

Center
for
Economic Research

Discussion paper

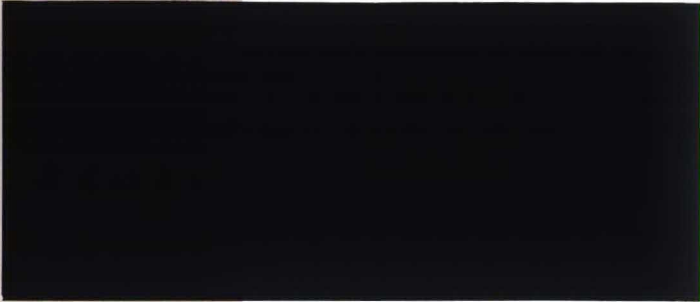
CRM
CBM
R

88414

1990

1990-32

32



No. 9032

**ROBUST BAYESIAN INFERENCE IN
ELLIPTICAL REGRESSION MODELS**

by Jacek Osiewalski
and Mark Steel

R46

518.922 1

May 1990

ISSN 0924-7815

ROBUST BAYESIAN INFERENCE
IN ELLIPTICAL REGRESSION MODELS*

Jacek Osiewalski
Academy of Economics
Kraków, Poland

Mark Steel
Tilburg University
Tilburg, The Netherlands

ABSTRACT

Broadening the stochastic assumptions on the error terms of regression models was prompted by the analysis of linear multivariate t models in Zellner (1976). We consider a possibly non-linear regression model under *any* multivariate elliptical data density, and examine Bayesian posterior and predictive results. The latter are shown to be robust with respect to the specific choice of a sampling density within this elliptical class. In particular, sufficient conditions for such model robustness are that we single out a precision factor τ^2 that does not influence the way the density changes over ellipsoids, and that we specify an improper prior density on τ^2 . Apart from the posterior distribution of this nuisance parameter τ^2 , the entire analysis will then be completely unaffected by departures from Normality. Similar results hold in finite mixtures of such elliptical densities, which can be used to approximate more general data processes.

* This paper was written while the first author was a visiting fellow at the Center for Operations Research and Econometrics (CORE), Louvain-la-Neuve, Belgium, and the second author held a research fellowship of the Royal Netherlands Academy of Arts and Sciences (KNAW). Useful discussions with Jean-Pierre Florens, Michel Mouchart and Jean-Marie Rolin are gratefully acknowledged, though the usual disclaimer applies.

1. INTRODUCTION

The Bayesian analysis of regression models with dependent non-Normal error terms has received considerable attention, especially since the seminal paper of Zellner (1976), who considered linear multivariate Student t regression models. This assumption was extended to scale mixtures of Normal distributions in Jammalamadaka *et al.* (1987) and in Chib *et al.* (1988) whereas Osiewalski (1990) and Chib *et al.* (1990) generalize, in addition, to nonlinear models. Here we shall examine a further generalization to the *entire class of multivariate elliptical or ellipsoidal densities*, as it was defined in e.g. Kelker (1970), Cambanis *et al.* (1981) or Dickey and Chen (1985).

In particular, we find that any multivariate elliptical regression model, combined with an improper reference prior on the “nuisance” scalar precision parameter τ^2 , will lead to exactly the same posterior and predictive analyses as in the Normal case. Thus, in this sense, our inference is fully *robust* with respect to changes in the specification of the sampling process within this wide class of elliptical densities. Remark that this property differs from robustness against extreme observations, as used e.g. in Ramsay and Novick (1980), who defined a concept of “ L robustness”. The latter relates to the sensitivity of the likelihood to the *data*, and is based on the influence function. Instead, we arrive at robustness of posterior and predictive results with respect to the *sampling model* itself, within a broad class of models that includes many L robust cases, like e.g. Student or Cauchy models. Thus, we focus on “model robustness” [see Berger (1985, p. 248)], and in particular, on what Box and Tiao (1973, p. 152) call “inference robustness”.

These robustness results are derived for multivariate elliptical distributions, and do not generally hold under independent non-Normal error terms. If we assume that the errors are independently *and* identically distributed according to some elliptical process other than the Normal, no such robustness occurs. The results in Box and Tiao (1973, Ch. 3), West (1984) and Bagchi and Guttman (1988) provide some evidence in this respect. However, if we start from a multivariate elliptical framework, where independence can only be accommodated under Normality [see Kelker (1970, Lemma 5)], the usual improper reference prior on τ^2 does the trick. Only posterior results for τ^2 are affected by departures from Normality, as in Zellner (1976). Given the nuisance character of this scale factor, however, these results are not explicitly stated here. Predictive inference and posterior infer-

ence on the parameter θ , which defines the location and shape of the ellipsoids, can then be conducted exactly as in the usual Normal case. The remaining parameter ν only serves to index the way the density function changes over ellipsoids, and the sample will only update the marginal prior on ν through its prior dependence on θ .

A finite mixture of elliptical data densities is then considered for cases in which we want to avoid a symmetry assumption. The mixing will be preserved in posterior and predictive analyses, which allows broadening the class of data densities, without really affecting the complexity of the ensuing analysis. It is just like mixing Normal distributions defined over different ellipsoids, where the mixing parameter λ will be revised by the sample, albeit in a rather moderate way.

Section 2 introduces the Bayesian model, on the basis of which we derive posterior and predictive results in Sections 3 and 4, respectively. The finite mixtures of data densities are examined in Section 5, whereas a final section summarizes some conclusions.

2. THE BAYESIAN MODEL

2.1. The Elliptical Sampling Model

A general form of elliptical, also known as ellipsoidal, distributions will be assumed for the sampling process. The observation vector y has an n -variate continuous elliptical distribution, given a set of exogenous variables X and a sufficient parameterization, say ω , if and only if its data density is

$$p(y | X, \omega) = |\tilde{V}(X, \bar{\eta})|^{-\frac{1}{2}} g_{n,\omega} [(y - h(X, \beta))' (\tilde{V}(X, \bar{\eta}))^{-1} (y - h(X, \beta))]. \quad (2.1)$$

In (2.1) $g_{n,\omega}(\cdot)$ is a nonnegative function, which for any n and ω has to fulfil the condition

$$\int_0^\infty u^{\frac{n}{2}-1} g_{n,\omega}(u) du = \Gamma\left(\frac{n}{2}\right) \pi^{-\frac{n}{2}}. \quad (2.2)$$

It can be shown [see Cambanis *et al.* (1981), Dickey and Chen (1985), Kelker (1970)] that (2.2) is both necessary and sufficient to make (2.1) a proper, normalized density function.

The location vector in (2.1) is the, possibly nonlinear, but known, function $h(X, \beta)$, and the scale matrix is $\tilde{V}(X, \bar{\eta})$, where \tilde{V} is positive definite symmetric

(PDS) and a known matrix function of X and $\bar{\eta}$. Therefore, $\beta \in B \subseteq \mathbb{R}^k$ and $\bar{\eta} \in \bar{H} \subseteq \mathbb{R}^q$ serve to define the isodensity ellipsoids of y . The labelling function $g_{n,\omega}$ that determines the density value for each of these ellipsoids [see e.g. Leamer (1978, p. 150)] is indexed by n and ω , which may contain parameters other than β and $\bar{\eta}$, introduced specifically for the purpose of describing $g_{n,\omega}$. Let us call these parameters $\nu \in N \subseteq \mathbb{R}^\ell$. A well-known example of this ν is found in the multivariate Student t distribution, where $\nu \in \mathbb{R}_+$ and

$$g_{n,\omega}(\cdot) = g_{n,\nu}(\cdot) = \frac{\Gamma(\frac{n+\nu}{2})}{\Gamma(\frac{n}{2})} (\nu\pi)^{-\frac{\nu}{2}} \left(1 + \frac{\cdot}{\nu}\right)^{-\frac{n+\nu}{2}}.$$

Indeed, from (2.1) we will then obtain a Student t data density with ν degrees of freedom, location vector $h(X, \beta)$ and precision matrix $\tilde{V}(X, \bar{\eta})^{-1}$, denoted by

$$p(y | X, \omega) = f_S^n(y | \nu, h(X, \beta), \tilde{V}(X, \bar{\eta})^{-1}). \quad (2.3)$$

A generalization of (2.3), where the dimension of ν is extended, can be found in Dickey and Chen (1985, p. 173). However, in some cases none of the parameters in ω will appear in $g_{n,\omega}(\cdot)$, which will then only depend on n , the dimension of y . If, in particular, we choose

$$g_{n,\omega}(\cdot) = g_n(\cdot) = (2\pi)^{-\frac{\nu}{2}} \exp\left(-\frac{\cdot}{2}\right),$$

our data density in (2.1) will be of the Normal form with mean $h(X, \beta)$ and covariance matrix $\tilde{V}(X, \bar{\eta})$:

$$p(y | X, \omega) = f_N^n(y | h(X, \beta), \tilde{V}(X, \bar{\eta})), \quad (2.4)$$

where $(\beta, \bar{\eta})$ is now a sufficient parameterization. Neither of these well-known special cases correspond to an indexing of $g_{n,\omega}(\cdot)$ by the parameters appearing in the definition of the ellipsoid, namely $(\beta, \bar{\eta})$. Starting from the more restricted class of elliptical distributions that can be expressed as scale mixtures of Normals [see e.g. Kelker (1970)], such a dependence can easily be introduced through the density of the mixing parameter. Examples appeared in Osiewalski (1990), Osiewalski and Steel (1990) and Chib *et al.* (1990), who considered some specific densities of the general form

$$\begin{aligned} p(y | X, \omega) &= \int_0^\infty f_N^n(y | h(X, \beta), \frac{\nu}{z} \tilde{V}(X, \bar{\eta})) \\ &\quad f_G\left(z \mid \frac{\nu + b_1(\beta, \bar{\eta})}{2}, \frac{b_2(\beta, \bar{\eta})}{2\nu}\right) dz \quad (2.5) \\ &= f_S^n(y | \nu + b_1(\beta, \bar{\eta}), h(X, \beta), \frac{\nu + b_1(\beta, \bar{\eta})}{b_2(\beta, \bar{\eta})} \tilde{V}(X, \bar{\eta})^{-1}), \end{aligned}$$

where $b_i(\beta, \bar{\eta})$, $i = 1, 2$, are positive functions of $(\beta, \bar{\eta})$ and $f_G(\cdot | \cdot, \cdot)$ denotes a gamma density function. In the (more general) framework of (2.1), the density in (2.5) can be reproduced by choosing

$$g_{n,\omega}(\cdot) = \frac{\Gamma\left(\frac{n + \nu + b_1(\beta, \bar{\eta})}{2}\right)}{\Gamma\left(\frac{\nu + b_1(\beta, \bar{\eta})}{2}\right) \pi^{\frac{n}{2}}} b_2(\beta, \bar{\eta})^{-\frac{n}{2}} \left(1 + \frac{\cdot}{b_2(\beta, \bar{\eta})}\right)^{-\frac{n + \nu + b_1(\beta, \bar{\eta})}{2}},$$

which clearly depends on $(\beta, \bar{\eta})$ as well.

The explanatory variables in X have a sampling distribution whose sufficient parameterization is denoted by λ . If we assume the joint prior on ω and λ is a product of $p(\omega)$ and $p(\lambda)$, both σ -finite, we can ignore the process of X for the purpose of conducting inference with (2.1). These assumptions, in fact, amount to operating a Bayesian cut [see e.g. Florens and Mouchart (1985) and Florens *et al.* (1990)].

We shall now introduce two restrictions on the class of models in (2.1), since this will open an avenue to drastic simplifications, using the property in (2.2), as will become clear in the sequel. Firstly, we restrict ourselves to those scale matrices $\tilde{V}(X, \bar{\eta})$ that can be written as

$$\tilde{V}(X, \bar{\eta}) = \frac{1}{\tau^2} V(X, \eta), \quad (2.6)$$

where $\tau^2 \in \mathbb{R}_+$ is a scalar precision parameter, implicitly reparameterizing $\bar{\eta}$ as (τ^2, η) . For notational convenience, we now define $\theta = (\beta, \eta)$ which contains all the information about the location and shape of the ellipsoids.

Secondly, we shall assume that this scalar precision parameter τ^2 does not index the function $g_{n,\omega}$, i.e.

$$g_{n,\omega}(\cdot) = g_{n,\theta,\nu}(\cdot). \quad (2.7)$$

This implies the interpretation for τ^2 can only be linked to the ellipsoids, and should, thus, be equivalent for all models in our class. Therefore, we can, at this level, before actually choosing a particular model, consider assigning a prior density to τ^2 , as its meaning does not vary over the elliptical class in (2.1), restricted by (2.6) and (2.7).

Our framework now excludes densities like (2.5), unless the functions $b_i(\beta, \bar{\eta}) = b_i(\beta, \tau^2, \eta)$ are constant in τ^2 , but cases where (2.7) is violated seem rather artificial and somewhat unlikely to occur in actual practice.

2.2. Prior Densities

We now face the task of completing the Bayesian model by assigning a prior distribution to $\omega = (\theta, \tau^2, \nu)$. Both θ and ν (if it appears) will typically be parameters of interest, and we shall leave the specification of their prior density completely free at this stage. We shall see in Section 3 that if we specify the ("usual") improper prior structure

$$p(\theta, \tau^2, \nu) = \frac{c}{\tau^2} p(\theta, \nu), \quad (2.8)$$

where c is a positive constant and $p(\theta, \nu)$ is functionally independent of τ^2 , the analysis will simplify greatly. More in particular, provided $g_{n,\omega}(\cdot)$ is *not* indexed by τ^2 , its actual form becomes completely irrelevant, so that both posterior and predictive analyses are *fully robust* with respect to any departures from Normality in the wide class of remaining elliptical densities.

3. POSTERIOR INFERENCE

Combining the general class of elliptical data densities in (2.1), restricted only by (2.6) and (2.7), with the improper prior family in (2.8), we obtain the joint density

$$p(y, \omega | X) = c p(\theta, \nu) (\tau^2)^{\frac{n}{2}-1} |V(X, \eta)|^{-\frac{1}{2}} g_{n,\theta,\nu}[\tau^2 d(y, X, \theta)], \quad (3.1)$$

where we have defined

$$d(y, X, \theta) = (y - h(X, \theta))' V(X, \eta)^{-1} (y - h(X, \theta)).$$

Let us now consider the transformation from (y, θ, τ^2, ν) to (y, θ, r^2, ν) , where

$$r^2 = \tau^2 d(y, X, \theta), \quad (3.2)$$

leading to

$$p(y, \theta, r^2, \nu | X) = c p(\theta, \nu) |V(X, \eta)|^{-\frac{1}{2}} d(y, X, \theta)^{-\frac{n}{2}} (\tau^2)^{\frac{n}{2}-1} g_{n,\theta,\nu}(\tau^2). \quad (3.3)$$

The function $g_{n,\theta,\nu}(\cdot)$ is not affected by the transformation in (3.2), since it does not involve τ^2 , so that property (2.2) can directly be applied to integrate out r^2 in (3.3). This leaves us with

$$p(y, \theta, \nu | X) = c \Gamma\left(\frac{n}{2}\right) \pi^{-\frac{n}{2}} p(\theta, \nu) |V(X, \eta)|^{-\frac{1}{2}} d(y, X, \theta)^{-\frac{n}{2}}, \quad (3.4)$$

which no longer depends on the form of $g_{n,\theta,\nu}(\cdot)$. The joint (improper) density of our parameters of interest and y is thus *completely robust* with respect to *any* departures from Normality in the class of elliptical data densities (2.1), restricted by (2.6) and (2.7), when τ^2 is treated by assuming the improper prior (2.8).

Let us now assume that the prior $p(\theta, \nu)$ is integrable in ν over N , which makes (3.4) integrable in ν , and also that the resulting density, $p(y, \theta | X) = \int_N p(y, \theta, \nu | X) d\nu$, is integrable in θ over $\Theta \subseteq B \times H$. We are then sure that the posterior of (θ, ν) is well defined as

$$p(\theta, \nu | y, X) \propto p(\theta, \nu) |V(X, \eta)|^{-\frac{1}{2}} d(y, X, \theta)^{-\frac{n}{2}}, \quad (3.5)$$

from which we can easily derive the posterior for the location and shape parameters θ .

Theorem 1. For *any* elliptical data density (2.1), fulfilling (2.6) and (2.7), and under an improper prior (2.8), which is integrable in ν , we obtain the *same* posterior of θ :

$$p(\theta | y, X) \propto p(\theta) |V(X, \eta)|^{-\frac{1}{2}} d(y, X, \theta)^{-\frac{n}{2}}, \quad (3.6)$$

where $p(\theta) = \int_N p(\theta, \nu) d\nu$ and we have assumed that (3.6) is integrable in θ over Θ . ■

Of course, (3.6) is exactly the posterior one obtains for the Normal data density (2.4), and may look even more familiar if we consider the simple linear case:

Corollary 1: In the special linear case of Theorem 1 where $h(X, \beta) = X\beta$ and $\Theta = \mathbb{R}^k \times H$, the posterior densities of θ are given by

$$p(\beta | \eta, y, X) = K(\eta)^{-1} p(\beta, \eta) f_S^k(\beta | n - k, \hat{\beta}, s^{-2} X'V(X, \eta)^{-1} X) \quad (3.7)$$

and

$$p(\eta | y, X) \propto K(\eta) |V(X, \eta)|^{-\frac{1}{2}} |X'V(X, \eta)^{-1} X|^{-\frac{1}{2}} (s^2)^{-\frac{1}{2}(n-k)}, \quad (3.8)$$

where

$$\hat{\beta} = (X'V(X, \eta)^{-1}X)^{-1} X'V(X, \eta)^{-1}y$$

$$s^2 = \frac{1}{n-k} (y - X\hat{\beta})' V(X, \eta)^{-1} (y - X\hat{\beta})$$

and $K(\eta)$, the inverse of the normalizing constant of (3.7), absorbs the prior information on η . ■

Implicitly, we have also made the assumption that X is of full column rank in Corollary 1, which implies $n \geq k$ in this linear case. If we specify a uniform prior on β , i.e. $p(\beta, \eta) \propto p(\eta)$, we simply have a Student t conditional posterior of β , which is proper if $n > k$. Moments of (3.7) then exist up to (not including) order $n - k$. Adding some prior information will typically lead to the existence of higher order moments. In particular, if $p(\beta, \eta)$ contains a Student t kernel for β with ν_0 degrees of freedom, the conditional posterior in (3.7) will be of a 2-0 poly- t form [see Drèze (1977) and Richard and Tompa (1980)], allowing for posterior moments up to order $\nu_0 + n - k$.

The invariance results obtained here are a direct consequence of the fact that, after integrating out τ^2 under (2.8), we have

$$p(y | X, \theta, \nu) = p(y | X, \theta) \propto d(y, X, \theta)^{-\frac{n}{2}}, \quad (3.9)$$

irrespective of the form of $g_{n, \theta, \nu}(\cdot)$. Therefore, we address a particular case of Hill (1969), who proposed specifying a spherical model without considering a scale parameter. In the framework of (2.1) he does *not* impose (2.6) and introduces sphericity by assuming $\tilde{V}(X, \tilde{\eta}) = I_n$ directly. The more “traditional” approach, in e.g. Zellner (1976), Jammalamadaka *et al.* (1987), Chib *et al.* (1988) and Osiewalski (1990), implicitly starts from the deeper level of parameterization used here and amounts to assuming (2.6). In that case sphericity is induced by taking $V(X, \eta) = I_n$. Hill’s (1969) specification of general spherical errors is thus made at a level of parameterization comparable to the one in (3.9). By not imposing (2.6), Hill’s approach is slightly more general, but at the cost of not obtaining the robustness that follows from (3.9). Nevertheless, Hill (1969) does introduce a scale factor in his discussion of Normality. At that level, our results imply that it is *not* the Normality assumption but the use of Jeffreys’ prior on this scale factor [as in (2.8)] that accounts for finding the “usual” posterior results. Therefore, provided one is willing to accept (2.6), Normality does not seem to be quite as restrictive as suggested by Hill.

From the joint posterior of θ and ν in (3.5), it is already obvious that ν only appears through the joint prior $p(\theta, \nu)$. This means that *given* θ the sample contains no information regarding ν , so that conditionally upon θ ν is not updated through the observations. Thus, under the conditions of Theorem 1, we always have that $p(\nu | \theta, y, X) = p(\nu | \theta)$, where the latter density is well defined since $p(\theta, \nu)$ is assumed to be integrable in ν . The *marginal* prior on ν , however, will generally be updated [see also Drèze and Richard (1983, p.522)], since it is given by the integral of (3.5) in θ over Θ , which can be written as

$$p(\nu | y, X) \propto \int_{\Theta} p(\nu | \theta) p(\theta | y, X) d\theta$$

where $p(\theta | y, X)$ was defined in (3.6). Thus, if $p(\nu | \theta)$ does not depend on θ (i.e. independence in probability if $p(\theta)$ is proper and functional independence if it is not) the sample cannot revise the marginal prior of ν either and we state:

Theorem 2: Under the conditions of Theorem 1, the prior structure for (θ, ν)

$$p(\theta, \nu) = p(\theta) p(\nu) \tag{3.10}$$

will prevent updating of the marginal prior information on ν , i.e.

$$p(\nu | y, X) = p(\nu). \tag{3.11}$$

■

The lack of dependence in (3.10), which is taken to be integrable in ν , will, for any proper elliptical sampling model (2.1) fulfilling (2.6) and (2.7), lead to posterior independence of θ and ν , provided we express our prior ignorance about τ^2 by the class of improper densities in (2.8), and if the joint posterior exists, which is assured if (3.6) is integrable in θ . This can be seen directly from (3.5), and, given the fact that the sample can only update ν through θ , this posterior independence will make sure that our marginal opinions regarding ν will not be revised through the observations. Chib *et al.* (1990) analyse the particular subclass of (2.1) where the elliptical densities can be described as scale mixtures of Normals. A prominent member of this subclass is the Student *t* model in (2.3), in which case Theorem 2 exactly reduces to their Corollary 4, stating a set of sufficient conditions for the impossibility to update the prior of the degrees-of-freedom parameter.

4. PREDICTIVE ANALYSIS

Alternatively, we can focus on the predictive properties of Bayesian models involving elliptical data densities as in (2.1) and improper priors as in (2.8), maintaining also (2.6) and (2.7).

For this purpose, we partition the n dimensional vector y as follows

$$y = \begin{pmatrix} y_{(1)} \\ y_{(2)} \end{pmatrix},$$

where $y_{(1)} \in \mathbb{R}^{n_1}$ and $n_1 < n$, and we are interested in forecasting $y_{(2)}$, given $y_{(1)}$ and X . Conformably, we partition

$$h(X, \beta) = \begin{bmatrix} h_{(1)}(X, \beta) \\ h_{(2)}(X, \beta) \end{bmatrix} \equiv \begin{pmatrix} h_{(1)} \\ h_{(2)} \end{pmatrix}$$

and

$$V(X, \eta) = \begin{bmatrix} V_{11}(X, \eta) & V_{12}(X, \eta) \\ V_{21}(X, \eta) & V_{22}(X, \eta) \end{bmatrix} \equiv \begin{pmatrix} V_{11} & V_{12} \\ V_{21} & V_{22} \end{pmatrix},$$

where the defining equalities are just used to economize on notation. From (3.4) it is immediately clear that the form of $g_{n, \theta, \nu}(\cdot)$ will not affect the predictive analysis either, and we obtain directly

$$p(y_{(1)}, y_{(2)}, \theta, \nu | X) = c \Gamma\left(\frac{n_1}{2}\right) \pi^{-\frac{n_1}{2}} p(\theta, \nu) |V_{11}|^{-\frac{1}{2}} a(y_{(1)}, X, \theta)^{-\frac{n_1}{2}} f_S^{n_2}(y_{(2)} | n_1, h_{(2)} + V_{21}V_{11}^{-1}(y_{(1)} - h_{(1)}), \frac{n_1}{a(y_{(1)}, X, \theta)} V_{22}^{-1}), \quad (4.1)$$

with $a(y_{(1)}, X, \theta) = (y_{(1)} - h_{(1)})' V_{11}^{-1}(y_{(1)} - h_{(1)})$ and $V_{22.1} = V_{22} - V_{21}V_{11}^{-1}V_{12}$. Given our assumption of integrability of the joint prior in ν , it is trivial to integrate it out, as in Section 3. The posterior of θ given the first subsample $y_{(1)}$ will be of exactly the same form as (3.6) in Theorem 1, but with $y_{(1)}$ instead of y throughout:

$$p(\theta | y_{(1)}, X) \propto p(\theta) |V_{11}|^{-\frac{1}{2}} a(y_{(1)}, X, \theta)^{-\frac{n_1}{2}}, \quad (4.2)$$

provided (4.2) is integrable in θ over Θ . The predictive density thus becomes the Student density in (4.1) of $y_{(2)}$, given $y_{(1)}$, X and θ , weighted by this posterior on the basis of $y_{(1)}$:

$$p(y_{(2)} | y_{(1)}, X) = \int_{\Theta} p(y_{(2)} | y_{(1)}, X, \theta) p(\theta | y_{(1)}, X) d\theta. \quad (4.3)$$

As was to be expected from (3.4), the general elliptical character of the data density does not induce any difference in our predictive analysis with respect to the Normal framework. Remark that integrating out τ^2 in (3.3), under the prior (2.8), always leads to a density of y , given θ and X , proportional to $d(y, X, \theta)^{-\frac{n}{2}}$, as was stressed in Section 3. This, of course, implies the Student density of $y_{(2)}$, given $y_{(1)}$, θ and X , but this Student t form will generally be lost when we integrate out θ in the predictive density as in (4.3).

A predictive analysis on the basis of (4.3) can be called for when e.g. the observations on $y_{(2)}$ are missing, whereas both $y_{(1)}$ and the entire X matrix are observed. However, in actual practice, it is often the case that only a submatrix of X , say X_1 , is jointly observed with $y_{(1)}$, so that the posterior information available for forecasting is only based on $y_{(1)}$ and X_1 . We then set out to predict $y_{(2)}$, given the observed $(y_{(1)}, X_1)$ and a set of exogenously given values for the remaining part of X , say X_2 . Given our maintained assumption of independence between X and ω , it is sufficient to assume

$$h_{(1)} = h_{(1)}(X_1, \beta) \quad \text{and} \quad V_{11} = V_{11}(X_1, \eta) \quad (4.4)$$

in order to have posterior independence between θ and X_2 .

Theorem 3: Under the conditions of Theorem 1 and (4.4) any elliptical sampling model (2.1) will allow conditional forecasting based on the predictive density

$$p(y_{(2)} | y_{(1)}, X) = \int_{\Theta} f_S^{n_2}(y_{(2)} | n_1, h_{(2)} + V_{21}V_{11}^{-1}(y_{(1)} - h_{(1)}), \quad (4.5)$$

$$\frac{n_1}{a(y_{(1)}, X_1, \theta)} V_{22.1}^{-1} p(\theta | y_{(1)}, X_1) d\theta.$$

and $p(\theta | y_{(1)}, X_1)$ is obtained from (4.2) but now with (4.4) holding. \blacksquare

The improper prior on τ^2 in (2.8), conditions (2.6)-(2.7) and the existence of the posterior thus lead to perfect predictive robustness which can be used in current practice under assumption (4.4).

In the simplest linear case with a uniform prior on $\theta = \beta$, we can write:

Corollary 2: If $h(X, \beta) = X\beta$, $V(X, \eta) = V$ is assumed known and $\Theta = \mathbb{R}^k$, then under a uniform prior on $\theta = \beta$ the predictive (4.5) in Theorem 3 reduces to the Student density

$$p(y_{(2)} | y_{(1)}, X) = f_S^{n_2}(y_{(2)} | n_1 - k, X_2\hat{\beta}_1 + V_{21}V_{11}^{-1}(y_{(1)} - X_1\hat{\beta}_1), s_1^{-2}W^{-1}) \quad (4.6)$$

with

$$\begin{aligned}\hat{\beta}_1 &= (X_1' V_{11}^{-1} X_1)^{-1} X_1' V_{11}^{-1} y_{(1)} \\ s_1^2 &= \frac{1}{n_1 - k} (y_{(1)} - X_1 \hat{\beta}_1)' V_{11}^{-1} (y_{(1)} - X_1 \hat{\beta}_1) \\ W &= (X_2 - V_{21} V_{11}^{-1} X_1) (X_1' V_{11}^{-1} X_1)^{-1} (X_2 - V_{21} V_{11}^{-1} X_1)' + V_{22.1}\end{aligned}$$

and $V_{22.1}$ defined as in (4.1). ■

A uniform prior of β has to be used for obtaining the Student predictive in (4.6), since we have left the class of prior densities that are natural conjugate for the Normal case (2.4) by assuming prior (functional) independence between β and τ^2 in (2.8).

5. FINITE MIXTURES OF ELLIPTICAL DATA DENSITIES

Although the class of sampling models described in (2.1), and restricted only by (2.6)-(2.7), can already cover many cases used in practical applications, it is still constrained to symmetry and, provided $g_{n,\omega}(\cdot)$ is strictly monotonic, to unimodality. If we wish to circumvent these restrictions, e.g. when faced with a blatantly skewed multimodal empirical data characterization, we can consider the use of finite mixtures of data densities as in (2.1), with (2.6) and (2.7) holding.

Finite mixtures of conjugate prior densities were used to approximate more general classes of priors in Dalal and Hall (1983) and Diaconis and Ylvisaker (1985), but here we introduce the mixing in the sampling model instead. This, of course, widens the family of data densities we can accommodate, and, in principle, $p(\theta, \nu)$ can also involve prior mixtures in our framework, although the latter point will not be elaborated here. We feel it is important to allow for a large enough class of sampling models, since the likelihood is (too) often felt to have some "external validity" [see Berger (1985, p. 249)], and therefore not questioned, whereas we "agree to disagree" on the formulation of the prior. In the terminology of Poirier (1988, p. 130) the "window" entertained should be large enough to interest a "sizeable audience of like-minded researchers". Assessment methods for finite mixtures are found in Dickey and Chen (1985, Section 5), based on elicited quantiles.

If we suitably extend $\theta = (\beta, \eta)$ and ν to parameterize a finite number of densities as in (2.1), each of which has the same scalar precision parameter τ^2 ,

we still have to introduce a mixing parameter λ . Let us, more in detail, analyse the case where we mix only two elliptical densities, implying that λ is scalar. The relevant sampling model becomes

$$p(y | X, \omega, \lambda) = \lambda \left| \frac{1}{\tau^2} V(X, \eta) \right|^{-\frac{1}{2}} g_{n, \omega} [(y - h(X, \beta))' (\frac{1}{\tau^2} V(X, \eta))^{-1} (y - h(X, \beta))] + (1 - \lambda) \left| \frac{1}{\tau^2} W(X, \eta) \right|^{-\frac{1}{2}} k_{n, \omega} [(y - m(X, \beta))' (\frac{1}{\tau^2} W(X, \eta))^{-1} (y - m(X, \beta))],$$

$$0 \leq \lambda \leq 1, \quad (5.1)$$

were both $g_{n, \omega}(\cdot)$ and $k_{n, \omega}(\cdot)$ satisfy condition (2.2), and $m(\cdot)$ and $W(\cdot)$ are known functions in \mathbb{R}^n and the space of all $n \times n$ PDS matrices, respectively. The nuisance parameter τ^2 does not index either of the functions $g_{n, \omega}(\cdot)$ and $k_{n, \omega}(\cdot)$, and we assume the improper prior structure, integrable in ν over N :

$$p(\omega, \lambda) = \frac{c}{\tau^2} p(\theta, \nu, \lambda). \quad (5.2)$$

As in Section 3, this will result in a joint density of $(y, \theta, \nu, \lambda | X)$ that no longer involves the functions $g_{n, \omega}(\cdot)$ or $k_{n, \omega}(\cdot)$. Under the prior in (5.2), mixing any elliptical data densities with common τ^2 has the same consequences for both posterior [on (θ, λ)] and predictive inference as the mixing of Normals. In particular, if the joint density of $(y, \theta, \lambda | X)$ is integrable in (θ, λ) , the posterior of (θ, λ) will be

$$p(\theta, \lambda | y, X) \propto p(\theta, \lambda) \left\{ \lambda \left| V(X, \eta) \right|^{-\frac{1}{2}} d(y, X, \theta)^{-\frac{n}{2}} + (1 - \lambda) \left| W(X, \eta) \right|^{-\frac{1}{2}} e(y, X, \theta)^{-\frac{n}{2}} \right\}, \quad (5.3)$$

with $d(y, X, \theta)$ as in (3.1) and $e(y, X, \theta) = (y - m(X, \beta))' W(X, \eta)^{-1} (y - m(X, \beta))$, whereas the prior density

$$p(\theta, \lambda) = \int_N p(\theta, \nu, \lambda) d\nu$$

must be at least integrable in those elements of θ that appear in only one of the mixed densities in (5.1), due to the summation character of mixtures. The posterior density in (5.3) is a generalization of (3.6), which it reduces to for $\lambda = 1$. For nondegenerate λ , however, the mixing in the data density (5.1) is carried over to the posterior. A convenient choice for the prior of λ may be a beta density, independent of θ , i.e.

$$p(\theta, \lambda) = p(\theta) f_B(\lambda | p, q) \quad (5.4)$$

with $0 \leq \lambda \leq 1$ and $p, q > 0$.

From (5.3) we then obtain the conditional posterior of λ as a mixture of beta densities

$$p(\lambda | \theta, y, X) = (pa_\theta + qb_\theta)^{-1} [pa_\theta f_B(\lambda | p+1, q) + qb_\theta f_B(\lambda | p, q+1)], \quad (5.5)$$

where we have defined

$$\begin{aligned} a_\theta &= |V(X, \eta)|^{-\frac{1}{2}} d(y, X, \theta)^{-\frac{p}{2}} \\ b_\theta &= |W(X, \eta)|^{-\frac{1}{2}} e(y, X, \theta)^{-\frac{q}{2}}. \end{aligned}$$

It is interesting to note that the prior mean of λ , given by $E(\lambda) = \frac{p}{p+q}$ is revised by the data evidence according to the relative posterior "fits" of the elliptical densities in the mixture. If the density multiplied by λ in (5.1) fits badly relative to the other one, a_θ will be much smaller than b_θ : in the extreme case that $a_\theta = 0$, we obtain

$$E(\lambda | \theta, y, X) = E(\lambda | y, X) = \frac{p}{p+q+1},$$

a downward revision of the mean by the sample information. The other extreme with $b_\theta = 0$ will lead to

$$E(\lambda | \theta, y, X) = E(\lambda | y, X) = \frac{p+1}{p+q+1},$$

which is larger than the prior mean. So, although the conditional posterior mean of λ generally depends on θ , the marginal mean will always be confined to the region $[\frac{p}{p+q+1}, \frac{p+1}{p+q+1}]$. Under a uniform prior for λ ($p = q = 1$), the posterior mean will be in $[\frac{1}{3}, \frac{2}{3}]$, an interval which will shrink very quickly if moderately strong prior information on λ is introduced. In the case that we choose $E(\lambda) = \frac{1}{2}$ (i.e. $p = q$) the length of this interval is only four times the prior variance of λ . It thus seems the data evidence can only mildly influence our opinions concerning λ .

The marginal posterior density of θ will be given by

$$p(\theta | y, X) \propto p(\theta) (pa_\theta + qb_\theta),$$

which can be written as the following mixture of the "individual" posteriors, each calculated as in (3.6) on the basis of one of the elliptical models in (5.1):

$$p(\theta | y, X) = (pK_a + qK_b)^{-1} [pK_a p_a(\theta | y, X) + qK_b p_b(\theta | y, X)], \quad (5.6)$$

where

$$\begin{aligned} p_a(\theta | y, X) &= K_a^{-1} p(\theta) a_\theta \\ p_b(\theta | y, X) &= K_b^{-1} p(\theta) b_\theta. \end{aligned}$$

This formulation clearly puts into focus the role of the normalizing constants K_a and K_b , which contribute to the "weights" in the same way as p and q . A similar function was performed by a_θ and b_θ in (5.5). Of course, (5.6) reduces to (3.6) for $q = 0$, in which case (5.4) groups all the prior mass at the point $\lambda = 1$.

From the posterior density in (5.3) it becomes apparent that, unless the functional forms of $h(\cdot)$ and $m(\cdot)$ or those of $V(\cdot)$ and $W(\cdot)$ differ, the mixing in (5.1) will not affect the inference at all. Indeed, then the posterior of λ in (5.5) will reduce to the beta density in the prior (5.4), and the posterior of θ will be the same as (3.6) in Section 3.

Let us now generalize the main results of this section to mixtures of $\ell > 2$ proper elliptical densities. We shall retain the improper prior as in (5.2) for the common nuisance parameter τ^2 , but λ will now be of dimension ℓ , and we shall, therefore, generalize the beta prior in (5.4) to a Dirichlet prior on λ , with the parameter vector $\alpha = (\alpha_1 \dots \alpha_\ell)'$, $\alpha_i > 0$, $\forall i$:

$$p(\lambda | \theta) = p(\lambda) = f_B^\ell(\lambda | \alpha), \quad (5.7)$$

where λ is restrained to the unit simplex (i.e. $\lambda_i > 0$, $\forall i$ and $\sum_{i=1}^{\ell} \lambda_i = 1$). Analogously to a_θ and b_θ in the case $\ell = 2$, we define c_i^θ for the i^{th} density in the mixture, and we denote by e^i the ℓ -dimensional vector with one in the i^{th} position and zeros elsewhere. Then we can state:

Theorem 4: Finite mixtures of ℓ elliptical densities, i.e. an obvious extension of (5.1), with common nuisance parameter τ^2 on which the improper prior (5.2) is defined, will, under (5.7), lead to

$$p(\lambda | \theta, y, X) = \left(\sum_{i=1}^{\ell} \alpha_i c_i^\theta \right)^{-1} \left[\sum_{i=1}^{\ell} \alpha_i c_i^\theta f_B^\ell(\lambda | \alpha + e^i) \right] \quad (5.8)$$

and

$$p(\theta | y, X) = \left(\sum_{i=1}^{\ell} \alpha_i K_i \right)^{-1} \left[\sum_{i=1}^{\ell} \alpha_i K_i p_i(\theta | y, X) \right], \quad (5.9)$$

where $p_i(\theta | y, X) = K_i^{-1} p(\theta) c_i^\theta$, $\forall i$, provided all these posterior densities are well defined. ■

Since the posterior results in (5.8) and (5.9) are also finite mixtures, their analysis is not more difficult than with a single elliptical sampling density. Just like in the previous section, prediction can also be based on mixed sampling models, now using the posterior densities for both λ and θ . Again, we end up with a mixture, as formally stated in the final theorem.

Theorem 5: Under the conditions of Theorem 4, we can base our predictions for a finitely mixed elliptical model on the predictive density

$$p(y_{(2)} | y_{(1)}, X) = \left(\sum_{i=1}^{\ell} \alpha_i L_i \right)^{-1} \left[\sum_{i=1}^{\ell} \alpha_i L_i p_i(y_{(2)} | y_{(1)}, X) \right], \quad (5.10)$$

which is itself a mixture of

$$p_i(y_{(2)} | y_{(1)}, X) = \int_{\Theta} p_i(y_{(2)} | y_{(1)}, X, \theta) p_i(\theta | y_{(1)}, X) d\theta,$$

where $p_i(y_{(2)} | y_{(1)}, X, \theta)$ is the Student t density in (4.1) now corresponding to the i^{th} data density in the mixture, and

$$p_i(\theta | y_{(1)}, X) = L_i^{-1} p(\theta) | V_{i1}^i |^{-\frac{1}{2}} a_i(y_{(1)}, X, \theta)^{-\frac{n_1}{2}},$$

as in (4.2), where each L_i must be finite, and indices i refer to the i^{th} data density throughout. ■

As in Section 4, if we wish to use posterior densities for θ , computed after observing $y_{(1)}$ and only part of X , namely X_1 , we need a bit more. Imposing condition (4.4) on every data density that is used in the sampling model will be sufficient.

We suggest approximating non-elliptical (e.g. asymmetric) sampling processes by such finite mixtures of elliptical densities, since the mixing will be preserved in both posterior and predictive analyses. We thus have a way of considerably broadening the class of data densities, without really adding to the complexity of the analysis.

6. CONCLUDING REMARKS

Under certain conditions, it was shown that Bayesian posterior and predictive analysis is perfectly robust with respect to the choice of a sampling density within the entire class of elliptical densities. Sufficient conditions are that we can single out a scale factor τ^2 that does not influence the way the density changes over ellipsoids, and that we specify an improper prior density on τ^2 .

Once the scale factor is then integrated out, the tails of the sampling density do not matter anymore, only the location and shape of the ellipsoids, parameterized by θ , are relevant. The posterior of θ will then be given by the simple expression in Theorem 1, which is the same as in the Normal case. The only purpose of the parameter ν is to describe the tails of the data density. Thus, if the latter become irrelevant, then, clearly, the sample can not directly revise our opinion about ν . It can only do so through revising θ if there is prior dependence between θ and ν . This is the object of Theorem 2.

Our conclusions are similar for prediction: given an improper prior on the nuisance parameter τ^2 , everything is just like in the Normal regression model. Theorem 3 summarizes these findings.

The results from Sections 1 through 4 can be related to previous work in this area; in particular, our paper extends the framework of scale mixtures of Normal densities, found in Jammalamadaka *et al.* (1987), Chib *et al.* (1988), Osiewalski (1990) and Chib *et al.* (1990), to general elliptical densities. It also broadens the linear regression model, used in the first two of the above references, to a possibly nonlinear one. Taking into account that only a diffuse prior for τ^2 was considered in the present paper, we can establish the following correspondences. Within the class of scale mixtures of Normals, Proposition 1 of Jammalamadaka *et al.* (1987) is a special case of our Corollary 2 for $V(X, \eta) = I_n$, whereas Theorem 3 generalizes Proposition 1 of Chib *et al.* (1988), who assumed linearity and a uniform prior on β . Both Theorems 1 and 3 extend results obtained under scale mixtures of Normals in Osiewalski (1990) to general elliptical densities, and Theorem 2 generalizes Theorem 2 in Chib *et al.* (1990) in the same way.

If the inherent symmetry of the single elliptical data density is found to be too restrictive, we can make use of finite mixtures of elliptical densities to approximate some non-elliptical data density. These mixtures are then carried over to posterior

and predictive results, without leading to an increase in complexity (see Theorems 4 and 5, respectively). Since we have updating of the mixing parameter λ as well, we could even think of incorporating this into a testing framework, where we intend to choose between several competing models. The updating of the prior of λ by the sample will indicate which model is most favoured on the grounds of posterior fit. Note that the contenders have to correspond to different ellipsoids, e.g. through different functional form or choice of regressors. Mixing e.g. a Normal and a Cauchy defined over the same ellipsoid will, of course, give the same results as with a single Normal data density. Also, we have seen that the sample information on λ can easily be drowned by moderately informative prior notions.

The findings in this paper generalize and explain many results that have appeared in the literature, and give remarkably weak sufficient conditions for robustness with respect to the data density. This provides us with a fairly strong argument in favour of using the standard Normal results in regression models, and gives an implicit motivation for stressing sensitivity with respect to the choice of the prior density instead.

REFERENCES

- Bagchi, P. and I. Guttman, 1988, Bayesian regression analysis under non-normal errors, in: J.M. Bernardo, M.H. DeGroot, D.V. Lindley and A.F.M. Smith, eds., *Bayesian Statistics 3* (Oxford University Press, Oxford), 543-552.
- Berger, J.O., 1985, *Statistical Decision Theory and Bayesian Analysis* (Springer-Verlag, New York).
- Box, G.E.P. and G.C. Tiao, 1973, *Bayesian Inference in Statistical Analysis* (Addison-Wesley, Reading).
- Cambanis, S., S. Huang and G. Simons, 1981, On the theory of elliptically contoured distributions, *Journal of Multivariate Analysis* 11, 368-385.
- Chib, S., J. Osiewalski and M. Steel, 1990, Posterior inference on degrees of freedom in multivariate- t regression models, mimeo.
- Chib, S., R.C. Tiwari and S.R. Jammalamadaka, 1988, Bayes prediction in regressions with elliptical errors, *Journal of Econometrics* 38, 349-360.
- Dalal, S. and W.J. Hall, 1983, Approximating priors by mixtures of natural conjugate priors, *Journal of the Royal Statistical Society B* 45, 278-286.
- Diaconis, P. and D. Ylvisaker, 1985, Quantifying prior opinion, in: J.M. Bernardo, M.H. DeGroot, D.V. Lindley and A.F.M. Smith, eds., *Bayesian Statistics 2* (North-Holland, Amsterdam), 133-156.
- Dickey, P. and C.-H. Chen, 1985, Direct subjective-probability modelling using ellipsoidal distributions, in: J.M. Bernardo, M.H. DeGroot, D.V. Lindley and A.F.M. Smith, eds., *Bayesian Statistics 2* (North-Holland, Amsterdam), 57-182.
- Drèze, J.H., 1977, Bayesian regression analysis using poly- t densities, *Journal of Econometrics* 6, 329-354.
- Drèze, J.H. and J.-F. Richard, 1983, Bayesian analysis of simultaneous equation systems, Ch. 3 in: Z. Griliches and M.D. Intriligator, eds., *Handbook of Econometrics* (North-Holland, Amsterdam).

- Florens, J.P. and M. Mouchart, 1985, Conditioning in dynamic models, *Journal of Time Series Analysis* 6, 15-34.
- Florens, J.P., M. Mouchart and J.M. Rolin, 1990, *Elements of Bayesian Statistics* (M. Dekker, New York).
- Hill, B.M., 1969, Foundations for the theory of least squares, *Journal of the Royal Statistical Society B* 31, 89-97.
- Jammalamadaka, S.R., R.C.Tiwari and S. Chib, 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 generalization, *Sankhyā A* 32, 419-430.
- Leamer, E.E., 1978, *Specification Searches* (Wiley, New York).
- Osiewalski, J., 1990, A note on Bayesian inference in a regression model with elliptical errors, *Journal of Econometrics*, forthcoming.
- Osiewalski, J. and M. Steel, 1990, Semi-conjugate prior densities in multivariate t regression models, CORE Discussion Paper 9018, Université Catholique de Louvain, Louvain-la-Neuve.
- Poirier, D.J., 1988, Frequentist and subjectivist perspectives on the problems of model building in economics, *Economic Perspectives* 2, 121-144.
- Ramsay, J.O. and M.R. Novick, 1980, PLU robust Bayesian decision theory: point estimation, *Journal of the American Statistical Association* 75, 901-907.
- Richard, J.-F. and H. Tompa, 1980, On the evaluation of poly- t density functions, *Journal of Econometrics* 12, 331-351.
- West, M., 1984, Outlier models and prior distributions in Bayesian linear regression, *Journal of the Royal Statistical Society B* 46, 431-439.
- Zellner, A., 1976, Bayesian and non-Bayesian analysis of the regression model with multivariate Student- t error terms, *Journal of the American Statistical Association* 71, 400-405.

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

(For previous papers please consult previous discussion papers.)

No.	Author(s)	Title
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 Cardinaly 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
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

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

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

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

No.	Author(s)	Title
9022	K. Kamiya and D. Talman	Linear Stationary Point Problems
9023	W. Emons	Good Times, Bad Times, and Vertical Upstream Integration
9024	C. Dang	The D_2 -Triangulation for Simplicial Homotopy Algorithms for Computing Solutions of Nonlinear Equations
9025	K. Kamiya and D. Talman	Variable Dimension Simplicial Algorithm for Balanced Games
9026	P. Skott	Efficiency Wages, Mark-Up Pricing and Effective Demand
9027	C. Dang and D. Talman	The D_1 -Triangulation in Simplicial Variable Dimension Algorithms for Computing Solutions of Nonlinear Equations
9028	J. Bai, A.J. Jakeman and M. McAleer	Discrimination Between Nested Two- and Three- Parameter Distributions: An Application to Models of Air Pollution
9029	Th. van de Klundert	Crowding out and the Wealth of Nations
9030	Th. van de Klundert and R. Gradus	Optimal Government Debt under Distortionary Taxation
9031	A. Weber	The Credibility of Monetary Target Announce- ments: An Empirical Evaluation
9032	J. Osiewalski and M. Steel	Robust Bayesian Inference in Elliptical Regression Models

P.O. BOX 90153. 5000 LE TILBURG. THE NETHERLANDS

Bibliotheek K. U. Brabant



17 000 011 17592 5