SEMI-CONJUGATE PRIOR DENSITIES IN MULTIVARIATE T REGRESSION MODELS

by Jacek Osiewalski and Mark F.J. Steel

February 1990
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 $\sigma^2$) 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

\[ y = X\beta + u, \quad (2.1) \]

where we observe the \( n \) dimensional vector \( y \), \( X \) is an \( n \times k \) known design matrix of rank \( k \), and \( \beta \) 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

\[ p(y|\beta, \tau^2) = \int_0^\infty p_N^n(y|X\beta, W(z)^2 \tau^2 V) p(z|\beta, \tau^2) dz, \quad (2.2) \]

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 \( W(z)^2 \tau^2 \), and where \( W(.) \) 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 \( W(z) \) the model simplifies to the usual Normal case.

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

\[ p(y|\beta, \tau^2) = p_S^n(y|\nu_0, X\beta, \tau^2 V^{-1}), \quad (2.3) \]
which is an n-variate Student t density with \( \nu_0 \) degrees of freedom (\( \nu_0 \) known, \( \nu_0 > 0 \)), location vector \( X \beta \) (the mean if \( \nu_0 > 1 \)), and precision matrix \( \nu_0^{-1} \) \( \Sigma^{-1} \) (i.e. with covariance matrix \( \nu_0(\nu_0-2)^{-1}\Sigma^{-1} \) if \( \nu_0 > 2 \)). In (2.3) both \( \beta \) and \( \tau^2 \) 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

\[
p(y, z, \beta, \tau^2) = p^n(y | X \beta, \nu(z)^2 \frac{1}{\tau^2} \Sigma) p(z | \beta, \tau^2) p(\beta, \tau^2),
\]

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

\[
\phi^2 = \nu(z)^{-2} \tau^2.
\]

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

\[
p(y, z, \beta, \phi^2) = p^n(y | X \beta, \phi^{-2} \Sigma) p(\beta, \phi^2, z),
\]

where \( y \) and \( z \) are independent given \( (\beta, \phi^2) \), and

\[
p(\beta, \phi^2, z) = \nu(z)^2 p(z | \beta, \tau^2) p(\beta, \tau^2)
\]

with \( \tau^2 = \nu(z)^2 \phi^2 \).

Remark that where \( p(y | \beta, \tau^2) \) is Student t in (2.3) since \( z \) had to be integrated out, \( p(y | \beta, \phi^2) \) is simply Normal and independent of \( z \) in (2.6), as the entire effect of \( z \) is included in \( \phi^2 \). This, of course, implies that if, in (2.7), \( p(\beta, \phi^2) \) 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
\[ p_{\text{NG}}(\beta, \sigma^2) = p_{N}^k(\beta|\bar{\beta}, \sigma^{-2\bar{A}^{-1}}) p_{G}(\sigma^2|\frac{\epsilon}{2}, \frac{1}{2}), \tag{2.8} \]

with \( e > 0, f > 0, \bar{\beta} \in \mathbb{R}^k \) and \( A \) a PDS \( k \times k \) matrix, then the posterior \( p(\beta, \sigma^2|y) \) 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 \((\beta, \tau^2)\) as, respectively,

\[ p(\beta, \tau^2) = \int_0^\infty p_{\text{NG}}(\beta, \sigma^2) m(z, \beta, \sigma^2) dz, \tag{2.9} \]

and

\[ p(\beta, \tau^2|y) = \int_0^\infty p_{\text{NG}}(\beta, \sigma^2|y) m(z, \beta, \sigma^2) dz, \tag{2.10} \]

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

\[ m(z, \beta, \sigma^2) = \gamma(z)^{-2} p(z|\beta, \sigma^2) \tag{2.11} \]

in both cases, where \( p(z|\beta, \sigma^2) \) is derived from (2.7).

These considerations lead to the following definition.

**Definition:** For a given data density (2.2), every prior density \( p(\beta, \tau^2) \) that corresponds to \( p(\beta, \sigma^2) \) in the Normal-gamma form (2.8) is called a semi-conjugate 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) \[ p(\beta) = p_{S}^k(\beta|\epsilon, \bar{\beta}, e, A), \]

(ii) \[ p(\beta|y) = p_{S}^k(\beta|\epsilon, \beta, e, \bar{T}, A), \]

where
\[
A_* = A + X'V^{-1}X, \quad \beta_* = \beta_0^{-1}(A\beta + X'V^{-1}y),
\]
\[
f_* = f + \tilde{\beta}'A\beta + y'V^{-1}y - \beta'_0A\beta_0, \quad \text{and} \quad e_* = e + n,
\]

(iii) \[E(\beta|y) = \beta_* = \beta_0^{-1}A\beta + \beta_0^{-1}X'V^{-1}y,\]
i.e. the posterior expectation of \(\beta\) 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 cannot 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 \(\beta\), it need not preserve the form of the joint prior on \((\beta, \tau^2)\) 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 \((\beta, \tau^2)\) and \(z\). Note that JTC implicitly imposed two conflicting independence conditions: between \(z\) and \((\beta, \tau^2)\) in their formulation of the data density, and between \(z\) and \((\beta, \tau^2)\) in their description of a "conjugate prior". They did not explicitly distinguish between the conditional distributions of \(z\) given \((\beta, \tau^2)\) and given \((\beta, \tau^2)\), 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 \(\beta\) and \(\sigma^2 = \tau^{-2}:\)
\[ p(\beta|\sigma^2) \propto \left( \frac{\nu_0 \sigma^2 + \bar{\beta}}{\nu} \right)^{-\frac{k}{2}} \left[ \frac{\nu}{\nu_0 \sigma^2 + f} (\beta - \bar{\beta})' A (\beta - \bar{\beta}) \right]^{-\frac{\nu + k}{2}} \]  

(3.1)

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

(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 \( \sigma^2 \) itself, but rather \( \sigma^2 (\nu-\nu_0)/f \) is F distributed with \((\nu_0,\nu-\nu_0)\) degrees of freedom. Formally, \( \sigma^2 \) has the following three-parameter inverted beta (IB) distribution (also known as beta prime):

\[ p(\sigma^2) = p_{\text{IB}} \left[ \sigma^2 \mid \frac{\nu-\nu_0}{2}, \frac{\nu_0}{2}, \frac{f}{\nu_0} \right]. \]  

(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 \( \lambda \) in

\[ p(\beta, \sigma^2) = p_S \left[ \beta \mid \nu, \bar{\beta}, \frac{\nu}{\nu_0 \sigma^2 + f} A \right] p_{\text{IB}} \left[ \sigma^2 \mid \frac{\nu-\nu_0-\lambda}{2}, \frac{\nu_0+\lambda}{2}, \frac{f}{\nu_0} \right], \]  

(3.4)

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

However, irrespective of the value we choose for \( \lambda \), (3.4) is not 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 $\beta$ 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 $l = 0$ 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 $p_{NG}(\beta, \nu^2)$ 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 $\beta$, we thus require numerical integration. Note also that the t-F prior in (3.4) cannot 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 any prior on $(\beta, \nu^2)$ that implies a Normal-gamma prior structure for $p(\beta, \nu^2)$ will be semi-conjugate for an elliptical model from the class in (2.2). Of course, the variable transformation from $\nu^2$ to $\nu^2$ in (2.5) will have consequences for the mixing density, and thus for the model.

Let us specify $y(z) = (\nu_0/z)^{*}$ as in Section 2, but let now $z$ have a $\chi^2$ density with $\nu_0 + l$ degrees of freedom, conditional on $(\beta, \nu^2)$ and not on $(\beta, \nu^2)$, which led to the model in (2.3) for $l = 0$. 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 $(\beta, \nu^2)$:

$$p(\beta, \nu^2) = p_N^k(\beta|\nu, \nu^{-2}A^{-1}) p_G^k(\nu^2|\nu_0 - l, 2, \nu_0 - l, 2).$$  \tag{3.5}
which makes the prior in (3.4) semi-conjugate for \((\beta, \tau^2)\). It is, however, important to realize that the independence between \(z\) and \((\beta, \tau^2)\) now assumed introduces a dependence of \(z\) on \((\beta, \tau^2)\), expressed in

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

(3.6)

where \(l\) no longer appears and we have defined

\[
d_\beta = f + (\beta - \tilde{\beta})'A(\beta - \tilde{\beta}).
\]

(3.7)

Note that, for \(l = 0\), (3.5)-(3.7) strictly corresponds to the assumptions in Section 3 of JTC. However, JTC implicitly treat different 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 \(\sigma^2 = \tau^{-2}\))

\[
p(y|\beta, \sigma^2) = p^n_S\left[y \left| \nu + k, X\beta, \frac{\nu + k}{\nu_0 \sigma^2 + d_\beta} \right.\right],
\]

(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 not 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 \(\beta\), given \(\sigma^2\), 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 \(\beta\) given \(\sigma^2\). Instead of a 2-1 poly-t density, the posterior \(p(\beta|\sigma^2, y)\) will then simplify 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 \((\nu_*, A_*, \beta_*, f_*)\) instead of \((\nu, A, \tilde{\beta}, f)\), where \(\nu_* = \nu + n\) and the rest was defined in Section 2. Note that \(\nu_0\) and \(l\), also appearing in (3.4), are not affected by the sample information. The fact that the functional form is here retained for \(\sigma^2\) as well is due to the assumed independence of \(z\) and \(\tau^2\), given \(\beta\).
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

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

We now have a Student model with $\nu + k > k$ degrees of freedom ($k$ is the dimension of $\beta$), 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 $\gamma$, now related to $\beta$ through the inequality $\gamma > d_\beta / (\nu + k)$. The precision parameter is now $\delta = \gamma^{-1}$, instead of $\tau^2$, and note that for any given value of $\beta$ it is restricted to lie in the bounded interval $(0, \nu + k)$. If we take the $t$-F prior in (3.4), then the reparameterization to $(\beta, \delta)$ leads to a beta distribution for $\delta$ given $\beta$:

$$p(\delta|\beta) = p_B(\delta | \frac{\nu^* k}{2}, \frac{\nu - \mu}{2}, \frac{\nu^* k}{d_\beta})$$ \hspace{1cm} (4.2)

where we have defined $\mu = \nu - \nu_0 - l > 0$, in order to neutralize one of the hyperparameters ($\nu_0, l$). Due to the reparameterization from $(\beta, \sigma^2)$ to $(\beta, \delta)$ only the sum $\nu^* + l$ affects data and prior densities now, so that $\mu$ replaces both $\nu_0$ 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 $\beta$, resulting from (3.4), namely

$$p(\beta) = p_S^k(\beta | \mu, \tilde{\beta}, \mu, A).$$  \hspace{1cm} (4.3)

is still semi-conjugate for $(\beta, \delta)$ in (4.1). The implied marginal density of $\delta$ is nonzero over $(0, (\nu + k)/f)$ and the conditional densities of $\beta$, given values of $\delta \in (0, (\nu + k)/f)$, are nonzero over ellipsoids $(\beta - \tilde{\beta})' A (\beta - \tilde{\beta}) < ((\nu + k)/\delta) - f$.

The semi-conjugate structure in (4.2)-(4.3) puts a lower bound $f / (\nu + k)$ on the variance parameter $\gamma$ and allows $\beta$ values far from the prior mean (in
the metric induced by A) only for large values of \( y \), 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 \( \mu < v \). If the model in (4.1) has relatively fat tails, then the prior on \( \beta \) 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 \( v \) chosen for the data density. Indeed, from (4.3) we can then directly assign values to \( \mu \), \( \bar{\beta} \) and \( A \), given \( f \), based on the prior mean and variance of \( \beta \) 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 \( \gamma \) implied by (3.4):

\[
E(\gamma) = \frac{v+k-2}{(v+k)(\mu-2)} f. \tag{4.4}
\]

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 \( \beta \) and we know both prior and posterior means of \( \gamma \).

In particular, we obtain for \( \beta \) the Student t posterior

\[
p(\beta|y) = p_\beta(\beta|\mu_*, \beta_*, f_*, A_*). \tag{4.5}
\]

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

\[
E(\gamma|y) = \frac{1}{v+k} \left[ f + \frac{v-\mu+\text{tr}(AA_*^{-1})}{\mu_*^2} f_* \right] \tag{4.6}
\]
from the t-F posterior density of \((\beta, \sigma^2)\). However, the t-beta prior in (4.2)-(4.3) is only semi-conjugate, so that the conditional posterior density of \(\delta\) is no longer in a simple beta form, but is given by

\[
p(\delta | \beta, y) = \left[ 1 + \frac{g_\beta}{d_\beta} \right]^{-\frac{\nu+2}{2}} \left[ 1 + \frac{g_\beta}{\nu+1} \delta \right]^{-\frac{\nu+2}{2}} P_B \left[ \frac{\delta}{2}, \frac{\nu-\frac{\mu}{2}}, \frac{\nu+k}{d_\beta} \right],
\]

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

\[
g_\beta = (y-X\beta)'(y-X\beta)^{-1}. \tag{4.8}
\]

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:

\[
p(\delta | \beta, y) \propto p(\delta | \beta) P_B(\delta | \nu + k/2 - 1, n/2 + 1, \nu + k/g_\beta), \tag{4.9}
\]

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

\[
p(z | \beta, \delta) = P_G(z | \nu + k/2, 1/2), \tag{4.10}
\]

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

\[
p(z | \beta, \xi) = P_G(z - \xi d_\beta | \nu - \mu/2, 1/2), \tag{4.11}
\]

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

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 $\beta$ 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 inter-
teresting 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 coeffi-
cients. 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 possibili-
ty to conduct analytical inference on part of the parameters. In addition,
we feel that the semi-conjugate concept has considerable scope for theore-
tically 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 $k \times k$ covariance matrix $C$:

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

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 $k \times k$ precision matrix $A$:

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

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

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

which becomes a $\chi^2$ for $e = \frac{v}{2}$ and $f = \frac{1}{2}$.

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

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

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

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

a special case of which is an $F$ density:

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

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

\[ p_S^k(x|a+e, b, \frac{a+e}{a+\bar{f}} A) p_{IB}^e(v|\frac{e-\bar{f}}{2}, \frac{a+\bar{f}}{a}) = \]
\[ = p_S^k(x|e-\bar{f}, b, \frac{e-\bar{f}}{2} A) p_{IB}^e(v|\frac{k+e-\bar{f}}{2}, \frac{a+\bar{f}}{a}, \frac{f+(x-b)'A(x-b)}{a}) \]
\[ -a < \bar{f} < e. \]  
(A.2)

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

\[ p(w,z) = p_{IB}^e(w|\frac{e}{2}, \frac{c}{2}, \frac{f h}{d}) p_G^e(z|\frac{c+e}{2}, \frac{f h + d w}{2}). \]  
(A.3)

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

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

\[ p(v) = p_{IB}^e(v|b, a, \frac{1}{c}). \]  
(A.4)

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

\[ p(v) = p_F(v|2b, 2a). \]  
(A.5)

\[ p_N^k(x|b, \frac{a}{ws} A^{-1}) p_G^e(w|\frac{e-\bar{f}}{2}, \frac{f s}{2a}) p_G^e(s|\frac{a+\bar{f}}{2}, \frac{g}{2}) = \]
\[ p_S(x|e-\lambda, b, \frac{e-\lambda}{f} A) \cdot p_{IB}(w|\frac{a+\mu}{2}, \frac{k+\mu}{2}, \frac{ag}{f+(x-b)'A(x-b)}) \times \\
p_G(s|\frac{k+e+a}{2}, \frac{ag+[f+(x-b)'A(x-b)]w}{2a}) \quad -a < \lambda < e. \quad (A.6) \]

\[ p^n_N(y|x\beta, \sigma^{-2} \nu) \cdot p^k_N(\beta|\bar{\beta}, \sigma^{-2} \nu^{-1}) \cdot p_G(\sigma^2|\frac{e}{2}, \frac{f}{2}) = \\
= p^n_N(\beta|\beta_*, \sigma^{-2} \nu^{-1}) \cdot p_G(\sigma^2|\frac{e}{2}, \frac{f}{2}) \cdot p_S(y|x\beta, \sigma^{-2} \nu^{-1})^{-1}, \quad (A.7) \]

where:

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

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

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

\[ p(y|\beta, \tau^2) = p^n_S(y|\nu_0, X\bar{\beta}, \tau^2 V^{-1}) = \int_0^\infty p^n_N(y|x\beta, \frac{\nu_0}{z}\frac{1}{2})\cdot p_G(z|\frac{\nu_0}{2}, \frac{1}{2}) \text{d}z, \]

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

\[ p(\beta, \tau^2) = p^k_S(\beta|\nu, \bar{\beta}, \lambda A) \cdot p_{IB}(\tau^2|\frac{\nu_0+\lambda}{2}, \frac{k+\mu}{2}, \frac{\nu_0}{d_\beta}). \]

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

\[ p(\beta, \tau^2) = \int_0^\infty p^k_S(\beta|\bar{\beta}, \frac{\nu_0}{\nu\tau^2} A^{-1}) \cdot p_G(\tau^2|\frac{\mu}{2}, \frac{f}{2\nu_0} \cdot \nu) \cdot p_G(\nu|\frac{\nu_0+\lambda}{2}, \frac{1}{2}) \text{d}\nu. \]
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

\[
p(y, \beta, \tau^2, z, v) = P_N(y|x, \beta, \frac{\nu_0}{z\tau^2}, v) P_N(\beta|\bar{\beta}, \frac{\nu_0}{\nu_2}, A^{-1}) P_G(\tau^2|\frac{M}{2}, \frac{f}{2\nu_0}, v)
\]

\[
\times P_G(v|\frac{\nu_0 + l}{2}, \frac{1}{2}) P_G(z|\frac{\nu_0}{2}, \frac{1}{2})
\]

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 \( (\lambda, z) \), 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

\[
p(\lambda) = P_{IB}(\lambda| \frac{\nu_0}{2}, \frac{\nu_0 + l}{2}, 1),
\]

\[
p(z|\lambda) = P_G(z|\frac{2\nu_0 + l}{2}, \frac{1+\lambda}{2}),
\]

\[
p(\beta, \tau^2|y, \lambda, z) = P_N(\beta|\bar{\beta}, \frac{\nu_0}{z\tau^2}, M_{\lambda}^{-1}) P_G(\tau^2|\frac{M_{\lambda}}{2}, \frac{f_{\lambda} z}{2\nu_0}),
\]

\[
p(y|z, \lambda) = P_S(y|\mu, x\bar{\beta}, f_{\lambda}(\lambda V + XA^{-1} X')^{-1}) = p(y|\lambda),
\]

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

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

and \( \mu_{\lambda} = \mu + n \).

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

\[
p(y, \beta, \tau^2, \lambda) = P_S(\beta|\mu_{\lambda}, \beta_{\lambda}, \frac{\mu_{\lambda}}{f_{\lambda}}, M_{\lambda}) \times
\]

\[
p_{IB}(\tau^2|\frac{\nu_0 (1+\lambda)}{2}, \frac{\nu_0 + \mu_{\lambda}}{2}, \frac{\nu_0 (1+\lambda)}{f_{\lambda} + (\beta - \beta_{\lambda})' M_{\lambda} (\beta - \beta_{\lambda})}) p(y|\lambda) p(\lambda).
\]

(B.1)
Using (A.4) and (A.2), one can equivalently write

\[ p(y, \beta, \sigma^2, \lambda) = p_{SB}(\beta \mid 2\nu_0 + \mu_\lambda, \beta_\lambda, \frac{2\nu_0 + \mu_\lambda}{\nu_0 \sigma^2 (1+\lambda) + f_{\lambda}} M_\lambda) \times \]

\[ p_{IB}(\sigma^2 \mid \frac{\mu_\lambda}{2}, \frac{\nu_0}{2}, \frac{f_{\lambda}}{\nu_0 (1+\lambda)}) p(y \mid \lambda) p(\lambda). \]  

(B.2)

Note that, conditionally on \( y \) and \( \lambda \), \((\beta, \sigma^2)\) has a t-F distribution, thus the posterior density \( p(\beta, \sigma^2 \mid y) \) is a continuous mixture of t-F densities with the following mixing density:

\[ p(\lambda \mid y) \propto p(y \mid \lambda) p(\lambda) \propto \left| M_\lambda \right|^{-\frac{1}{2}} f_{\lambda}^{-1} - \frac{2\nu_0 + \mu_\lambda}{2} - \frac{2\nu_0 + \mu_\lambda}{2} (1+\lambda). \]

The first factor on the r.h.s. of (B.1) equals \( p(\beta \mid y, \lambda) \), and the second factor on the r.h.s. of (B.2) is \( p(\sigma^2 \mid y, \lambda) \). For \( \lambda=0 \), implying that \( \mu = \nu_\alpha \) in Zellner's notation, these two densities, as well as \( p(\lambda \mid y) \), 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 \( (\sigma^2, \tau, \theta) \) to \( (\sigma^2, \tau, \lambda) \), where \( \lambda = \tau^2 / \theta^2 \), is proportional to \( \lambda^{3/2} / \tau \); obviously, it is proportional to \( \tau / \lambda^{3/2} \). (Differences in notation: Zellner's \( \tau, \theta \) and \( \nu_s \) are equal to our \( z \), \( v \) and \( f_{\lambda} \), respectively; thus, Zellner's \( \lambda = \tau^2 / \theta^2 \) 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 \( p(\sigma^2, \lambda \mid y) \), one will obtain the same \( p(\sigma^2 \mid y, \lambda) \) and \( p(\lambda \mid y) \) as we do for \( \lambda=0 \). The remaining differences are due to obvious misprints; in Zellner's formula for \( c_\lambda \) after (A.9), \( \beta' A \bar{\beta} \) should be multiplied by \( \lambda \), as it is in his (A.4), and the exponent of \( (1+\lambda) \) in (A.10) should read \( -\nu_0 \), not \( \nu_0 \). Correcting these misprints one obtains from Zellner's (A.10) the same \( p(\beta \mid y, \lambda) \) as ours for \( \lambda=0 \).
References


<table>
<thead>
<tr>
<th>No.</th>
<th>Author(s)</th>
<th>Title</th>
</tr>
</thead>
<tbody>
<tr>
<td>8801</td>
<td>Th. van de Klundert and F. van der Ploeg</td>
<td>Fiscal Policy and Finite Lives in Interdependent Economies with Real and Nominal Wage Rigidities</td>
</tr>
<tr>
<td>8802</td>
<td>J.R. Magnus and B. Pesaran</td>
<td>The Bias of Forecasts from a First-order Autoregression</td>
</tr>
<tr>
<td>8804</td>
<td>F. van der Ploeg and A.J. de Zeeuw</td>
<td>Perfect Equilibrium in a Model of Competitive Arms Accumulation</td>
</tr>
<tr>
<td>8805</td>
<td>M.F.J. Steel</td>
<td>Seemingly Unrelated Regression Equation Systems under Diffuse Stochastic Prior Information: A Recursive Analytical Approach</td>
</tr>
<tr>
<td>8806</td>
<td>Th. Ten Raa and E.N. Wolff</td>
<td>Secondary Products and the Measurement of Productivity Growth</td>
</tr>
<tr>
<td>8807</td>
<td>F. van der Ploeg</td>
<td>Monetary and Fiscal Policy in Interdependent Economies with Capital Accumulation, Death and Population Growth</td>
</tr>
<tr>
<td>8901</td>
<td>Th. Ten Raa and P. Kop Jansen</td>
<td>The Choice of Model in the Construction of Input-Output Coefficients Matrices</td>
</tr>
<tr>
<td>8902</td>
<td>Th. Nijman and F. Palm</td>
<td>Generalized Least Squares Estimation of Linear Models Containing Rational Future Expectations</td>
</tr>
<tr>
<td>8903</td>
<td>A. van Soest, I. Woittiez, A. Kapteyn</td>
<td>Labour Supply, Income Taxes and Hours Restrictions in The Netherlands</td>
</tr>
<tr>
<td>8904</td>
<td>F. van der Ploeg</td>
<td>Capital Accumulation, Inflation and Long-Run Conflict in International Objectives</td>
</tr>
<tr>
<td>8905</td>
<td>Th. van de Klundert and A. van Schaik</td>
<td>Unemployment Persistence and Loss of Productive Capacity: A Keynesian Approach</td>
</tr>
<tr>
<td>8907</td>
<td>J. Osiewalski</td>
<td>Posterior Densities for Nonlinear Regression with Equicorrelated Errors</td>
</tr>
<tr>
<td>8908</td>
<td>M.F.J. Steel</td>
<td>A Bayesian Analysis of Simultaneous Equation Models by Combining Recursive Analytical and Numerical Approaches</td>
</tr>
<tr>
<td>No.</td>
<td>Author(s)</td>
<td>Title</td>
</tr>
<tr>
<td>------</td>
<td>---------------------------</td>
<td>-----------------------------------------------------------------------------------------</td>
</tr>
<tr>
<td>8909</td>
<td>F. van der Ploeg</td>
<td>Two Essays on Political Economy</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(i) The Political Economy of Overvaluation</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(ii) Election Outcomes and the Stockmarket</td>
</tr>
<tr>
<td>8910</td>
<td>R. Gradus and</td>
<td>Corporate Tax Rate Policy and Public and Private Employment</td>
</tr>
<tr>
<td></td>
<td>A. de Zeeuw</td>
<td></td>
</tr>
<tr>
<td>8911</td>
<td>A.P. Barten</td>
<td>Allais Characterisation of Preference Structures and the Structure of Demand</td>
</tr>
<tr>
<td>8912</td>
<td>K. Kamiya and</td>
<td>Simplicial Algorithm to Find Zero Points of a Function with Special Structure on a</td>
</tr>
<tr>
<td></td>
<td>A.J.J. Talman</td>
<td>Simplotope</td>
</tr>
<tr>
<td>8913</td>
<td>G. van der Laan and</td>
<td>Price Rigidities and Rationing</td>
</tr>
<tr>
<td></td>
<td>A.J.J. Talman</td>
<td></td>
</tr>
<tr>
<td>8914</td>
<td>J. Osiewalski and</td>
<td>A Bayesian Analysis of Exogeneity in Models</td>
</tr>
<tr>
<td></td>
<td>M.F.J. Steel</td>
<td>Pooling Time-Series and Cross-Section Data</td>
</tr>
<tr>
<td>8915</td>
<td>R.P. Gilles, P.H. Ruys</td>
<td>On the Existence of Networks in Relational Models</td>
</tr>
<tr>
<td></td>
<td>and J. Shou</td>
<td></td>
</tr>
<tr>
<td>8916</td>
<td>A. Kapteyn, P. Kooreman</td>
<td>Quantity Rationing and Concavity in a Flexible Household Labor Supply Model</td>
</tr>
<tr>
<td></td>
<td>and A. van Soest</td>
<td></td>
</tr>
<tr>
<td>8917</td>
<td>F. Canova</td>
<td>Seasonalities in Foreign Exchange Markets</td>
</tr>
<tr>
<td>8918</td>
<td>F. van der Ploeg</td>
<td>Monetary Disinflation, Fiscal Expansion and the Current Account in an Interdependent</td>
</tr>
<tr>
<td></td>
<td></td>
<td>World</td>
</tr>
<tr>
<td>8919</td>
<td>W. Bossert and</td>
<td>On the Uniqueness of Cardinally Interpreted Utility Functions</td>
</tr>
<tr>
<td></td>
<td>F. Stehling</td>
<td></td>
</tr>
<tr>
<td>8920</td>
<td>F. van der Ploeg</td>
<td>Monetary Interdependence under Alternative Exchange-Rate Regimes</td>
</tr>
<tr>
<td>8921</td>
<td>D. Canning</td>
<td>Bottlenecks and Persistent Unemployment: Why Do Booms End?</td>
</tr>
<tr>
<td>8922</td>
<td>C. Fershtman and</td>
<td>Price Cycles and Booms: Dynamic Search Equilibrium</td>
</tr>
<tr>
<td></td>
<td>A. Fishman</td>
<td></td>
</tr>
<tr>
<td>8923</td>
<td>M.B. Canzoneri and</td>
<td>Is the European Community an Optimal Currency Area? Optimal Tax Smoothing versus the Cost</td>
</tr>
<tr>
<td></td>
<td>C.A. Rogers</td>
<td>of Multiple Currencies</td>
</tr>
<tr>
<td>8924</td>
<td>F. Groot, C. Withagen</td>
<td>Theory of Natural Exhaustible Resources: The Cartel-Versus-Fringe Model Reconsidered</td>
</tr>
<tr>
<td></td>
<td>and A. de Zeeuw</td>
<td></td>
</tr>
<tr>
<td>No.</td>
<td>Author(s)</td>
<td>Title</td>
</tr>
<tr>
<td>------</td>
<td>---------------------------------</td>
<td>----------------------------------------------------------------------</td>
</tr>
<tr>
<td>8925</td>
<td>O.P. Attanasio and G. Weber</td>
<td>Consumption, Productivity Growth and the Interest Rate</td>
</tr>
<tr>
<td>8926</td>
<td>N. Rankin</td>
<td>Monetary and Fiscal Policy in a 'Hartian' Model of Imperfect Competition</td>
</tr>
<tr>
<td>8927</td>
<td>Th. van de Klundert</td>
<td>Reducing External Debt in a World with Imperfect Asset and Imperfect Commodity Substitution</td>
</tr>
<tr>
<td>8928</td>
<td>C. Dang</td>
<td>The D_2-Triangulation of R^n for Simplicial Algorithms for Computing Solutions of Nonlinear Equations</td>
</tr>
<tr>
<td>8929</td>
<td>M.F.J. Steel and J.F. Richard</td>
<td>Bayesian Multivariate Exogeneity Analysis: An Application to a UK Money Demand Equation</td>
</tr>
<tr>
<td>8930</td>
<td>F. van der Ploeg</td>
<td>Fiscal Aspects of Monetary Integration in Europe</td>
</tr>
<tr>
<td>8931</td>
<td>H.A. Keuzenkamp</td>
<td>The Prehistory of Rational Expectations</td>
</tr>
<tr>
<td>8932</td>
<td>E. van Damme, R. Selten and E. Winter</td>
<td>Alternating Bid Bargaining with a Smallest Money Unit</td>
</tr>
<tr>
<td>8933</td>
<td>H. Carlsson and E. van Damme</td>
<td>Global Payoff Uncertainty and Risk Dominance</td>
</tr>
<tr>
<td>8934</td>
<td>H. Huizinga</td>
<td>National Tax Policies towards Product-Innovating Multinational Enterprises</td>
</tr>
<tr>
<td>8935</td>
<td>C. Dang and D. Talman</td>
<td>A New Triangulation of the Unit Simplex for Computing Economic Equilibria</td>
</tr>
<tr>
<td>8936</td>
<td>Th. Nijman and M. Verbeek</td>
<td>The Nonresponse Bias in the Analysis of the Determinants of Total Annual Expenditures of Households Based on Panel Data</td>
</tr>
<tr>
<td>8937</td>
<td>A.P. Barten</td>
<td>The Estimation of Mixed Demand Systems</td>
</tr>
<tr>
<td>8938</td>
<td>G. Marini</td>
<td>Monetary Shocks and the Nominal Interest Rate</td>
</tr>
<tr>
<td>8939</td>
<td>W. Güth and E. van Damme</td>
<td>Equilibrium Selection in the Spence Signaling Game</td>
</tr>
<tr>
<td>8940</td>
<td>G. Marini and P. Scaramozzino</td>
<td>Monopolistic Competition, Expected Inflation and Contract Length</td>
</tr>
<tr>
<td>8941</td>
<td>J.K. Dagsvik</td>
<td>The Generalized Extreme Value Random Utility Model for Continuous Choice</td>
</tr>
<tr>
<td>No.</td>
<td>Author(s)</td>
<td>Title</td>
</tr>
<tr>
<td>-----</td>
<td>-----------</td>
<td>-------</td>
</tr>
<tr>
<td>8942</td>
<td>M.F.J. Steel</td>
<td>Weak Exogenity in Misspecified Sequential Models</td>
</tr>
<tr>
<td>8943</td>
<td>A. Roell</td>
<td>Dual Capacity Trading and the Quality of the Market</td>
</tr>
<tr>
<td>8944</td>
<td>C. Hsiao</td>
<td>Identification and Estimation of Dichotomous Latent Variables Models Using Panel Data</td>
</tr>
<tr>
<td>8945</td>
<td>R.P. Gilles</td>
<td>Equilibrium in a Pure Exchange Economy with an Arbitrary Communication Structure</td>
</tr>
<tr>
<td>8946</td>
<td>W.B. MacLeod and J.M. Malcomson</td>
<td>Efficient Specific Investments, Incomplete Contracts, and the Role of Market Alternatives</td>
</tr>
<tr>
<td>8947</td>
<td>A. van Soest and A. Kapteyn</td>
<td>The Impact of Minimum Wage Regulations on Employment and the Wage Rate Distribution</td>
</tr>
<tr>
<td>8948</td>
<td>P. Kooreman and B. Melenberg</td>
<td>Maximum Score Estimation in the Ordered Response Model</td>
</tr>
<tr>
<td>8949</td>
<td>C. Dang</td>
<td>The $D_2$-Triangulation for Simplicial Deformation Algorithms for Computing Solutions of Nonlinear Equations</td>
</tr>
<tr>
<td>8950</td>
<td>M. Cripps</td>
<td>Dealer Behaviour and Price Volatility in Asset Markets</td>
</tr>
<tr>
<td>8951</td>
<td>T. Wansbeek and A. Kapteyn</td>
<td>Simple Estimators for Dynamic Panel Data Models with Errors in Variables</td>
</tr>
<tr>
<td>8952</td>
<td>Y. Dai, G. van der Laan, D. Talman and Y. Yamamoto</td>
<td>A Simplicial Algorithm for the Nonlinear Stationary Point Problem on an Unbounded Polyhedron</td>
</tr>
<tr>
<td>8953</td>
<td>F. van der Ploeg</td>
<td>Risk Aversion, Intertemporal Substitution and Consumption: The CARA-LQ Problem</td>
</tr>
<tr>
<td>8954</td>
<td>A. Kapteyn, S. van de Geer, H. van de Stadt and T. Wansbeek</td>
<td>Interdependent Preferences: An Econometric Analysis</td>
</tr>
<tr>
<td>8955</td>
<td>L. Zou</td>
<td>Ownership Structure and Efficiency: An Incentive Mechanism Approach</td>
</tr>
<tr>
<td>8957</td>
<td>E. van Damme</td>
<td>Signaling and Forward Induction in a Market Entry Context</td>
</tr>
<tr>
<td>No.</td>
<td>Author(s)</td>
<td>Title</td>
</tr>
<tr>
<td>-----</td>
<td>---------------------------------------------------------------------------</td>
<td>-------------------------------------------------------------------------------------------</td>
</tr>
<tr>
<td>9001</td>
<td>A. van Soest, P. Kooreman and A. Kapteyn</td>
<td>Coherency and Regularity of Demand Systems with Equality and Inequality Constraints</td>
</tr>
<tr>
<td>9002</td>
<td>J.R. Magnus and B. Pesaran</td>
<td>Forecasting, Misspecification and Unit Roots: The Case of AR(1) Versus ARMA(1,1)</td>
</tr>
<tr>
<td>9003</td>
<td>J. Driffill and C. Schultz</td>
<td>Wage Setting and Stabilization Policy in a Game with Renegotiation</td>
</tr>
<tr>
<td>9004</td>
<td>M. McAleer, M.H. Pesaran and A. Bera</td>
<td>Alternative Approaches to Testing Non-Nested Models with Autocorrelated Disturbances: An Application to Models of U.S. Unemployment</td>
</tr>
<tr>
<td>9005</td>
<td>Th. ten Raa and M.F.J. Steel</td>
<td>A Stochastic Analysis of an Input-Output Model: Comment</td>
</tr>
<tr>
<td>9006</td>
<td>M. McAleer and C.R. McKenzie</td>
<td>Keynesian and New Classical Models of Unemployment Revisited</td>
</tr>
<tr>
<td>9007</td>
<td>J. Osiewalski and M.F.J. Steel</td>
<td>Semi-Conjugate Prior Densities in Multivariate t Regression Models</td>
</tr>
</tbody>
</table>