, and thus restricting the original parameter space.

Keywords:

Bayesian econometrics, Student t errors, linear regression, conjugate prior densities, posterior inference.

1. Introduction

Part of the literature in Bayesian econometrics has been directed towards broadening the distributional assumptions on the error terms of the multiple regression model.
Zellner (1976) considers Student t errors and concludes that inference still remains relatively simple with diffuse priors. Jammalamadaka, Tiwari and Chib (1987), henceforth denoted by JTC, consider spherically distributed errors and state that under diffuse prior assumptions prediction is unaffected by such departures from Normality. Chib, Tiwari and Jammalamadaka (1988) extend these results to a case of elliptical errors and Osiewalski (1989) examines posterior and predictive inference in elliptical cases for possibly nonlinear models and general prior structures.

An intriguing aspect of such excursions outside the exponential family of distributions, is whether one can still find "conjugate" prior distributions, and, if so, in which sense of the word.
A quick perusal of the literature shows that the concept of conjugate prior distributions is not always defined in the same way.
Traditionally, a conjugate family is defined either by the property of proportionality to the likelihood (or, more generally, having the same functional form) or by being closed under sampling. Where Raiffa and Schlaiffer (1961) stress that the latter property is very desirable, they base their actual definition of conjugate priors on the former characteristic, a definition also adopted by e.g. Zellner (1971, p. 21) and Press (1972, p. 76).
Berger (1985, p. 130) uses the term "conjugate family" for a class of prior distributions that is closed under sampling, and reserves the term "natural conjugate" for priors with the same functional form as the likelihood. The latter property is stressed as being crucial for the construction of conjugate families in DeGroot (1970, p. 164), yet he defines these families as being closed under sampling.
Both concepts coincide for exponential (and uniform) families of data distributions, where the dimension of the sufficient statistics remains

fixed under independent sampling (see e.g. DeGroot 1970, pp. 159-164). Views only start to differ when we venture outside these standard cases.

Diaconis and Ylvisaker (1979, 1985) claim that these "traditional" concepts are essentially vacuous, and Hartigan (1983, p. 72) dismisses the concept of conjugate priors altogether, as a "chimera" (in his index, p. 142). It is shown by Diaconis and Ylvisaker (1979) that the "usual" conjugate families can be characterized unambiguously by a posterior mean that is linear in the observations. Goel and DeGroot (1980) investigate this linearity in a more general form for linear regression models, and suggest that "perhaps one should use linear posterior expectation as the defining property of a conjugate distribution" (p. 895). Outside the exponential family setting, this concept was elaborated for location parameter problems by Diaconis and Ylvisaker (1985).

In spite of the general theoretical debate, the restrictive Normal linear regression framework (with a covariance matrix known up to a scalar factor σ^2) allows a widely accepted form of conjugate prior, namely the Normal-inverted gamma distribution, or, equivalently, the Normal-gamma distribution if we parameterize in terms of $\tau^2 = \sigma^{-2}$. Such priors have the same functional form as the likelihood, are closed under sampling and give linear posterior means of regression coefficients.
For a regression model with multivariate t errors, however, the meaning of the term "conjugate prior" is not so straightforward. References to this question in the literature are fairly scarce and usually remain rather vague. Zellner (1976, p. 403) proposed a t-F prior structure as a natural conjugate distribution, since its functional form corresponds to the likelihood, but showed in an appendix that the posterior density will not retain the t-F structure, but will rather be a continuous mixture of t-F densities.
The commonly used conjugate prior for Normal linear regression, i.e. the Normal-gamma distribution, is the starting point for JTC (their Section 3) who suggest that the conjugate prior density for the spherical family is given by their formulae (6)-(8). They also state that for the specific multivariate t regression model considered by Zellner (1976), their approach gives his t-F prior as a conjugate one.

However, as we shall show in Sections 2 and 3, the correspondence between such "conjugate" priors in non-Normal elliptical models and Normal-gamma priors in Normal models is not as simple as it seemed in JTC, and, furthermore, this strategy cannot result in finding Zellner's (1976) t-F prior for his model. We borrow from JTC the basic idea of relating a "conjugate" prior for regression with spherical (or elliptical) errors to the Normal-gamma prior for regression with Normal errors. In Section 2 we formalize this idea by defining a semi-conjugate prior as the prior which, under a certain transformation, is equivalent to the standard conjugate prior in the Normal regression framework. Since our definition requires coherency with the uncontroversial reference case, but neither proportionality to the likelihood nor being closed under sampling, we use the term "semi-conjugate prior". There is another good reason to do so. Our semi-conjugate priors imply multivariate Student t marginal distributions of the regression coefficients a priori as well as a posteriori, so the functional form of the joint prior is (at least) partially preserved in the posterior.

Section 3 slightly broadens Zellner's (1976) t-F prior family, examines the conjugate properties of these priors in Zellner's model, and shows that the t-F prior is semi-conjugate, but for a different model. In Section 4 we find, after the appropriate transformation, a semi-conjugate prior for the standard multivariate t regression model considered by Zellner. This will allow us to conduct marginal posterior inference on the regression coefficients analytically under informative prior densities for the multivariate t model, an analysis which required numerical integration in Zellner (1976).

Some concluding remarks are grouped in a final section.

Appendix A explains our notation for probability density functions, and gives some useful properties, whereas Appendix B contains the posterior analysis for Zellner's model under the slightly generalized t-F prior.

2. Semi-Conjugate Prior Densities

Consider the standard linear regression model

$$y = X\beta + u, \tag{2.1}$$

where we observe the n dimensional vector y, X is an $n \times k$ known design matrix of rank k, and β is a vector of k unknown regression coefficients. In line with the literature, we take X to be a given matrix of observations, although we fully realize that in actual practice X will typically be random. The latter case can easily be incorporated within our present framework if X is independent of all the parameters in the implied conditional model, and we remark that such a condition holds under the slightly stronger assumption of a Bayesian cut (see Florens and Mouchart 1985). The stochastic nature of the model, however, does not follow from the usual practice of assuming a Normal distribution for the error vector u. Instead, we take a scale mixture of Normal densities, leading to the following sampling model

$$p(y|\beta,\tau^2) = \int_0^\infty p_N^n(y|X\beta,\psi(z)^2 \frac{1}{\tau^2} V)\ p(z|\beta,\tau^2)dz, \tag{2.2}$$

where V is a known $n \times n$ PDS matrix, z is a continuous positive random variable, that introduces the mixing on the scale parameter $\psi(z)^2/\tau^2$, and where $\psi(.)$ is differentiable in $(0,\infty)$. $\tau^2 > 0$, finally, is the "standard" unknown scalar precision parameter. For the notation of density functions, as well as certain useful properties, we refer to Appendix A.

If we choose $V = I_n$, we are in the class of spherical distributions as in JTC, and for any constant $\psi(z)$ the model simplifies to the usual Normal case.

However, for nonconstant $\psi(z)$, $p(y|\beta,\tau^2)$ is some non-Normal elliptical density. In particular, if we choose $\psi(z) = (\nu_0/z)^{\frac{1}{2}}$ and z as a χ^2 density with ν_o degrees of freedom (thus independent of β and τ^2), then

$$p(y|\beta,\tau^2) = p_S^n(y|\nu_0,X\beta,\tau^2V^{-1}), \tag{2.3}$$

which is an n-variate Student t density with ν_0 degrees of freedom (ν_0 known, $\nu_0 > 0$), location vector $X\beta$ (the mean if $\nu_0 > 1$), and precision matrix $\tau^2 V^{-1}$ (i.e. with covariance matrix $\nu_0(\nu_0-2)^{-1}\tau^{-2}V$ if $\nu_0 > 2$). In (2.3) both β and τ^2 are unknown parameters, on which we wish to conduct Bayesian inference. With $V = I_n$ the data density in (2.3) corresponds to the one in Zellner (1976), who parameterizes in terms of the variance parameter $\sigma^2 = \tau^{-2}$. The role of scale mixtures of Normal distributions follows from Kelker's (1970) Theorem 10, discussed in Dickey and Chen (1985, p. 169).

Let us first consider the elliptical model in (2.2), and specify a prior density $p(\beta,\tau^2)$ on its parameters. The joint density then becomes

$$p(y,z,\beta,\tau^2) = p_N^n(y|X\beta,\psi(z)^2 \frac{1}{\tau^2} V)\ p(z|\beta,\tau^2)\ p(\beta,\tau^2), \tag{2.4}$$

and we shall use the transformation $(\tau^2,z) \rightarrow (\varphi^2,z)$ with

$$\varphi^2 = \psi(z)^{-2}\tau^2, \tag{2.5}$$

i.e. φ^2 is just the inverted scale parameter of our Normal density in (2.2). This leads to

$$p(y,z,\beta,\varphi^2) = p_N^n(y|X\beta,\varphi^{-2}V)\ p(\beta,\varphi^2,z), \tag{2.6}$$

where y and z are <u>independent given (β,φ^2)</u>, and

$$p(\beta,\varphi^2,z) = \psi(z)^2 p(z|\beta,\tau^2)\ p(\beta,\tau^2) \tag{2.7}$$

with $\tau^2 = \psi(z)^2\varphi^2$.

Remark that where $p(y|\beta,\tau^2)$ is Student t in (2.3) since z had to be integrated out, $p(y|\beta,\varphi^2)$ is simply Normal and independent of z in (2.6), as the entire effect of z is included in φ^2. This, of course, implies that if, in (2.7), $p(\beta,\varphi^2)$ happens to be in the well-known Normal-gamma form, which is conjugate (in all three senses mentioned in the introduction) for the Normal linear regression model, and is denoted by

$$p_{NG}(\beta,\varphi^2) = p_N^k(\beta|\bar{\beta},\varphi^{-2}A^{-1})\ p_G(\varphi^2|\tfrac{e}{2},\tfrac{f}{2}), \tag{2.8}$$

with $e > 0$, $f > 0$, $\bar{\beta} \in \mathbb{R}^k$ and A a PDS $k\times k$ matrix, then the posterior $p(\beta,\varphi^2|y)$ will have the same functional form, given e.g. in (A.7). In the case that (2.8) applies, we can write prior and posterior densities for our original parameters (β,τ^2) as, respectively,

$$p(\beta,\tau^2) = \int_0^\infty p_{NG}(\beta,\varphi^2)m(z,\beta,\varphi^2)dz, \tag{2.9}$$

and

$$p(\beta,\tau^2|y) = \int_0^\infty p_{NG}(\beta,\varphi^2|y)m(z,\beta,\varphi^2)dz, \tag{2.10}$$

with φ^2 expressed as in (2.5) and using the same mixing function

$$m(z,\beta,\varphi^2) = \psi(z)^{-2}p(z|\beta,\varphi^2) \tag{2.11}$$

in both cases, where $p(z|\beta,\varphi^2)$ is derived from (2.7).
These considerations lead to the following definition.

<u>Definition:</u> For a given data density (2.2), every prior density $p(\beta,\tau^2)$ that corresponds to $p(\beta,\varphi^2)$ in the Normal-gamma form (2.8) is called a <u>semi-conjugate</u> prior density.

Note that our definition of a semi-conjugate prior for scale mixtures of Normal distributions is coherent with the usual notion of the conjugate prior for the Normal regression model. It also gives three obvious necessary conditions; if $p(\beta,\tau^2)$ is a semi-conjugate prior, then

(i) $p(\beta) = p_S^k(\beta|e,\bar{\beta},\frac{e}{f}A)$,

(ii) $p(\beta|y) = p_S^k(\beta|e_*,\beta_*,\frac{e_*}{f_*}A_*)$,

where

$$A_* = A + X'V^{-1}X, \quad \beta_* = A_*^{-1}(A\bar{\beta}+X'V^{-1}y),$$

$$f_* = f + \bar{\beta}'A\bar{\beta} + y'V^{-1}y - \beta_*'A_*\beta_*, \text{ and } e_* = e + n,$$

(iii) $E(\beta|y) = \beta_* = A_*^{-1}A\bar{\beta} + A_*^{-1}X'V^{-1}y,$

i.e. the posterior expectation of β is linear with respect to y.
For non-Normal distributions, (iii) may appear to be in conflict with the results in Goel and DeGroot (1980). In fact, our elliptical densities in (2.2) are not covered by their framework, which assumes that the covariance matrix of the errors is fully known. In addition, their assumption of independence can not be satisfied by non-Normal jointly elliptical errors (see Kelker 1970, Lemma 5).
Although a semi-conjugate prior density implies a Student t form for both the marginal prior and the marginal posterior of β, it need not preserve the form of the joint prior on (β,τ^2) in the posterior analysis, nor does it necessarily possess the functional form of the likelihood.

JTC's description of a "conjugate prior for the spherical family", given by their formulae (6)-(8), only corresponds to our (2.8), (2.9) and (2.11) under independence between (β,φ^2) and z. Note that JTC implicitly imposed two conflicting independence conditions: between z and (β,τ^2) in their formulation of the data density, and between z and (β,φ^2) in their description of a "conjugate prior". They did not explicitly distinguish between the conditional distributions of z given (β,φ^2) and given (β,τ^2), using G(z) (in their notation) for both. Therefore, they were led to the conclusion that Zellner's (1976) t-F prior results from (2.9) as a "conjugate" prior for his multivariate t model.
That the latter conclusion is not warranted will be seen in the next section, which will address the specific case of Student t errors.

3. Conjugate and Semi-Conjugate Properties of the t-F prior

From the form of the likelihood function in (2.3) (with $V = I_n$) Zellner (1976) deduced the following prior structure on β and $\sigma^2 = \tau^{-2}$:

$$p(\beta|\sigma^2) \propto \left[\frac{\nu_0\sigma^2+f}{\nu}\right]^{-\frac{k}{2}} \left[\nu + \frac{\nu}{\nu_0\sigma^2+f}\,(\beta-\bar{\beta})'A(\beta-\bar{\beta})\right]^{-\frac{\nu+k}{2}} \tag{3.1}$$

$$p(\sigma^2) \propto \left[\frac{\nu_0\sigma^2}{f}\right]^{\frac{\nu_0}{2}-1} \left[1 + \frac{\nu_0\sigma^2}{f}\right]^{-\frac{\nu}{2}}, \tag{3.2}$$

with the hyperparameters $\nu > \nu_0$, $f > 0$, $\bar{\beta} \in \mathbb{R}^k$ and A any PDS matrix of dimension $k\times k$, while $\nu_0 > 0$ is defined in (2.3). Note that Zellner wrote the prior in terms of $\nu_a = \nu-\nu_0$ and $s_a^2 = f/\nu_a$.

Zellner (1976, p. 403) stated that (3.1)-(3.2) constitutes a natural conjugate prior for the multivariate t case, and called it the product of a conditional multivariate t density in (3.1) and a marginal F density in (3.2). This accounts for the term t-F prior, which we shall use in spite of the fact that not σ^2 itself, but rather $\sigma^2(\nu-\nu_0)/f$ is F distributed with $(\nu_0,\nu-\nu_0)$ degrees of freedom. Formally, σ^2 has the following three-parameter inverted beta (IB) distribution (also known as beta prime):

$$p(\sigma^2) = p_{IB}\left[\sigma^2 \middle| \frac{\nu-\nu_0}{2}, \frac{\nu_0}{2}, \frac{f}{\nu_0}\right]. \tag{3.3}$$

By construction, the t-F prior has a similar functional form as the likelihood, and is thus natural-conjugate in this sense. In fact, exactly the same functional form would require that the exponent of the first factor in (3.2) be $\nu_0/2$ instead of $(\nu_0/2)-1$.

In order to cover this case as well, we slightly extend the class of priors to nonzero ℓ in

$$p(\beta,\sigma^2) = p_S^k\left[\beta \middle| \nu,\bar{\beta}, \frac{\nu}{\nu_0\sigma^2+f}\,A\right]\, p_{IB}\left[\sigma^2 \middle| \frac{\nu-\nu_0-\ell}{2}, \frac{\nu_0+\ell}{2}, \frac{f}{\nu_0}\right], \tag{3.4}$$

where now $\nu > \nu_0 + \ell > 0$ (and still $\nu_0 > 0$), so that no links between ν and ν_0 remain. Taking $\ell = 0$ leads to (3.1)-(3.2), while $\ell = 2$ ensures exactly the same functional form as the likelihood.

However, irrespective of the value we choose for ℓ, (3.4) is <u>not</u> a semi-conjugate prior structure for the model in (2.3), since e.g. the necessary condition (ii) is not satisfied: although the marginal prior of β has a Student t form, the posterior does not. Appendix B proves that $p(\beta|y)$ is, instead, a continuous mixture of t densities. In fact, this was essentially shown in Zellner's (1976) appendix, but only for $\ell = 0$ and subject to slight errors. Therefore, we have to conclude that the t-F prior in (3.1)-(3.2) [or (3.4)] cannot correspond to a Normal-gamma structure $p_{NG}(\beta,\varphi^2)$ for (2.3), contrary to JTC's statement in their Section 3.

The reason why the posterior does not preserve the functional form of the prior is rather obvious: the conditional posterior density $p(\beta|\sigma^2,y)$ is in the 2-0 poly-t form with one t kernel coming from the likelihood and one from the prior in (3.1). Since both kernels are usually different, they do not reduce to a single t density. Of course, this should not surprise us, since the model we treat is outside the exponential family of data densities. In order to obtain the marginal posterior density of β, we thus require numerical integration. Note also that the t-F prior in (3.4) can not be called conjugate in the sense of "giving linear posterior expectations" either, since neither $E(\beta|\sigma^2,y)$ nor $E(\beta|y)$ are linear in y.

We now set out to examine whether, within the class of linear regression models with Student t errors, we can find one for which the t-F prior in (3.4) is semi-conjugate and thus leads to analytical marginal posterior inference [condition (ii) in Section 2].

Using the definition from Section 2 we know that <u>any</u> prior on (β,τ^2) that implies a Normal-gamma prior structure for $p(\beta,\varphi^2)$ will be semi-conjugate for an elliptical model from the class in (2.2). Of course, the variable transformation from τ^2 to φ^2 in (2.5) will have consequences for the mixing density, and thus for the model.
Let us specify $\psi(z) = (\nu_0/z)^{\frac{1}{2}}$ as in Section 2, but let now z have a χ^2 density with $\nu_0+\ell$ degrees of freedom, conditional on (β,φ^2) and <u>not</u> on (β,τ^2), which led to the model in (2.3) for $\ell = 0$. Under these assumptions, it can be shown [using (A.1)-(A.4)] that the t-F prior in (3.4) is implied by the following Normal-gamma prior structure on (β,φ^2):

$$p(\beta,\varphi^2) = p_N^k(\beta|\bar{\beta},\varphi^{-2}A^{-1})\ p_G\left[\varphi^2\left|\frac{\nu-\nu_0-\ell}{2},\ \frac{f}{2}\right.\right], \tag{3.5}$$

which makes the prior in (3.4) semi-conjugate for (β,τ^2). It is, however, important to realize that the independence between z and (β,φ^2) now assumed introduces a dependence of z on (β,τ^2), expressed in

$$p(z|\beta,\tau^2) = p_G\left[z\left|\frac{\nu+k}{2},\ \frac{\nu_0+d_\beta\tau^2}{2\nu_0}\right.\right], \tag{3.6}$$

where ℓ no longer appears and we have defined

$$d_\beta = f + (\beta-\bar{\beta})'A(\beta-\bar{\beta}). \tag{3.7}$$

Note that, for $\ell = 0$, (3.5)-(3.7) strictly corresponds to the assumptions in Section 3 of JTC. However, JTC implicitly treat <u>different</u> models in their Sections 2 and 3, corresponding to different choices of the mixing density. For their Section 3 the relevant model is (2.2) mixed with the density in (3.6), which becomes (in terms of $\sigma^2 = \tau^{-2}$)

$$p(y|\beta,\sigma^2) = p_S^n\left[y\left|\nu+k, X\beta, \frac{\nu+k}{\nu_0\sigma^2+d_\beta}V^{-1}\right.\right], \tag{3.8}$$

instead of (2.3). So (3.8) is the Student t model for which the t-F prior in (3.4) is semi-conjugate, which does <u>not</u> hold for (2.3), nor for Zellner's (1976) model. This semi-conjugate prior can not be proportional to any likelihood following from (3.8), since the latter has a 1-1 poly-t form in β, given σ^2, and $p(\beta|\sigma^2)$ is just a Student t density. Note, however, that the fixed constants appearing in (3.8) are also the hyperparameters of the prior in (3.4). This particular feature implies that the kernel appearing in the denominator of the 1-1 poly-t likelihood will exactly cancel out with the Student t kernel of the prior for β given σ^2. Instead of a 2-1 poly-t density, the posterior $p(\beta|\sigma^2,y)$ will then simplifly to a Student t form. The joint posterior density $p(\beta,\sigma^2|y)$ will retain the t-F form of (3.4), but with the updated hyperparameters (ν_*,A_*,β_*,f_*) instead of $(\nu,A,\bar{\beta},f)$, where $\nu_* = \nu+n$ and the rest was defined in Section 2. Note that ν_0 and ℓ, also appearing in (3.4), are not affected by the sample information. The fact that the functional form is here retained for σ^2 as well is due to the assumed independence of z and φ^2, given β.

4. Semi-Conjugate Priors for the Standard Multivariate t Model

Let us now reparameterize the data density (3.8), obtained in the previous section, in the following way: define $\gamma = (\nu_0\sigma^2+d_\beta)/(\nu+k)$ and consider

$$p(y|\beta,\gamma) = p_S^n(y|\nu+k, X\beta, (\gamma V)^{-1}). \tag{4.1}$$

We now have a Student model with $\nu+k > k$ degrees of freedom (k is the dimension of β), ruling out very fat (e.g. Cauchy) tails.
Apart from that restriction, however, we have a standard multivariate Student t with the same location vector and relative precision matrix V^{-1} as in (2.3), but with a different variance parameter γ, now related to β through the inequality $\gamma > d_\beta/(\nu+k)$. The precision parameter is now $\delta = \gamma^{-1}$, instead of τ^2, and note that for any given value of β it is restricted to lie in the <u>bounded</u> interval $(0,\frac{\nu+k}{d_\beta})$. If we take the t-F prior in (3.4), then the reparameterization to (β,δ) leads to a beta distribution for δ given β:

$$p(\delta|\beta) = p_B(\delta|\frac{\mu+k}{2}, \frac{\nu-\mu}{2}, \frac{\nu+k}{d_\beta}), \tag{4.2}$$

where we have defined $\mu = \nu-\nu_0-\ell > 0$, in order to neutralize one of the hyperparameters (ν_0,ℓ). Due to the reparameterization from (β,σ^2) to (β,δ) only the sum $\nu_0+\ell$ affects data and prior densities now, so that μ replaces both ν_0 and ℓ. Since coherency was the principle underlying our definition of a semi-conjugate prior, it should not be affected by changes in the parameterization. Therefore, the combination of (4.2) with the marginal prior on β, resulting from (3.4), namely

$$p(\beta) = p_S^k(\beta|\mu,\bar{\beta},\frac{\mu}{f} A), \tag{4.3}$$

is still semi-conjugate for (β,δ) in (4.1). The implied marginal density of δ is nonzero over $(0,(\nu+k)/f)$ and the conditional densities of β, given values of $\delta \in (0,(\nu+k)/f)$, are nonzero over ellipsoids $(\beta-\bar{\beta})'A(\beta-\bar{\beta}) < ((\nu+k)/\delta)-f$.
The semi-conjugate structure in (4.2)-(4.3) puts a lower bound $f/(\nu+k)$ on the variance parameter γ and allows β values far from the prior mean (in

the metric induced by A) only for large values of γ, i.e. for noisy data processes. Contrary to the case considered in the previous section, the hyperparameters of the prior do not appear in the data density, so that there is room for prior elicitation. Apart from the obvious restrictions that A be PDS, $\mu > 0$ and $f > 0$, there is however, an upper bound on the degrees of freedom in (4.3) induced by (4.2), namely $\mu < \nu$. If the model in (4.1) has relatively fat tails, then the prior on β must have even fatter tails (with a difference of more than k degrees of freedom). In addition, we assume that $\mu > 2$, which provides us with an easy way of eliciting the hyperparameters $(\mu, f, \bar{\beta}, A)$ in the t-beta prior (4.2)-(4.3), given values of ν chosen for the data density. Indeed, from (4.3) we can then directly assign values to μ, $\bar{\beta}$ and A, given f, based on the prior mean and variance of β and the desired tail behaviour (e.g. through the existence of prior moments). More importantly, a choice for f can then be made on the basis of the prior mean for the variance parameter γ implied by (3.4):

$$E(\gamma) = \frac{\nu+k-2}{(\nu+k)(\mu-2)} f. \tag{4.4}$$

So, provided we are willing to choose $\mu > 2$ and to accept more than $\mu+k$ degrees of freedom for the multivariate t errors, we have an entirely analytical marginal posterior analysis for β and we know both prior and posterior means of γ.

In particular, we obtain for β the Student t posterior

$$p(\beta|y) = p_S^k(\beta|\mu_*, \beta_*, \frac{\mu_*}{f_*} A_*), \tag{4.5}$$

where the hyperparameters A_*, β_* and f_* are defined in Section 2 and $\mu_* = \mu+n$. The posterior mean of γ will be given by

$$E(\gamma|y) = \frac{1}{\nu+k}\left[f + \frac{\nu-\mu+tr(AA_*^{-1})}{\mu_*-2} f_*\right] \tag{4.6}$$

from the t-F posterior density of (β,σ^2). However, the t-beta prior in (4.2)-(4.3) is only <u>semi</u>-conjugate, so that the conditional posterior density of δ is no longer in a simple beta form, but is given by

$$p(\delta|\beta,y) = \left[1 + \frac{g_\beta}{d_\beta}\right]^{\frac{\mu_*+k}{2}} \left[1 + \frac{g_\beta}{\nu+k}\,\delta\right]^{-\frac{\nu_*+k}{2}} p_B\left[\delta\left|\frac{\mu_*+k}{2},\ \frac{\nu-\mu}{2},\ \frac{\nu+k}{d_\beta}\right.\right], \quad (4.7)$$

on the support $(0,(\nu+k)/d_\beta)$, where we have defined

$$g_\beta = (y-X\beta)'V^{-1}(y-X\beta). \quad (4.8)$$

Remark that the density in (4.7) can also be written as proportional to a product of a beta and an inverted beta kernel. In particular, we can single out the beta density corresponding to the prior $p(\delta|\beta)$ in (4.2) to obtain:

$$p(\delta|\beta,y) \propto p(\delta|\beta)\ p_{IB}(\delta|\frac{\nu+k}{2}-1,\ \frac{n}{2}+1,\ \frac{\nu+k}{g_\beta}), \quad (4.9)$$

where the correct support is induced by the prior. Note that, for $\delta = \gamma^{-1}$, the Student t model (4.1) can be presented as a mixture of Normal distributions $p_N^n(y|X\beta,[(\nu+k)/z\delta]V)$ with the mixing density

$$p(z|\beta,\delta) = p_G(z|\frac{\nu+k}{2},\ \frac{1}{2}), \quad (4.10)$$

independent of (β,δ). If we define the inverted scale parameter $\zeta = \delta z/(\nu+k)$, which is the counterpart of φ^2 in Section 2, then the prior density (4.2)-(4.3) and the mixing density (4.10) correspond to a Normal-gamma prior for (β,ζ) [i.e. (4.2)-(4.3) is semi-conjugate] and to

$$p(z|\beta,\zeta) = p_G(z-\zeta d_\beta|\frac{\nu-\mu}{2},\ \frac{1}{2}), \quad (4.11)$$

which is only nonzero when $z > \zeta d_\beta$, i.e. in the support of (β,δ), and implicitly imposes the other prior constraint that $\mu < \nu$. Clearly, now z depends on ζ, given β, which results in a difference between prior and posterior functional forms for δ in (4.2) and (4.9), respectively.

5. Concluding Remarks

In the general case of a linear regression model with errors distributed as a scale mixture of multivariate Normal densities, we have defined a semi-conjugate prior structure as the prior which, under a certain transformation, corresponds to the Normal-gamma (i.e. conjugate) prior for the Normal regression model. Since the transformation used does not affect the regression coefficient vector, any semi-conjugate prior leads to the same Student t marginal prior and posterior densities for this vector as in the reference case of a Normal regression model with conjugate prior.
Focusing on multivariate Student t errors in particular, the semi-conjugate prior (4.2)-(4.3) allows us to conduct a fully analytical marginal posterior analysis of the regression coefficients under informative (i.e. non-diffuse) prior assumptions. In addition, we can easily evaluate the posterior mean of the variance parameter γ.
The price to pay for this is that we have to restrict ourselves to a specific subset of the parameter space $\mathbb{R}^k \times \mathbb{R}_+$. The latter is a direct analogue of the pitfalls usually encountered in the natural conjugate framework for Normal models. These pitfalls can generally lead to a deceptively strong influence of the prior (see Richard 1973, p. 181), if we do not carefully assess all its (implicit) consequences.
Of course, our definition of semi-conjugate prior densities in Section 2 was formulated in terms of a broader class of elliptical models in (2.2), but it is by no means obvious that a semi-conjugate prior exists for any specific member of this class. For example, we have found the semi-conjugate t-beta prior for the multivariate t model, but only when the amount of degrees of freedom exceeds the dimension of the regression coefficient vector. Once we are willing to accept the latter restriction, we can analytically obtain the marginal posterior density of β and the posterior mean of the variance, which will suffice for most practical purposes. However, implicit restrictions on our semi-conjugate prior structures should not be overlooked.

Due to its internal coherency and its ensuing invariance with respect to reparameterizations, the concept of semi-conjugate prior densities seems intuitively appealing. It is also relatively easy to check and has interesting consequences [(i)-(iii) in Section 2]. It only coincides with one of the three definitions of "conjugate" mentioned in the introduction, namely the linearity of posterior expectations for the regression coefficients. Fully preserving the functional form of the prior typically seems a hopeless cause outside the exponential framework, and we do not suggest taking the same functional form as the likelihood in these cases. The latter property, rather useless in itself, is traded in for the possibility to conduct analytical inference on part of the parameters. In addition, we feel that the semi-conjugate concept has considerable scope for theoretically interesting results.

Appendix A. Probability density functions

A.1. Definitions

A k-variate Normal density on $x \in \mathbb{R}^k$ with mean vector $b \in \mathbb{R}^k$ and PDS k×k covariance matrix C:

$$p_N^k(x|b,C) = [(2\pi)^k|C|]^{-\frac{1}{2}}\exp -\frac{1}{2}(x-b)'C^{-1}(x-b).$$

A k-variate Student t density on $x \in \mathbb{R}^k$ with $r > 0$ degrees of freedom, location vector $b \in \mathbb{R}^k$ and PDS k×k precision matrix A:

$$p_S^k(x|r,b,A) = \frac{\Gamma(\frac{r+k}{2})}{\Gamma(\frac{r}{2})(r\pi)^{k/2}}\,|A|^{\frac{1}{2}}[1 + \frac{1}{r}(x-b)'A(x-b)]^{-\frac{r+k}{2}}.$$

A gamma density on $z > 0$ with $e,f > 0$:

$$p_G(z|e,f) = f^e[\Gamma(e)]^{-1}z^{e-1}\exp(-fz),$$

which becomes a χ^2_ν for $e = \frac{\nu}{2}$ and $f = \frac{1}{2}$.

A beta density on $v \in (0,c)$ with $a,b > 0$:

$$P_B(v|a,b,c) = \frac{\Gamma(a+b)}{c\Gamma(a)\Gamma(b)}\,(\frac{v}{c})^{a-1}(1- \frac{v}{c})^{b-1}.$$

A three-parameter inverted beta or beta prime density on $w > 0$ with $a,b,c > 0$ (see Zellner 1971, p. 376):

$$p_{IB}(w|a,b,c) = \frac{\Gamma(a+b)}{c\Gamma(a)\Gamma(b)}\,(\frac{w}{c})^{b-1}(1+ \frac{w}{c})^{-(a+b)},$$

a special case of which is an F density:

$$p_F(w|\nu_1,\nu_2) = p_{IB}\left[w\left|\frac{\nu_2}{2}, \frac{\nu_1}{2}, \frac{\nu_2}{\nu_1}\right.\right].$$

An m-0 poly-t density has a kernel which is composed of the product of m Student t kernels, whereas an m-n poly-t density is defined as the ratio of an m-0 poly-t to an n-0 poly-t density. Such densities, based on the work of Dickey (1968), are presented in Drèze (1977), whereas algorithms for their analysis are discussed in Richard and Tompa (1980).

A.2. Some properties

$$p_N^k(x|b,\frac{1}{v}A^{-1})\; p_G(v|\frac{e}{2},\frac{f}{2}) =$$

$$= p_S^k(x|e,b,\frac{e}{f}A)\; p_G(v|\frac{e+k}{2},\frac{f+(x-b)'A(x-b)}{2}). \tag{A.1}$$

$$p_S^k(x|a+e,b,\frac{a+e}{av+f}A)\; p_{IB}(v|\frac{e-\ell}{2},\frac{a+\ell}{2},\frac{f}{a}) =$$

$$= p_S^k(x|e-\ell,b,\frac{e-\ell}{f}A)\; p_{IB}(v|\frac{k+e-\ell}{2},\frac{a+\ell}{2},\frac{f+(x-b)'A(x-b)}{a}),$$

$$-a < \ell < e. \tag{A.2}$$

If $p(v,z) = p(v)p(z) = p_G(v|\frac{c}{2},\frac{d}{2})\; p_G(z|\frac{e}{2},\frac{f}{2})$ and $w = h\frac{v}{z}$ (h is a positive constant), then

$$p(w,z) = p_{IB}(w|\frac{e}{2},\frac{c}{2},\frac{fh}{d})\; p_G(z|\frac{c+e}{2},\frac{fh+dw}{2h}). \tag{A.3}$$

This is a simple generalization of the well-known theorem that the ratio of two independent χ^2 variables (both divided by their degrees of freedom) has an F distribution.

If $p(w) = p_{IB}(w|a,b,c)$ and $v = \frac{1}{w}$, then

$$p(v) = p_{IB}(v|b,a,\frac{1}{c}). \tag{A.4}$$

If $p(w) = p_{IB}(w|a,b,c)$ and $v = \frac{a}{bc}w$, then

$$p(v) = p_F(v|2b,2a). \tag{A.5}$$

$$p_N^k(x|b,\frac{a}{ws}A^{-1})\; p_G(w|\frac{e-\ell}{2},\frac{fs}{2a})\; p_G(s|\frac{a+\ell}{2},\frac{g}{2}) =$$

$$= p_S^k(x|e-\ell, b, \frac{e-\ell}{f} A)\ p_{IB}(w|\frac{a+\ell}{2}, \frac{k+e-\ell}{2}, \frac{ag}{f+(x-b)'A(x-b)}) \times$$

$$\times\ p_G(s|\frac{k+e+a}{2}, \frac{ag+[f+(x-b)'A(x-b)]w}{2a}),\ -a < \ell < e. \qquad (A.6)$$

$$p_N^n(y|X\beta, \varphi^{-2}V)\ p_N^k(\beta|\bar{\beta}, \varphi^{-2}A^{-1})\ p_G(\varphi^2|\frac{e}{2}, \frac{f}{2}) =$$

$$= p_N^k(\beta|\beta_*, \varphi^{-2}A_*^{-1})\ p_G(\varphi^2|\frac{e_*}{2}, \frac{f_*}{2})\ p_S^n(y|e, X\bar{\beta}, \frac{e}{f}(V+XA^{-1}X')^{-1}), \quad (A.7)$$

where:

$$e_* = e+n, \quad A_* = A + X'V^{-1}X, \quad \beta_* = A_*^{-1}(A\bar{\beta}+X'V^{-1}y)$$

$$f_* = f + \bar{\beta}'A\bar{\beta} + y'V^{-1}y - \beta_*'A_*\beta_* = f + (y-X\bar{\beta})'(V+XA^{-1}X')^{-1}(y-X\bar{\beta}).$$

This identity summarizes the standard Bayesian calculations in the case of Normal linear regression with Normal-gamma (i.e. conjugate) prior. See also Raiffa and Schlaifer (1961, p. 58).

Appendix B. Posterior corresponding to Zellner's (1976) model and a t-F prior

Consider the data density

$$p(y|\beta,\tau^2) = p_S^n(y|\nu_0, X\beta, \tau^2V^{-1}) = \int_0^\infty p_N^n(y|X\beta, \frac{\nu_0}{z\tau^2} V)\ p_G(z|\frac{\nu_0}{2}, \frac{1}{2})dz,$$

and the prior density [using the hyperparameters $(\nu_0, \ell, \mu, f, \bar{\beta}, A)$]

$$p(\beta,\tau^2) = p_S^k(\beta|\mu, \bar{\beta}, \frac{\mu}{f} A)\ p_{IB}(\tau^2|\frac{\nu_0+\ell}{2}, \frac{k+\mu}{2}, \frac{\nu_0}{d_\beta}).$$

When expressed in terms of $\sigma^2 = (\tau^2)^{-1}$ and using (3.7), these densities strictly correspond to (3.4). According to (A.6) one can write

$$p(\beta,\tau^2) = \int_0^\infty p_N^k(\beta|\bar{\beta}, \frac{\nu_0}{v\tau^2} A^{-1})\ p_G(\tau^2|\frac{\mu}{2}, \frac{f}{2\nu_0} v)\ p_G(v|\frac{\nu_0+\ell}{2}, \frac{1}{2})dv.$$

Now $p(y,\beta,\tau^2) = p(y|\beta,\tau^2)\ p(\beta,\tau^2)$ can be treated as a marginal density from the joint density

$$p(y,\beta,\tau^2,z,v) = p_N^n(y|X\beta,\frac{\nu_0}{z\tau^2}\ V)\ p_N^k(\beta|\bar{\beta},\frac{\nu_0}{v\tau^2}\ A^{-1})\ p_G(\tau^2|\frac{\mu}{2},\ \frac{f}{2\nu_0}\ v)$$

$$\times\ p_G(v|\frac{\nu_0+\ell}{2},\ \frac{1}{2})\ p_G(z|\frac{\nu_0}{2},\ \frac{1}{2}),$$

where v and z are (marginally) independent. Replacing v by $\lambda = \frac{v}{z}$, with the use of (A.3), and applying (A.7) conditionally on (λ,z), one obtains the following factorization:

$$p(y,\beta,\tau^2,z,\lambda) = p(\lambda)\ p(z|\lambda)\ p(\beta,\tau^2|y,z,\lambda)\ p(y|z,\lambda)$$

where

$$p(\lambda) = p_{IB}(\lambda|\frac{\nu_0}{2},\ \frac{\nu_0+\ell}{2},\ 1),$$

$$p(z|\lambda) = p_G(z|\frac{2\nu_0+\ell}{2},\ \frac{1+\lambda}{2}),$$

$$p(\beta,\tau^2|y,\lambda,z) = p_N^k(\beta|\beta_\lambda,\frac{\nu_0}{z\tau^2}\ M_\lambda^{-1})\ p_G(\tau^2|\frac{\mu_*}{2},\ \frac{f_\lambda z}{2\nu_0}),$$

$$p(y|z,\lambda) = p_S^n(y|\mu,X\bar{\beta},\frac{\mu}{f}(\lambda V+XA^{-1}X')^{-1}) = p(y|\lambda),$$

$$M_\lambda = X'V^{-1}X + \lambda A,\qquad \beta_\lambda = M_\lambda^{-1}(\lambda A\bar{\beta}+X'V^{-1}y),$$

$$f_\lambda = \lambda f + (y-X\bar{\beta})'(V + \frac{1}{\lambda}\ XA^{-1}X')^{-1}(y-X\bar{\beta}),$$

and $\mu_* = \mu+n$.

Integrating out z, with the use of (A.6), leads to

$$p(y,\beta,\tau^2,\lambda) = p_S^k(\beta|\mu_*,\beta_\lambda,\ \frac{\mu_*}{f_\lambda}\ M_\lambda)\ \times$$

$$p_{IB}(\tau^2|\nu_0+\frac{\ell}{2},\ \frac{k+\mu_*}{2},\ \frac{\nu_0(1+\lambda)}{f_\lambda+(\beta-\beta_\lambda)'M_\lambda(\beta-\beta_\lambda)})\ p(y|\lambda)\ p(\lambda). \tag{B.1}$$

Using (A.4) and (A.2), one can equivalently write

$$p(y,\beta,\sigma^2,\lambda) = p_S^k(\beta|2\nu_0+\mu_*+\ell,\beta_\lambda, \frac{2\nu_0+\mu_*+\ell}{\nu_0\sigma^2(1+\lambda)+f_\lambda} M_\lambda) \times$$

$$p_{IB}(\sigma^2|\frac{\mu_*}{2}, \nu_0+\frac{\ell}{2}, \frac{f_\lambda}{\nu_0(1+\lambda)})\ p(y|\lambda)\ p(\lambda). \tag{B.2}$$

Note that, conditionally on y <u>and</u> λ, (β,σ^2) has a t-F distribution, thus the posterior density $p(\beta,\sigma^2|y)$ is a continuous mixture of t-F densities with the following mixing density:

$$p(\lambda|y) \propto p(y|\lambda)\ p(\lambda) \propto |M_\lambda|^{-\frac{1}{2}} f_\lambda^{-\frac{\mu_*}{2}} \lambda^{\frac{k+\nu_0+\mu+\ell}{2}-1} (1+\lambda)^{-\frac{2\nu_0+\ell}{2}}.$$

The first factor on the r.h.s. of (B.1) equals $p(\beta|y,\lambda)$, and the second factor on the r.h.s. of (B.2) is $p(\sigma^2|y,\lambda)$. For ℓ=0, implying that $\mu = \nu_a$ in Zellner's notation, these two densities, as well as $p(\lambda|y)$, are expected to correspond strictly to those obtained by Zellner (1976, appendix). Yet, they differ, mainly in degrees-of-freedom parameters. These differences are caused by slight errors in Zellner's calculations. After his (A.5), Zellner states that the Jacobian of the transformation from $(\sigma^2,\tau,\vartheta)$ to (σ^2,τ,λ), where $\lambda = \tau^2/\vartheta^2$, is proportional to $\lambda^{3/2}/\tau$; obviously, it is proportional to $\tau/\lambda^{3/2}$. (Differences in notation: Zellner's τ, ϑ and $\nu s^2+g(\lambda)$ are equal to our $z^{-\frac{1}{2}}$, $v^{-\frac{1}{2}}$ and f_λ, respectively; thus, Zellner's $\lambda = \tau^2/\vartheta^2$ is the same as our $\lambda = v/z$). In addition, the factorization of Zellner's (A.8a) into (A.8b) and (A.8c) is not fully correct; (A.8c) should be divided by $[(1+\lambda)\nu_0]/[\nu s^2+g(\lambda)]$, from the normalizing constant of (A.8b). When one takes into account the true value of the Jacobian and the correct factorization of $p(\sigma^2,\lambda|y)$, one will obtain the same $p(\sigma^2|y,\lambda)$ and $p(\lambda|y)$ as we do for ℓ=0. The remaining differences are due to obvious misprints; in Zellner's formula for c_λ, after (A.9), $\bar{\beta}'A\bar{\beta}$ should be multiplied by λ, as it is in his (A.4), and the exponent of $(1+\lambda)$ in (A.10) should read $-\nu_0$, not ν_0. Correcting these misprints one obtains from Zellner's (A.10) the same $p(\beta|y,\lambda)$ as ours for ℓ=0.

References

Berger, J.O. (1985), Statistical Decision Theory and Bayesian Analysis, New York: Springer Verlag.

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

DeGroot, M.H. (1970), Optimal Statistical Decisions, New York: McGraw-Hill.

Diaconis, P., and Ylvisaker, D. (1979), "Conjugate Priors for Exponential Families", The Annals of Statistics, 7, 269-281.

——— (1985), "Quantifying Prior Opinion", in Bayesian Statistics 2, eds. J.M. Bernardo, M.H. DeGroot, D.V. Lindley, and A.F.M. Smith, Amsterdam: North-Holland, pp. 133-156.

Dickey, J.M. (1968), "Three Multidimensional-Integral Identities With Bayesian Applications", The Annals of Mathematical Statistics, 39, 1615-1627.

Dickey, J.M., and Chen, C.-H. (1985), "Direct Subjective-Probability Modelling Using Ellipsoidal Distributions", in Bayesian Statistics 2, eds. J.M. Bernardo, M.H. DeGroot, D.V. Lindley, and A.F.M. Smith, Amsterdam: North Holland, pp. 157-182.

Drèze, J.H. (1977), "Bayesian Regression Analysis Using Poly-t Densities", Journal of Econometrics, 6, 329-354.

Florens, J.P., and Mouchart, M. (1985), "Conditioning in Dynamic Models", Journal of Time Series Analysis, 6, 15-34.

Goel, P., and DeGroot, M.H. (1980), "Only Normal Distributions Have Linear Posterior Expectations in Linear Regression", Journal of the American Statistical Association, 75, 895-900.

Hartigan, J.A. (1983), Bayes Theory, New York: Springer Verlag.

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

Kelker, D. (1970), "Distribution Theory of Spherical Distributions and a Location-Scale Parameter Generalization", Sankhyā, Ser. A, 32, 419-430.

Osiewalski, J. (1989), "A Note on Bayesian Inference in a Regression Model with Elliptical Errors", CORE Discussion Paper 8940, Université Catholique de Louvain.

Press, S.J. (1972), Applied Multivariate Analysis, New York: Holt, Rinehart and Winston.

Raiffa, H., and Schlaifer, R. (1961), Applied Statistical Decision Theory, Boston: Harvard University.

Richard, J.F. (1973), Posterior and Predictive Densities for Simultaneous Equation Models, New York: Springer Verlag.

Richard, J.F., and Tompa, H. (1980), "On the Evaluation of Poly-t Density Functions", Journal of Econometrics, 12, 331-351.

Zellner, A. (1971), An Introduction to Bayesian Inference in Econometrics, New York: John Wiley.

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:

No.	Author(s)	Title
8801	Th. van de Klundert and F. van der Ploeg	Fiscal Policy and Finite Lives in Interdependent Economies with Real and Nominal Wage Rigidity
8802	J.R. Magnus and B. Pesaran	The Bias of Forecasts from a First-order Autoregression
8803	A.A. Weber	The Credibility of Monetary Policies, Policymakers' Reputation and the EMS-Hypothesis: Empirical Evidence from 13 Countries
8804	F. van der Ploeg and A.J. de Zeeuw	Perfect Equilibrium in a Model of Competitive Arms Accumulation
8805	M.F.J. Steel	Seemingly Unrelated Regression Equation Systems under Diffuse Stochastic Prior Information: A Recursive Analytical Approach
8806	Th. Ten Raa and E.N. Wolff	Secondary Products and the Measurement of Productivity Growth
8807	F. van der Ploeg	Monetary and Fiscal Policy in Interdependent Economies with Capital Accumulation, Death and Population Growth
8901	Th. Ten Raa and P. Kop Jansen	The Choice of Model in the Construction of Input-Output Coefficients Matrices
8902	Th. Nijman and F. Palm	Generalized Least Squares Estimation of Linear Models Containing Rational Future Expectations
8903	A. van Soest, I. Woittiez, A. Kapteyn	Labour Supply, Income Taxes and Hours Restrictions in The Netherlands
8904	F. van der Ploeg	Capital Accumulation, Inflation and Long-Run Conflict in International Objectives
8905	Th. van de Klundert and A. van Schaik	Unemployment Persistence and Loss of Productive Capacity: A Keynesian Approach
8906	A.J. Markink and F. van der Ploeg	Dynamic Policy Simulation of Linear Models with Rational Expectations of Future Events: A Computer Package
8907	J. Osiewalski	Posterior Densities for Nonlinear Regression with Equicorrelated Errors
8908	M.F.J. Steel	A Bayesian Analysis of Simultaneous Equation Models by Combining Recursive Analytical and Numerical Approaches

No.	Author(s)	Title
8909	F. van der Ploeg	Two Essays on Political Economy (i) The Political Economy of Overvaluation (ii) Election Outcomes and the Stockmarket
8910	R. Gradus and A. de Zeeuw	Corporate Tax Rate Policy and Public and Private Employment
8911	A.P. Barten	Allais Characterisation of Preference Structures and the Structure of Demand
8912	K. Kamiya and A.J.J. Talman	Simplicial Algorithm to Find Zero Points of a Function with Special Structure on a Simplotope
8913	G. van der Laan and A.J.J. Talman	Price Rigidities and Rationing
8914	J. Osiewalski and M.F.J. Steel	A Bayesian Analysis of Exogeneity in Models Pooling Time-Series and Cross-Section Data
8915	R.P. Gilles, P.H. Ruys and J. Shou	On the Existence of Networks in Relational Models
8916	A. Kapteyn, P. Kooreman and A. van Soest	Quantity Rationing and Concavity in a Flexible Household Labor Supply Model
8917	F. Canova	Seasonalities in Foreign Exchange Markets
8918	F. van der Ploeg	Monetary Disinflation, Fiscal Expansion and the Current Account in an Interdependent World
8919	W. Bossert and F. Stehling	On the Uniqueness of Cardinally Interpreted Utility Functions
8920	F. van der Ploeg	Monetary Interdependence under Alternative Exchange-Rate Regimes
8921	D. Canning	Bottlenecks and Persistent Unemployment: Why Do Booms End?
8922	C. Fershtman and A. Fishman	Price Cycles and Booms: Dynamic Search Equilibrium
8923	M.B. Canzoneri and C.A. Rogers	Is the European Community an Optimal Currency Area? Optimal Tax Smoothing versus the Cost of Multiple Currencies
8924	F. Groot, C. Withagen and A. de Zeeuw	Theory of Natural Exhaustible Resources: The Cartel-Versus-Fringe Model Reconsidered

No.	Author(s)	Title
8925	O.P. Attanasio and G. Weber	Consumption, Productivity Growth and the Interest Rate
8926	N. Rankin	Monetary and Fiscal Policy in a 'Hartian' Model of Imperfect Competition
8927	Th. van de Klundert	Reducing External Debt in a World with Imperfect Asset and Imperfect Commodity Substitution
8928	C. Dang	The D_1-Triangulation of R^n for Simplicial Algorithms for Computing Solutions of Nonlinear Equations
8929	M.F.J. Steel and J.F. Richard	Bayesian Multivariate Exogeneity Analysis: An Application to a UK Money Demand Equation
8930	F. van der Ploeg	Fiscal Aspects of Monetary Integration in Europe
8931	H.A. Keuzenkamp	The Prehistory of Rational Expectations
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

No.	Author(s)	Title
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
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

No.	Author(s)	Title
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