Posterior and predictive densities for nonlinear regression

Osiewalski, J.

TILBURG UNIVERSITY
KATHOLIEKE UNIVERSITEIT BRABANT
POSTBOX 90153
5000 LE TILBURG
THE NETHERLANDS

DEPARTMENT OF ECONOMICS

RESEARCH MEMORANDUM

POSTERIOR AND PREDICTIVE DENSITIES FOR NONLINEAR REGRESSION A PARTLY LINEAR MODEL CASE

Jacek Osiewalski

FEW 333

Posterior and Predictive Densities for Nonlinear Regression.
A Partly Linear Model Case

Jacek Osiewalski

Academy of Economics
ul. Rakowicka 27
31-510 Kraków
Poland

Tilburg, June 1988.

The paper was written at Tilburg University during a research stay supported by the Netherlands Organization for Scientific Research (NWO). I greatly benefited from discussions with Mark F.J. Steel and from his comments on the first draft.

1. Introduction

The paper continues the Bayesian analysis of nonlinear regression models, that is models of known functional form (nonlinear in parameters) and with an additive error term. In this area of Bayesian research, Zellner (1971, §6.2), Sankar (1970), H. Tsurumi and Y. Tsurumi (1976), Harkema and Schim van der Loeff (1977) focus their attention on the estimation of CES production functions parameters; Box and Tiao (1973, p. 436) present an approximate Bayesian approach based on linearization; Eaves (1983) considers a reference prior - in the sense of Bernando (1979) - and gives an illustration of discrepancies between exact and approximate posterior densities; Broemeling (1985, p. 104-116) presents general formulae of posterior and predictive densities and points at some easy special cases and at useful approximations.
In Osiewalski (1987) and Osiewalski and Goryl (1986, 1988) - all in Polish - posterior densities and moments for some specific nonlinear models (logistic growth function, Törnquist-type Engel curves) under Jeffreys' (or reference) priors are derived.

This paper generalizes the approach adopted previously for the CES functions and deals with Bayesian estimation and prediction for those nonlinear regression models which are linear in some parameters (say, $\beta_1,\dots,\beta_k$) given values of the remaining parameters (say, $\eta_1,\dots,\eta_q$). This class of nonlinear regression models is worth considering since the exact Bayesian analysis with an appropriately chosen prior requires only q- or (q+1)-dimensional numerical integrations, irrespective of k.

The plan of the paper is as follows: some introductory remarks on Bayesian nonlinear regression are given in Section 2; in Section 3 the class of "partly linear" nonlinear models is introduced; posterior and predictive densities and moments under the assumption of improper uniform prior of $\beta = (\beta_1,\dots,\beta_k)'$ are derived in Section 4, where the correspondence of this prior with Jeffreys' prior is discussed as well. The use of a finite mixture of normal or t distributions as a prior of β is considered in Section 5 and corresponding posterior and predictive densities are derived there. Concluding remarks and comments on applications are given in Section 6.

1.1. Notation and main identities

Throughout the paper p(.) denotes a probability density function (PDF) with special notation for PDF's of gamma, normal and t distributions. For $x \in R^k$, $p_N^k(x|c,W)$ denotes a k-variate normal PDF with a mean vector c and a covariance matrix W, and $p_S^k(x|r, c, T)$ denotes a k-variate Student t PDF with r degrees of freedom, σ noncentrality vector c and a precision matrix T.

For $w \in R_+$,

$$p_\gamma(w|a,b) = \frac{b^a}{\Gamma(a)} w^{a-1} \exp(-bw),$$

that is a gamma PDF with parameters a > 0, b > 0. The following identities are used:

$$\int_{R^k} p_N^m(y|Qx + a, S)\, p_N^k(x|b,c)\, dx = p_N^m(y|Qb + a, S + QCQ'), \tag{1.1}$$

$$p_N^k(x|c, \tfrac{1}{w} A^{-1})\, p_\gamma(w|\tfrac{a}{2},\tfrac{b}{2}) =$$

$$= p_\gamma(w|\tfrac{a+k}{2}, \tfrac{1}{2}[b + (x-c)' A(x-c)])\, p_S^k(x|a, c, \tfrac{a}{b} A), \tag{1.2}$$

and

$$\int_0^\infty p_N^k(x|c, \tfrac{1}{w} A^{-1})\, p_\gamma(w|\tfrac{a}{2},\tfrac{b}{2})\, dw = p_S^k(x|a, c, \tfrac{a}{b} A), \tag{1.3}$$

which is an immediate consequence of (1.2).

2. A Bayesian approach to nonlinear regression models

Let us consider the nonlinear regression model

$$y_t = h(z_t;\vartheta) + u_t \quad , \quad u_t \sim iiN(0,\sigma^2) \tag{2.1}$$

where z_t is a $r \times 1$ vector of independent variables, ϑ is a $K \times 1$ unknown parameter vector, $z_t \in R^r$, $\vartheta \in \Theta$ (Θ is a full-dimensional subset of R^k), $h: R^r \times \Theta \to R$ is a known function; σ^2 is an unknown nuisance parameter. We assume that $h(z_t;\vartheta)$ is not linear in ϑ (given z_t) and that this function (as a function of ϑ given z_t) is sufficiently well-behaved to insure the existence of certain derivatives and integrals which appear in Bayesian analysis. We treat z_t as a known nonstochastic vector[1)] and assume that (given $z_1,\dots,z_n, z_{n+1},\dots,z_{n+m}$) one observes $y = (y_1,\dots,y_n)'$ and one has to make inferences about ϑ and to forecast $\tilde{y} = (\tilde{y}_1,\dots,\tilde{y}_m)' = (y_{n+1},\dots,y_{n+m})'$.
Let

$$Z = (z_1 \; z_2 \; \dots \; z_n)', \; \tilde{Z} = (z_{n+1} \; \dots \; z_{n+m})', \; \omega = \sigma^{-2}$$

and let $p(y,\tilde{y}|Z,\tilde{Z},\vartheta,\omega)$ denote the joint density of current and future observations given the values of independent variables and parameters. In our i.i.d. case

$$p(y,\tilde{y}|Z,\tilde{Z},\vartheta,\omega) = p(y|Z,\vartheta,\omega)\; p(\tilde{y}|\tilde{Z},\vartheta,\omega)$$

---

1) Of course, this assumption can be relaxed. If z_t is weakly exogenous for (ϑ,σ^2) over the sample period and parameters of the marginal process generating z_t are prior independent of (ϑ,σ^2), then all the posterior results (obtained in Sections 4, 5) hold. If, additionally, z_t is strongly exogenous for (ϑ,σ^2) over the forecast period, then all the predictive results hold. For definitions of weak and strong exogeneity see Engle, Hendry and Richard (1983).

and all the densities are densities of appropriate normal distributions: $(n+m)$-, n- and m-dimensional, respectively.

In the Bayesian approach, all inferences about ϑ are based on the marginal posterior PDF $p(\vartheta|y,Z)$ obtained from the joint posterior PDF

$$p(\vartheta,\omega|y,Z) \propto p(\vartheta,\omega).p(y|Z,\vartheta,\omega)$$

where $p(\vartheta,\omega)$ is the prior PDF. Bayesian prediction of $\tilde{y}$ is based on the predictive PDF

$$p(\tilde{y}|y,Z,\tilde{Z}) = \int_{\Theta}\int_{0}^{\infty} p(\tilde{y}|\tilde{Z},\vartheta,\omega)\; p(\vartheta,\omega|y,Z)\; d\omega\; d\vartheta; \tag{2.2}$$

see Zellner (1971, ch. 2). When the prior density of ϑ and ω is composed of a gamma PDF on ω and an independent prior on ϑ:

$$p(\vartheta,\omega) \propto p(\vartheta)\; p_{\gamma}(\omega|\tfrac{e}{2},\, \tfrac{f}{2}),\; \vartheta \in \Theta,\; \omega \in R_{+},$$

then - rewriting Broemeling's formulae (3.81) and (3.77)-(3.80) in our notation - we have, using the natural-conjugate properties of gamma densities for our model,

$$p(\vartheta|y,Z) \propto p(\vartheta) \left\{ f + \sum_{t=1}^{n} [y_t - h(z_t;\vartheta)]^2 \right\}^{-\frac{1}{2}(e+n)} \tag{2.3}$$

as the unnormalized posterior PDF of ϑ and

$$p(\tilde{y}_j|y,Z,z_{n+j}) = \int_{\Theta} p_S^1\left[\tilde{y}_j \,\middle|\, e+n,\; h(z_{n+j};\vartheta),\; \frac{e+n}{f + \sum_{t=1}^{n}[y_t - h(z_t;\vartheta)]^2}\right] \times$$

$$\times\; p(\vartheta|y,Z)\; d\vartheta \tag{2.4}$$

as the (marginal) univariate predictive PDF. In this general formulation of the model, the posterior and predictive PDF's are not tractable and (generally) require K-dimensional numerical integrations in order to obtain the normalizing constant, moments, univariate posterior densities etc. Broemeling (1985, p. 107) writes: "One may sum up the Bayesian analysis of nonlinear regression, when ϑ is scalar, by saying a complete analysis is possible (...); however, if ϑ is of dimension greater than or equal to two, a Bayesian analysis becomes more difficult" and when ϑ is of dimension 3 or greater "the numerical integration problems become impractical". On the oher hand, Broemeling (1985, p. 108) realizes that "there are some special cases, where an exact and complete Bayesian analysis is possible" and gives as an example

$$h(z_t,\vartheta) = \vartheta_2\, h_1(z_t,\vartheta_1),\ \vartheta_1 \in R,\ \vartheta_2 \in R.$$

Let us note, however, that in econometric literature much more complicated functional forms of h were successfully analyzed, namely the forms obtained by taking logarithms of both sides of different CES production functions with multiplicative lognormal errors; see Sankar (1970), H. Tsurumi and Y. Tsurumi (1976).

Bayesian estimation of 5 or more unkown parameters of CES functions required bivariate or trivariate numerical integration; great analytical simplifications were possible because the models were linear in some parameters and uniform priors for these parameters led to "partly tractable" posteriors. The obvious conclusion is that the Bayesian analysis of a given nonlinear model should exploit linearities in order to become "more practical". The aim of this paper is to provide general formulae of posterior and predictive densities and moments for the case of a nonlinear model which is linear in some parameters. The approach, used previously for some specific cases, is generalized into three main directions:

1) a general - not specific - form of "partly linear" model is considered,
2) not only uniform improper but also some proper informative priors are allowed,
3) not only posterior, but also predictive PDF's are derived.

3. Partly linear regression models

Let us restrict our considerations to nonlinear models of the following functional form:

$$h(z_t,\vartheta) = x_0(z_t,\eta) + \beta_1 x_1(z_t,\eta) + \ldots + \beta_k x_k(z_t,\eta), \tag{3.1}$$

where $\vartheta = (\beta', \eta')'$, $\beta = (\beta_1,\ldots,\beta_k)' \in R^k$
$\eta = (\eta_1,\ldots,\eta_q)' \in H \subset R^q$, $k + q = K$,

$x_i(z_t,\eta)$ for $i = 0, 1,\ldots,k$ are known functions (sufficiently well-behaved), and H is a (full-dimensional) set of admissible values of η. That is, we are interested in models where it is possible to divide a parameter vector (ϑ) into two separate subvectors (β and η) in such a way that - given η - the model is linear with respect to β. For n observations ($t = 1,\ldots,n$) and m values to be predicted ($t = n + 1,\ldots,n + m$) we have

$$\begin{bmatrix} y \\ \tilde{y} \end{bmatrix} = \begin{bmatrix} w_\eta \\ \tilde{w}_\eta \end{bmatrix} + \begin{bmatrix} X_\eta \\ \tilde{X}_\eta \end{bmatrix}\beta + \begin{bmatrix} u \\ \tilde{u} \end{bmatrix},$$

$$\begin{bmatrix} u \\ \tilde{u} \end{bmatrix} \sim N(0, \tfrac{1}{\omega} I_{n+m}),$$

where

$$w_\eta = \begin{bmatrix} x_0(z_1,\eta) \\ \cdot \\ \cdot \\ \cdot \\ x_0(z_n,\eta) \end{bmatrix}, \qquad X_\eta = \begin{bmatrix} x_1(z_1,\eta) & \ldots & x_k(z_1,\eta) \\ \cdot & & \cdot \\ \cdot & & \cdot \\ \cdot & & \cdot \\ x_1(z_n,\eta) & \ldots & x_k(z_n,\eta) \end{bmatrix},$$

$$\tilde{w}_\eta = \begin{bmatrix} x_0(z_{n+1},\eta) \\ \cdot \\ \cdot \\ \cdot \\ x_0(z_{n+m},\eta) \end{bmatrix}, \quad \tilde{X}_\eta = \begin{bmatrix} x_1(z_{n+1},\eta) & \cdots & x_k(z_{n+1},\eta) \\ \cdot & & \cdot \\ \cdot & & \cdot \\ \cdot & & \cdot \\ x_1(z_{n+m},\eta) & \cdots & x_k(z_{n+m},\eta) \end{bmatrix},$$

$$u = (u_1, \ldots, u_n)' \quad , \quad \tilde{u} = (u_{n+1}, \ldots, u_{n+m})'.$$

The data distribution and the distribution of future values are independent normal distributions:

$$p(y,\tilde{y}|Z,\tilde{Z},\beta,\eta,\omega) = p(y|Z,\beta,\eta,\omega)\ p(\tilde{y}|\tilde{Z},\beta,\eta,\omega),$$

$$p(y|Z,\beta,\eta,\omega) = p_N^n(y|X_\eta\beta + w_\eta,\ \omega^{-1}I_n),$$

$$p(\tilde{y}|\tilde{Z},\beta,\eta,\omega) = p_N^m(\tilde{y}|\tilde{X}_\eta\beta + \tilde{w}_\eta,\ \omega^{-1}I_m).$$

Let us assume that the n×k matrix X_η is of full column rank (k) for every $\eta \in H$. Now $X_\eta'X_\eta$ is a nonsingular k×k matrix and we can define

$$b_\eta = (X_\eta'X_\eta)^{-1}\ X_\eta'(y - w_\eta),\ s_\eta = (y - w_\eta - X_\eta b_\eta)'\ (y - w_\eta - X_\eta b_\eta).$$

The following equality holds:

$$(y - w_\eta - X_\eta\beta)'\ (y - w_\eta - X_\eta\beta) = s_\eta + (\beta - b_\eta)'\ X_\eta'X_\eta(\beta - b_\eta),$$

which enables us to write the data density (or the likelihood function) in the more convenient form

$$p(y|Z,\beta,\eta,\omega) = (2\pi)^{-\frac{n}{2}}\ \omega^{\frac{n}{2}}\ \exp\left\{-\ \frac{\omega}{2}[s_\eta + (\beta-b_\eta)'X_\eta X_\eta(\beta-b_\eta)]\right\}. \qquad (3.2)$$

Except for the dependence of X_η, b_η, s_η on unknown η, the likelihood function looks like its counterpart for the usual linear model. In order to take advantage of this similarity, an appropriate class of priors should be considered. In the linear case (η known) Jeffreys' improper prior and proper natural conjugate normal-gamma priors of (β,ω) give completely analytical posterior and predictive results, and independent priors of the form: Student t (or normal for β and gamma for ω lead to univariate numerical integrations only, so these facts suggest the classes of priors worth considering in our nonlinear - partly linear case. We assume here that $\theta = (\beta'\eta')'$ and ω are independent a priori:

$$p(\beta,\eta,\omega) = p(\beta,\eta)\ p(\omega) \tag{3.3}$$

and

$$p(\omega) \propto \omega^{\frac{e}{2}-1} \exp(-\tfrac{f}{2}.\omega);$$

$e = f = 0$ correspond to the improper prior $p(\omega) \propto \omega^{-1}$, if $e > 0$ and $f > 0$ then $p(\omega) = p_{\gamma}(\omega|\frac{e}{2}, \frac{f}{2})$.

We will consider three types of priors of (β,η):

$$p_1(\beta,\eta) \propto g(\eta),$$

$$p_i(\beta,\eta) = p_i(\beta|\eta)\ p(\eta), \quad i = 2,3.$$

For $i = 2,3$, $p(\eta)$ is some marginal prior density of η and $p_i(\beta/\eta)$ are informative conditional priors of β (given η) which are finite mixtures of normal ($i=2$) or t ($i=3$) densities.

In the case of $p_1(\beta,\eta)$, $g(\eta)$ need not be the marginal prior since we do not impose the assumption of prior independence between β and η, but only between (β,η) and ω. A clarifying example of this is given by Jeffreys' prior (4.2) in Subsection 4.2.
Of course, the assumption (3.3) will complicate the analysis in the cases of informative priors of β. Conditionally on η, the natural-conjugate framework would be more convenient. That is, we could assume that $p(\beta,\omega|\eta)$ is a finite mixture of normal-gamma densities and that $p(\eta)$ is any given PDF. Then, the similar lines of reasoning as in subsection 5.2 would lead

to analytical integrations with respect to β and ω, and only q-dimensional numerical integration (with respect to η) would be required. But, as in the linear case, the dependence between β and ω in the natural-conjugate normal-gamma prior may cause some serious problems, e.g. it may require rather strong prior beliefs about ω in order to achieve some preassumed prior moments of β. The assumption (3.3) enables to avoid such troubles at the expense of one more dimension in numerical integration. In particular, this assumption allows for expressing strong beliefs about β and the lack of opinion about ω simultaneously.

4. Bayesian analysis with an improper uniform prior density for β

4.1. Posterior and predictive PDF's

For the likelihood function (3.2) and for the prior density

$$p_1(\beta,\eta,\omega) \propto g(\eta)\ \omega^{\frac{e}{2}-1} \exp(-\tfrac{f}{2}.\omega),$$

$$\beta \in R^k,\ \eta \in H \subset R^q,\ \omega \in R_+,$$

we obtain the following joint posterior PDF:

$$p_1(\beta,\eta,\omega|y,Z) \propto g(\eta)\ (f + s_\eta)^{-\frac{e+n-k}{2}} |X_\eta' X_\eta|^{-\frac{1}{2}}\ p_\gamma\left[\omega|\frac{e+n-k}{2},\ \frac{f+s_\eta}{2}\right] \times$$

$$\times\ p_N^k\left[\beta|b_\eta,\ \frac{1}{\omega}(X_\eta' X_\eta)^{-1}\right].$$

Now the joint posterior PDF can easily be represented as a product of appropriate marginal and conditional PDF's:

$$p_1(\beta,\eta,\omega|y,Z) = p_1(\eta|y,Z)\ p_1(\omega|y,Z,\eta)\ p_1(\beta|y,Z,\eta,\omega),$$

where

$$p_1(\eta|y,Z) \propto g(\eta)|X_\eta' X_\eta|^{-\frac{1}{2}} (f + s_\eta)^{-\frac{e+n-k}{2}},$$

$$p_1(\omega|y,Z,\eta) = p_\gamma\left[\omega|\frac{e+n-k}{2},\ \frac{f+s_\eta}{2}\right],$$

$$p_1(\beta|y,Z,\eta,\ \omega) = p_N^k\left[\beta|b_\eta,\ \frac{1}{\omega}(X_\eta' X_\eta)^{-1}\right].$$

Since - according to (1.2) -

$$p_N^k\left[\beta|b_\eta,\ \frac{1}{\omega}(X_\eta' X_\eta)^{-1}\right]\ p_\gamma\left[\omega|\frac{e+n-k}{2},\ \frac{f+s_\eta}{2}\right] =$$

$$= p_\gamma\left[\omega \middle| \frac{e+n}{2}, \frac{1}{2}[f + s_\eta + (\beta - b_\eta)'X_\eta'X_\eta(\beta - b_\eta)]\right] p_S^k\left[\beta \middle| e+n-k, b_\eta, \frac{e+n-k}{f+s_\eta} X_\eta'X_\eta\right],$$

then the joint posterior PDF can be written equivalently as

$$p_1(\beta,\eta,\omega|y,Z) = p_1(\eta|y,Z)\ p_1(\beta|y,Z,\eta)\ p_1(\omega|y,Z,\beta,\eta),$$

where

$$p_1(\beta|y,Z,\eta) = p_S^k\left[\beta \middle| \bar{e}, b_\eta, \frac{\bar{e}}{f + s_\eta} X_\eta'X_\eta\right],$$

$$p_1(\omega|y,Z,\beta,\eta) = p_\gamma\left[\omega \middle| \frac{e + n}{2}, \frac{d}{2}\right],$$

and

$$\bar{e} = e + n - k \quad , \quad d = f + s_\eta + (\beta - b_\eta)'\ X_\eta'X_\eta(\beta - b_\eta).$$

For inferences about η, the marginal posterior PDF $p_1(\eta|y,Z)$ should be used. This PDF is - in general - intractable and numerical integrations will usually be required to calculate a normalizing constant, moments and univariate marginal densities. For inferences about β, the marginal posterior PDF

$$p_1(\beta|y,Z) = \int_H p_1(\eta|y,Z)\ p_1(\beta|y,Z,\eta)\ d\eta$$

is appropriate. Since the conditional posterior PDF, $p_1(\beta|y,Z,\eta)$, is in Student t form, we have analytical formulae for conditional posterior moments and also for univariate PDF's $p_1(\beta_i|y,Z,\eta)$, $i = 1,\ldots,k$, which are of univariate Student t form.

Now the marginal posterior moments and densities can be calculated as

$$E_1(\beta|y,Z) = \int_H p_1(\eta|y,Z)\ b_\eta d\eta,$$

$$E_1(\beta\beta'|y,Z) = \int_H p_1(\eta|y,Z) \left[\frac{f + s_\eta}{\bar{e} - 2} (X_\eta'X_\eta)^{-1} + b_\eta . b_\eta'\right] d\eta,$$

$$p_1(\beta_i|y,Z) = \int_H p_1(\eta|y,Z)\ p_1(\beta_i|y,Z,\eta)\ d\eta.$$

For the calculation of mixed moments one can use the formula

$$E_1(\beta.\eta'|y,Z) = \int_H p_1(\eta|y,Z)\ b_\eta . \eta'\ d\eta.$$

In this paper ω is treated as a nuisance parameter, but if some inferences about ω are necessary then the marginal posterior PDF

$$p_1(\omega|y,Z) = \int_H p_1(\eta|y,Z)\ p_1(\omega|y,Z,\eta)\ d\eta$$

is appropriate and the known properties of $p_1(\omega|y,Z,\eta) = p_\gamma\left[\omega|\frac{\bar{e}}{2}, \frac{f + s_\eta}{2}\right]$ can be used.

In order to derive the predictive PDF $p_1(\tilde{y}|y,Z,\tilde{Z})$ according to (2.2), let us notice that in our case we can write

$$p_1(\tilde{y}|y,Z,\tilde{Z}) = \int_H \int_0^\infty \int_{R^k} p(\tilde{y}|\tilde{Z},\beta,\eta,\omega)\ p_1(\beta|y,Z,\eta,\omega)\ d\beta\ \times$$

$$\times\ p_1(\omega|y,Z,\eta)\ d\omega\ p_1(\eta|y,Z)\ d\eta.$$

Since the first and second densities after the integral signs are normal and the third density is gamma, then by successive analytical integrations based on (1.1) and (1.3) one obtains

$$p_1(\tilde{y}|y,Z,\tilde{Z}) = \int_H p_S^m\left[\tilde{y}|\bar{e},\ \tilde{w}_\eta + \tilde{X}_\eta b_\eta,\ \frac{\bar{e}}{f + s_\eta}\left[I_m + \tilde{X}_\eta (X_\eta'X_\eta)^{-1}\tilde{X}_\eta'\right]^{-1}\right] \times$$

$$\times\ p_1(\eta|y,Z)\ d\eta.$$

Moments and univariate densities of the predictive distribution can be calculated in the similar way as in the case of marginal posterior distribution of β. It means that, in the case of improper uniform prior of β, Bayesian estimation and prediction requires numerical integrations over

$H \subset R^q$, irrespective of k (the dimension of β). Of course, k plays a great role in calculating values of the integrand, since a k×k matrix $(X_\eta' X_\eta)$ should be inverted for every η.

4.2. Uniform prior of β and Jeffreys' rule

Since the use of a prior from the class $p_1(\beta,\eta,\omega)$ greatly simplifies the forms of posterior and predictive PDF's, let us comment on its uniformity in β. Intuitively, the uniform prior of β represents vague prior knowledge about $\beta_1,\ldots,\beta_k$ and indeed it was used as a noninformative prior in the case of CES function by H. Tsurumi and Y. Tsurumi (1976)[2) and Sankar (1970).
But does this prior follow from any formal principle (as it does in the case of linear model, where the uniform prior can be justified in several ways)?

Let us return to the general model (2.1), that is

$$y_t = h(z_t,\vartheta) + u_t \;,\; u_t \sim iiN(0,\sigma^2),$$

and denote by D the n×K matrix of first-order partial derivatives

$$d_{ti} = \frac{\partial h(z_t,\vartheta)}{\partial \vartheta_i} \qquad (t = 1, \ldots,n \;;\; i = 1, \ldots,K).$$

We can write the information matrix (based on n observations) as

$$I(\vartheta,\sigma) = \left[\begin{array}{c|c} \sigma^{-2} D'D & 0 \\ \hline 0 & 2n\sigma^{-2} \end{array}\right],$$

so an application of Jeffrey's rule separately for ϑ and for σ gives

----

2) In fact, they assumed the uniform prior of all structucal parameters.

$$p_J(\vartheta) \propto |D'D|^{\frac{1}{2}} \quad , \quad p_J(\sigma) \propto \sigma^{-1},$$

and

$$p_J(\vartheta,\sigma) = p_J(\vartheta)\; p_J(\sigma) \propto \sigma^{-1}|D'D|^{\frac{1}{2}}$$

or, equivalently, in terms of $\omega = \sigma^{-2}$

$$p_J(\vartheta,\omega) \propto \omega^{-1}|D'D|^{\frac{1}{2}}.$$

As Eaves (1983) pointed out, $p_J(\vartheta,\sigma)$ is also a reference prior in the sense of Bernardo (1979), assuming that ϑ is a parameter of interest and σ is a nuisance parameter.
For the partly linear model:

$$h(z_t;\vartheta) = h(z_t;\, \beta,\eta) = x_0(z_t,\eta) + \beta_1 x_1(z_t,\eta) + \ldots + \beta_k x_k(z_t,\eta),$$

D can be partitioned as $D = [D_1\ D_2]$, where D_1 is $n\times k$ and consists of the following derivatives:

$$\frac{\partial h(z_t;\, \beta,\eta)}{\partial \beta_i} = x_i(z_t,\eta) \qquad \begin{bmatrix} t = 1, \ldots, n \\ i = 1, \ldots, q \end{bmatrix},$$

that is $D_1 = X_\eta$, and D_2 is $n\times q$ and consists of the following derivatives

$$\frac{\partial h(z_t;\, \beta,\eta)}{\partial \eta_j} = \frac{\partial x_0(z_t,\eta)}{\partial \eta_j} + \beta_1\frac{\partial x_1(z_t,\eta)}{\partial \eta_j} + \ldots + \beta_k\frac{\partial x_k(z_t,\eta)}{\partial \eta_j}$$

$$(t = 1, \ldots, n;\ j = 1, \ldots, q);$$

D_2 can be presented as $D_2 = \left[\dot{w}_1 + \dot{X}_1\beta \ \ \ldots \ \ \dot{w}_q + \dot{X}_q\beta\right]$, where $\dot{w}_j$ is a $n\times 1$ vector consisting of $\frac{\partial x_0(z_t,\eta)}{\partial \eta_j}$ (t = 1,...,n) and $\dot{X}_j$ is a n×k matrix consisting of $\frac{\partial x_i(z_t,\eta)}{\partial \eta_j}$ (t = 1,...,n; i = 1,...,k). Thus in the case of the partly linear model we obtain

$$p_J(\vartheta) = p_J(\beta,\eta) \propto \begin{vmatrix} X_\eta' X_\eta & X_\eta' D_2 \\ D_2' X_\eta & D_2' D_2 \end{vmatrix}^{\frac{1}{2}}$$

and Jeffreys' (or Bernardo's reference) prior may depend on β since D_2 depends on β for nonzero $\dot{X}_1,\dots,\dot{X}_q$. This leads to the following conclusion: for the special subclass of partly linear models where $x_1, \dots, x_k$ do not depend on η, that is for

$$h(z_t;\, \beta,\eta) = x_0(z_t,\eta) + \beta_1 x_1(z_t) + \dots + \beta_k x_k(z_t), \tag{4.1}$$

Jeffreys' (or reference) prior takes the form[3)]

$$p_J(\beta,\, \eta,\, \omega) \propto \omega^{-1}\, g_J(\eta), \tag{4.2}$$

where

$$g_J(\eta) \propto \begin{vmatrix} X_\eta' X_\eta & X_\eta' \dot{W} \\ \dot{W}' X_\eta & \dot{W}' \dot{W} \end{vmatrix}^{\frac{1}{2}}, \quad \dot{W} = [\dot{w}_1 \ \dots \ \dot{w}_q],$$

for other cases Jeffreys' prior usually depends on β.

It should be noticed, however, that there are other models (functional forms) which lead to reference priors not as convenient as (4.2) but still

---

3) Note that $g_J(\eta)$ is not a marginal prior, but a way of notation indicating that $p_J(\beta,\eta)$ does not depend on β.

allowing for analytical integrations with respect to β. Let us consider the following functional form:

$$h(z_t;\ \beta,\eta) = \beta_1 x_1(z_t,\eta) + \beta_2 x_2(z_t) + \ldots + \beta_k x_k(z_t), \qquad 4.3)$$

where $x_0(z_t,\eta) \equiv 0$ and only one x_i (say, x_1) depends on η. In this case only first columns of $\dot{X}_1, \ldots, \dot{X}_q$ are nonzero, so D_2 can be presented as $D_2 = \beta_1 G$, where G consists of those nonzero columns of $\dot{X}_1, \ldots, \dot{X}_q$. Now Jeffreys' prior takes the form

$$\begin{aligned} &p_J(\beta,\eta,\omega) \propto \omega^{-1} |\beta_1|^q\ g_J(\eta), \\ &g_J(\eta) \propto \begin{vmatrix} X_\eta' X_\eta & X_\eta' G \\ G' X_\eta & G'G \end{vmatrix}^{1/2}, \end{aligned} \qquad (4.4)$$

and the corresponding joint posterior PDF takes the form

$$p_J(\beta,\eta,\omega|y,Z) \propto g_J(\eta)\ s_\eta^{-\frac{n-k}{2}} |X_\eta' X_\eta|^{-\frac{1}{2}}\ |\beta_1|^q\ p_N^k\left[\beta|b_\eta,\ \frac{1}{\omega}(X_\eta' X_\eta)^{-1}\right] \times$$

$$\times\ p_\gamma\left[\omega|\frac{n-k}{2},\ \frac{s_\eta}{2}\right].$$

Now it is obvious that

$$p_J(\beta,\eta|y,Z) \propto g_J(\eta)\ s_\eta^{-\frac{n-k}{2}} |X_\eta' X_\eta|^{-\frac{1}{2}}\ |\beta_1|^q\ p_S^k\left[\beta|n-k,\ b_\eta,\ \frac{n-k}{s_\eta}\ X_\eta' X_\eta\right]$$

and that posterior analysis involves higher-order moments of t distribution; see Osiewalski (1987) and Osiewalski and Goryl (1988) for detailed derivations (as well as examples) in some specific cases with $k = 1$ and $q \le 2$.

Obviously, even in the relatively simple situations described above, the form of $p_J(\eta)$ may be so complicated that we do not expect

Jeffreys' prior to become widely used. The practice of using uniform priors for all structural parameters[4], like in H. Tsurumi and Y. Tsurumi (1976), could be justified by Savage's "precise measurement" (or "stable estimation") principle - see DeGroot (1970) - but only when the number of observations is "reasonably large"; now the problem of the form of the prior is replaced by the question whether our sample is large enough to rely on inferences corresponding to the uniform prior. It should be stressed that the choice of some simple prior of the form $p_1(\beta, \eta, \omega)$ as a "noninformative" one may have only practical (convenience) and intuitive justifications.

---

4) When assuming such a prior for one specific parameterization of a given nonlinear model one should be aware of the consequences of reparameterization. For example, if ε ($\varepsilon > 0$) is the elasticity-of-substitution parameter in the CES function (see Subsection 6.2), then the notationally most convenient (and usually used) parameterization is in terms of $\rho = (1-\varepsilon)/\varepsilon = \varepsilon^{-1} - 1$. If we assume, like in Tsurumi and Tsurumi (1976), $p(\rho) = \text{const}$ as a "noninformative" prior, we obtain a rather strange-looking form $p(\varepsilon) \propto \varepsilon^{-2}$, which gives an infinite prior probability that $\varepsilon < a$, but a finite prior probability that $\varepsilon > a$ ($a > 0$, say: $a = 1$). Parameterizing directly in terms of ε and assuming $p(\varepsilon) \propto \varepsilon^{-1}$, like in Sankar (1970), gives much more reasonable expression of prior ignorance; it is equivalent to $p(\rho) \propto (1 + \rho)^{-1}$, which is nonuniform in ρ. This difference in "noninformative" priors may lead to different posterior results in small samples.

5. Posterior and predictive PDF's corresponding to priors in the form of finite mixtures

5.1. Advantages of finite-mixture priors

In this section we allow for expressing prior beliefs about β in such a way that still enables analytical integrations of posterior PDF with respect to β. Of course, normal or t priors of β are the most convenient informative priors from the analytical and numerical point of view. On the other hand, they can prove too restrictive in practice because of their symmetry and unimodality. In order to obtain more flexible (but still convenient) classes of priors, finite mixtures of normal or t distributions seem worth considering.[5] As simple examples show, finite mixtures of univariate normal distributions can produce priors of quite different shapes: multimodal, asymmetric, phatykurtic - even if the number of components of the mixture is very small; some preliminary work on expressing prior beliefs in the form of such mixtures was done by Bijak (1987), but elicitation problems are outside the scope of this paper and need separate considerations. Here we are interested in the form and tractability of posterior and predictive PDF's corresponding to finite-mixture priors. Let us consider the general case first; $l(\delta|\text{data})$ denotes the likelihood function, where $\delta \in \Delta$ is a vector of parameters, and $p(\tilde{y}|\text{data}, \delta)$ denotes the conditional PDF of future observations ($\tilde{y}$) given data and parameters. If $p_g(\delta)$ is the prior density then, obviously, the posterior and predicitve densities are given by

$$p_g(\delta|\text{data}) = K_g^{-1}\, p_g(\delta)\, l(\delta|\text{data}),$$

$$p_g(\tilde{y}|\text{data}) = \int_\Delta p(\tilde{y}|\text{data}, \delta)\, p_g(\delta|\text{data})\, d\delta,$$

where $K_g = \int_\Delta p_g(\delta)\, l(\delta|\text{data})\, d\delta.$

---

5) The idea of taking mixtures as priors comes from Dalal and Hall (1983).

But when the prior is represented by a finite mixture of such $p_g(\delta)$ for $g = 1,\ldots,G$ (with weights c_g which are positive and sum up to 1), that is when

$$p(\delta) = \sum_{g=1}^{G} c_g\, p_g(\delta),$$

then

$$p(\delta|\text{data}) = \frac{p(\delta)\; l(\delta|\text{data})}{\int_\Delta p(\delta)\; l(\delta|\text{data})d\delta} = \frac{\sum_g c_g\, p_g(\delta)\; l(\delta|\text{data})}{\sum_g c_g\, K_g} =$$

$$= \sum_g \bar{c}_g\, p_g(\delta|\text{data}),$$

where

$$\bar{c}_g = c_g\, K_g / \sum_g c_g\, K_g,$$

and

$$p(\tilde{y}|\text{data}) = \int_\Delta p(\tilde{y}|\text{data},\, \delta)\; p(\delta|\text{data}) =$$

$$= \sum_g \bar{c}_g\, p_g(\tilde{y}|\text{data}),$$

thus the "overall" posterior and predictive PDF's are in the form of finite mixtures of "individual" posterior and predictive PDF's. If, for every g, there exist posterior and predictive moments about 0, then the corresponding "overall" moments about 0 are simply weighted averages of "individual" ones; the same property holds for marginal densities. And, what is most important for our purposes, if all "individual" posterior densities are analytically tractable (with respect to some δ_i) then the "overall" posterior is tractable (with respect to this δ_i). In the case considered in the paper we insist on taking advantages of partial linearity of the model, thus we insist on analytical integrability of the posterior PDF with respect to β. This is the reason of our special interest in mixtures of normal or t distributions of β. Assuming mixtures, we can proceed in two equivalent ways: to sum up "individual" results (weighted appropriately) or to derive directly "overall" results (which in the case of $G = 1$ are the same as "individual" ones). We adopt the second approach in the rest of this section.

Finite-mixture priors could be interpreted as representing prior information coming from several different (jointly exhaustive and mutually exclusive) sources. But generally mixtures can be treated merely as an useful approximation of some preassigned shape of prior density. Such an attitude was adopted by Dalal and Hall (1983) and is adopted here as well.

5.2. Mixtures of normal distributions

We assume the following conditional prior density[6)] of β given η:

$$p_2(\beta|\eta) = \sum_{g=1}^{G} c_g\, p_N^k\left[\beta | a_g,\ A_g^{-1}\right],$$

where $G \geq 1$, $c_g > 0$, $\sum_g c_g = 1$, $a_g \in R^k$ and A_g are PDS of order k.
The joint prior density for all parameters takes the form

$$p_2(\beta,\ \eta,\ \omega) = p_2(\beta|\eta)\ p(\eta)\ p(\omega) \propto$$

$$\propto p(\eta)\ \omega^{\frac{e}{2}-1}\ \exp(-\tfrac{f}{2}.\omega)\ \sum_g c_g |A_g|^{\frac{1}{2}} \exp\left[-\tfrac{1}{2}\ (\beta - a_g)'\ A_g(\beta - a_g)\right].$$

For this prior and the likelihood given by (3.2), Bayes' theorem leads to the following joint posterior PDF:

---

6) Our analysis here, as well as in Subsection 5.3, allows for prior dependence between β and η; parameters of $p_i(\beta|\eta)$, $i = 2, 3$, can be some functions of η. Obviously, the special case: $p_i(\beta|\eta) = p_i(\beta)$ seems the most convenient one in practice.

$$p_2(\beta,\eta,\omega|y,Z) \propto p_2(\beta,\eta,\omega)\ p(y|Z,\beta,\eta,\omega) \propto$$

$$\propto p(\eta)\ \omega^{\frac{e+n}{2}-1} \exp\left[-\frac{f+s_\eta}{2}.\omega\right].\sum_g c_g|A_g|^{\frac{1}{2}} \exp\{-\tfrac{1}{2}[(\beta-a_g)'\ A_g(\beta-a_g) +$$

$$+ \omega(\beta-b_\eta)'\ X_\eta'X_\eta(\beta-b_\eta)]\}.$$

For $\bar{A}_g = A_g + \omega X_\eta'X_\eta$ and $\bar{a}_g = \bar{A}_g^{-1}(A_g a_g + \omega X_\eta'X_\eta b_\eta)$ we have

$$(\beta - a_g)'\ A_g(\beta - a_g) + \omega(\beta - b_\eta)'X_\eta'X_\eta(\beta - b_\eta) =$$

$$= (\beta - \bar{a}_g)'\ \bar{A}_g(\beta - \bar{a}_g) + d_g,$$

where

$$d_g = a_g'A_g a_g + \omega b_\eta'X_\eta'X_\eta b_\eta - \bar{a}_g'\bar{A}_g\bar{a}_g; \quad d_g \geq 0.$$

Now the joint posterior PDF can be written as

$$p_2(\beta,\eta,\omega|y,Z) \propto$$

$$\propto p(\eta)\ \omega^{\frac{e+n}{2}-1} \exp\left[-\frac{f+s_\eta}{2}.\omega\right] \sum_g c_g\left|A_g.\bar{A}_g^{-1}\right|^{\frac{1}{2}} \exp\left[-\frac{d_g}{2}\right]\ p_N^k(\beta|\bar{a}_g,\bar{A}_g^{-1}).$$

Let us denote

$$C_g = c_g\left|A_g.\bar{A}^{-1}\right|^{\frac{1}{2}} \exp\left[-\frac{d_g}{2}\right],$$

$$C = \sum_g C_g \quad , \quad \bar{c}_g = C^{-1}C_g;$$

now

$$p_2(\beta,\eta,\omega,|y,Z) \propto p(\eta)\ \omega^{\frac{e+n}{2}-1} \exp\left[-\frac{f+s_\eta}{2}.\omega\right]\ C \sum_g \bar{c}_g\ p_N^k\left[\beta|\bar{a}_g,\ \bar{A}_g^{-1}\right].$$

Since $\bar{c}_g > 0$ and $\sum_g \bar{c}_g = 1$, then

$$\int_{R^k} \sum_g \bar{c}_g\, p_N^k\left[\beta|\bar{a}_g,\ \bar{A}_g^{-1}\right] d\beta = \sum_g \bar{c}_g \int_{R^k} p_N^k\left[\beta|\bar{a}_g,\ \bar{A}_g^{-1}\right] d\beta = 1,$$

$$p_2(\eta,\omega|y,Z) = \int_{R^k} p_2(\beta,\eta,\omega|y,Z)\, d\beta \propto$$

$$\propto p(\eta)\ \omega^{\frac{e+n}{2}-1} \exp\left[-\frac{f+s_\eta}{2}.\omega\right] C,$$

and $p_2(\beta|y,Z,\eta,\omega) = \sum_g \bar{c}_g\, p_N^k\left[\beta|\bar{a}_g,\ \bar{A}_g^{-1}\right]$.

The joint posterior PDF is now expressed as a product of the marginal posterior PDF of (η,ω):

$$p_2(\beta,\eta,\omega|y,Z) = p_2(\eta,\omega|y,Z)\ p_2(\beta|y,Z,\eta,\omega),$$

the latter density being the mixture of k-dimensional normal PDF's. Inferences about η (and ω, if necessary) will be based on the marginal posterior of (η,ω); in order to calculate its normalizing constant, moments and univariate marginal densities numerical integrations will be required. For inferences about β, its marginal posterior is appropriate. The marginal posterior PDF of β can be expressed as the following integral

$$p_2(\beta|y,Z) = \int_H \int_0^\infty p_2(\eta,\omega|y,Z)\ p_2(\beta|y,Z,\eta,\omega)\ d\omega\ d\eta.$$

Since conditional moments and univariate densities are given by known analytical formulae then marginal moments and univariate densities can be calculated as follows:

$$E_2(\beta|y,Z) = \int_H \int_0^\infty p_2(\eta,\omega|y,Z) \sum_g \bar{c}_g \bar{a}_g\ d\omega\ d\eta,$$

$$E_2(\beta\beta'|y,Z) = \int_H \int_0^\infty p_2(\eta,\omega|y,Z) \sum_g \bar{c}_g\left[\bar{A}_g^{-1} + \bar{a}_g.\bar{a}_g'\right] d\omega\ d\eta,$$

$$p_2(\beta_i|y,Z) = \int_H \int_0^\infty p_2(\eta,\omega|y,Z) \sum_g \bar{c}_g \; p_N^1\left[\beta_i \middle| \left[\bar{a}_g\right]_i, \left[\bar{A}_g^{-1}\right]_{ii}\right] d\omega \, d\eta;$$

similarly for mixed moments of β and η:

$$E_2(\eta.\beta'|y,Z) = \int_H \int_0^\infty p_2(\eta,\omega|y,Z).\eta.\sum_g \bar{c}_g \bar{a}_g' \, d\omega \, d\eta.$$

In order to derive the predictive PDF $p_2(\tilde{y}|y,Z,\tilde{Z})$ according to (2.2), let us write is as

$$p_2(\tilde{y}|y,Z,\tilde{Z}) = \int_H \int_0^\infty \int_{R^k} p(\tilde{y}|\tilde{Z},\beta,\eta,\omega) \; p_2(\beta|y,Z,\eta,\omega) \, d\beta \times$$

$$\times \; p_2(\eta,\omega|y,Z) \, d\omega \, d\eta.$$

By analytical integration with respect to β - on the basis of (1.1) - we obtain

$$p_2(\tilde{y}|y,Z,\tilde{Z}) = \int_H \int_0^\infty p_2(\eta,\omega|y,Z) \sum_g \bar{c}_g \; p_N^m\left[\tilde{y} \middle| \tilde{w}_\eta + \tilde{X}_\eta \bar{a}_g, \right.$$

$$\left. \tilde{X}_\eta \bar{A}_g^{-1} \tilde{X}_\eta' + \frac{1}{\omega} I_m\right] d\omega \, d\eta,$$

and it is easy to deduce the formulae for moments and univariate densities of the predictive distribution.

In the case of the prior considered here, Bayesian estimation and prediction requires (q+1)-dimensional numerical integrations, irrespective of k. For practical reasons, only mixtures of a small number of terms are attractive. Let us note that there is no special advantage of the type of prior assumed here over any other type of informative prior if $k = 1$ (β is a scalar parameter). For any $p(\beta,\eta)$, $\beta \in R^k$ and $\eta \in R^q$, we can always integrate ω out analytically and then the dimension of integrals to be calculated is $K = k + q$ (see Section 2); in the case of $p_2(\beta,\eta)$, considered here, we integrate β out analytically and then (q+1)-dimensional numerical integration (over η and ω) remains. When $k = 1$, both approaches lead to different, but (q+1)-dimensional integrals.

5.3. Mixtures of t distributions

Finite mixtures of normal distributions are quite flexible, except for their tail behaviour which is essentially the same as in the case of one normal distribution. In order to obtain fatter tails of the prior distribution of β given η, a finite mixture of k-variate Student t distributions can be applied. This mixture can formally be treated as a marginal distribution from the following mixture of normal-gamma distributions of β and an additional parameter $\tau > 0$:

$$p_3(\beta,\tau|\eta) = \sum_{g=1}^{G} c_g \; p_N^k\left[\beta|a_g, \; (\tau A_g)^{-1}\right] \; p_\gamma\left[\tau|\frac{l_g}{2}, \frac{v_g}{2}\right], \tag{5.1}$$

since the integration with respect to τ leads to

$$p_3(\beta|\eta) = {}_0\!\int^\infty p_3(\beta,\tau|\eta)d\tau = \sum_{g=1}^{G} c_g \; p_S^k\left[\beta|l_g, \; a_g, \; \frac{l_g}{v_g} A_g\right].$$

We adopt (5.1) as a starting point for the derivation of posterior and predictive results corresponding to the prior $p_3(\beta|\eta)$.[7] The joint prior density of β, η, ω and the additional parameter τ is as follows:

$$p_3(\beta,\eta,\omega,\tau) \propto p(\eta) \; \omega^{\frac{e}{2}-1} \; \exp(-\frac{f}{2}.\omega) \; \sum_g c_g \; p_N^k\left[\beta|a_g, (\tau A_g)^{-1}\right] \times$$

$$\times \; p_\gamma\left[\tau|\frac{l_g}{2}, \frac{v_g}{2}\right].$$

For this prior and the likelihood given by (3.2) one obtains the following joint posterior PDF:

---

7) Multiplying $p_3(\beta,\eta,\omega) = p_3(\beta|\eta) \; p(\eta) \; p(\omega)$ by the likelihood (3.2) and integrating ω out, we would obtain a finite mixture of double-t (2-0 poly-t) densities as a conditional posterior PDF of β given η (and some marginal posterior PDF of η). Thus (5.1) enables us to perform analytical integrations with respect to β in the same way as in Dickey (1968).

$$p_3(\beta,\eta,\omega,\tau|y,Z) \propto p_3(\beta,\eta,\omega,\tau)\ p(y|Z,\beta,\eta,\omega) \propto p(\eta)\ \omega^{\frac{e+n}{2}-1} \times$$

$$\times \exp\left[-\frac{f+s_\eta}{2}.\omega\right] \sum_g c_g\ p_\gamma\left[\tau\left|\frac{l_g}{2},\ \frac{v_g}{2}\right.\right] \tau^{\frac{k}{2}}|A_g|^{\frac{1}{2}} \exp\left\{-\frac{\tau}{2}[(\beta-a_g)'A_g(\beta-a_g) +\right.$$

$$\left. + \frac{\omega}{\tau}\ (\beta-b_\eta)'\ X_\eta'X_\eta(\beta-b_\eta)]\right\}.$$

Let us denote $\lambda = \frac{\omega}{\tau}$ and define matrices $\bar{A}_g$, vectors $\bar{a}_g$ and scalars d_g similarly as in the previous subsection (merely replacing ω by λ):

$$\bar{A}_g = A_g + \lambda\ X_\eta'X_\eta \quad , \quad \bar{a}_g = \bar{A}_g^{-1}(A_g a_g + \lambda\ X_\eta'X_\eta b_\eta),$$

$$d_g = a_g'A_g a_g + \lambda\ b_\eta'X_\eta'X_\eta b_\eta - \bar{a}_g'\bar{A}_g\bar{a}_g \qquad (\text{always } d_g \geq 0).$$

Now we can present the joint posterior PDF in the following form:

$$p_3(\beta,\eta,\omega,\tau|y,Z) \propto$$

$$\propto p(\eta)\ \omega^{\frac{e+n}{2}-1} \exp\left[-\frac{f+s_\eta}{2}.\omega\right] \sum_g c_g\left|A_g\bar{A}_g^{-1}\right|^{\frac{1}{2}} p_\gamma\left[\tau\left|\frac{l_g}{2},\ \frac{v_g}{2}\right.\right] \times$$

$$\times \exp\left[-\frac{\tau}{2}\ d_g\right]\ p_N^k\left[\beta|\bar{a}_g,\ (\tau\bar{A}_g)^{-1}\right].$$

After the transformation: $(\omega,\tau) \longrightarrow (\lambda,\tau)$, with the Jacobian equal to τ, one obtains

$$p_3(\beta,\eta,\lambda,\tau|y,Z) \propto p(\eta)\ \lambda^{\frac{e+n}{2}-1}\ \tau^{\frac{e+n}{2}} \exp\left[-\frac{f+s_\eta}{2}\ \lambda\tau\right] \times$$

$$\times \sum_g c_g\left|A_g\bar{A}_g^{-1}\right|^{\frac{1}{2}} p_\gamma\left[\tau\left|\frac{l_g}{2},\ \frac{v_g}{2}\right.\right] \exp\left[-\frac{\tau}{2}\ d_g\right]\ p_N^k\left[\beta|\bar{a}_g,\ \left[\tau\bar{A}_g\right]^{-1}\right].$$

Defining

$$\bar{l}_g = l_g + e + n \quad , \quad \bar{v}_g = v_g + d_g + (f + s_\eta)\lambda,$$

$$C_g = c_g \left| A_g \bar{A}_g^{-1} \right|^{\frac{1}{2}} \left[\Gamma\left(\frac{l_g}{2}\right) \right]^{-1} \Gamma\left(\frac{\bar{l}_g}{2}\right) v_g^{\frac{l_g}{2}} \bar{v}_g^{-\frac{\bar{l}_g}{2}},$$

$$C = C(\eta,\lambda) = \sum_g C_g \quad , \quad \bar{c}_g = C^{-1} C_g$$

one can write

$$p_3(\beta,\eta,\lambda,\tau|y,Z) \propto p(\eta)\, \lambda^{\frac{e+n}{2}-1} \sum_g C_g\, p_\gamma\left(\tau \middle| \frac{\bar{l}_g}{2}, \frac{\bar{v}_g}{2}\right) p_N^k\left(\beta \middle| \bar{a}_g, (\tau\bar{A}_g)^{-1}\right)$$

or

$$p_3(\beta,\eta,\lambda,\tau|y,Z) = p_3(\eta,\lambda|y,Z)\, p_3(\beta,\tau|y,Z,\eta,\lambda),$$

$$p_3(\eta,\lambda|y,Z) \propto p(\eta)\, \lambda^{\frac{e+n}{2}-1} C(\eta,\lambda),$$

$$p_3(\beta,\tau|y,Z,\eta,\lambda) = \sum_g \bar{c}_g\, p_\gamma\left(\tau \middle| \frac{\bar{l}_g}{2}, \frac{\bar{v}_g}{2}\right) p_N^k\left(\beta \middle| \bar{a}_g, (\tau\bar{A}_g)^{-1}\right).$$

Since integrations with respect to β and τ can be performed analytically, the above-presented forms of the joint posterior PDF seem relatively convenient. For estimation purposes, τ should be integrated out in order to obtain

$$p_3(\beta|y,Z,\eta,\lambda) = {}_0\!\int^{\infty} p_3(\beta,\tau|y,Z,\eta,\lambda)d\tau = \sum_g \bar{c}_g\, p_S^k\left(\beta \middle| \bar{l}_g, \bar{a}_g, \frac{\bar{l}_g}{\bar{v}_g}\bar{A}_g\right)$$

and then

$$p_3(\beta,\eta,\lambda|y,Z) = p_3(\beta|y,Z,\eta,\lambda)\, p_3(\eta,\lambda|y,Z).$$

Calculation of the normalizing constant of $p_3(\eta,\lambda|y,Z)$ as well as inferences about β and η (in terms of posterior first- and second-order moments and univariate marginal densities) require numerical integrations in the

(η,λ)-space; for β, the appropriate formulae can be easily derived on the basis of known properties of t distributions.

In order to obtain and analyse the predictive distribution corresponding to t-mixture prior one can rewrite $p(\tilde{y}|\tilde{Z},\beta,\eta,\omega)$ in terms of λ and τ (instead of ω):

$$p(\tilde{y}|\tilde{Z},\beta,\eta,\lambda,\tau) = p(\tilde{y}|\tilde{Z},\beta,\eta,\omega = \lambda\tau) = p_N^m\left[\tilde{y}|\tilde{w}_\eta + \tilde{X}_\eta\beta,\ \frac{1}{\lambda\tau} I_m\right]$$

and then derive analytically, according to (1.1) and (1.3),

$$p_3(\tilde{y}|y,Z,\tilde{Z},\eta,\lambda) = {}_0\!\int^\infty \int_{R^k} p(\tilde{y}|\tilde{Z},\beta,\eta,\lambda,\tau)\ p_3(\beta,\tau|y,Z,\eta,\lambda)\ d\beta\ d\tau =$$

$$= \sum_g \bar{c}_g\ {}_0\!\int^\infty p_\gamma\left[\tau|\frac{\bar{l}_g}{2},\ \frac{\bar{v}_g}{2}\right] \int_{R^k} p_N^m\left[\tilde{y}|\tilde{w}_\eta + \tilde{X}_\eta\beta,\ \frac{1}{\lambda\tau} I_m\right]\ p_N^k\left[\beta|\bar{a}_g,\right.$$

$$\left.(\tau\bar{A}_g)^{-1}\right] d\beta\ d\tau = \sum_g \bar{c}_g\ p_S^m\left[\tilde{y}|\bar{l}_g,\ \tilde{w}_\eta + \tilde{X}_\eta\bar{a}_g,\ \frac{\bar{l}_g}{\bar{v}_g}\left[\tilde{X}_\eta\bar{A}_g^{-1}\tilde{X}'_\eta + \frac{1}{\lambda} I_m\right]^{-1}\right].$$

Now it is obvious that the analysis of the predictive PDF

$$p_3(\tilde{y}|y,Z,\tilde{Z}) = \int_H {}_0\!\int^\infty p_3(\tilde{y}|y,Z,\tilde{Z},\eta,\lambda)\ p_3(\eta,\lambda|y,Z)\ d\lambda\ d\eta$$

exploits known properties of t distributions and requires (q+1)-dimensional numerical integrations, irrespective of k and m (the dimensions of β and $\tilde{y}$, respectively).

6. Concluding remarks and comments on applications

6.1. Discussion of the results

Let us treat the model under consideration, that is

$$y_t = x_0(z_t,\eta) + \beta_1 x_1(z_t,\eta) + \ldots + \beta_k x_k(z_t,\eta) + u_t,$$

$$u_t \sim iiN(0,\sigma^2),\ \eta \in H \subset R^q \quad , \quad \beta = (\beta_1,\ldots,\beta_k)' \in R^k,$$

not as a special case of the nonlinear regression, but as an useful general representation of linear and nonlinear regression models. In order to achieve this generality, we allow for $q = 0$ (η does not exist) or $k = 0$ (β does not exist), but with obvious restriction that $k + q \geq 1$ (there exists at least one unknown parameter). The following situations are possible:

(i) $q = 0$; the model is linear in all its parameters. From Subsections 4.1, 5.2 and 5.3, conditional posterior and predictive results (given η) remain valid; of course part of them (Subsection 4.1) are standard and well-known, but the use of finite-mixture priors seems new even in the linear context.

(ii) $k = 0$; there is no possibility to represent a given nonlinear model in the "partly linear" form. We have a "completely nonlinear" model: $\vartheta = \eta$, $h(z_t,\vartheta) = x_0(z_t,\eta)$, and we are back in Section 2 with Bayesian estimation and prediction based on (2.3) and (2.4).

(iii) $k > 0$ and $q > 0$; we have the partly linear regression model considered in Sections 3-5.

The main conclusion is that exact Bayesian analysis is possible if q is small, irrespective of k; of course the meaning of "small" is not precise and depends on computer facilities.
If, additionally, $k = 0$ or $k = 1$, then at most $(q+1)$-dimensional numerical integrations are required irrespective of the choice of the prior density

$p(\beta,\eta)$. However, if $k > 1$, then the classes of prior densities adopted in the paper seem specially attractive, since they are flexible and always lead to at most $(q+1)$-dimensional numerical integration.
If q is large (too large to perform integrations numerically) then integration by Monte Carlo methods or approximations are required. But - if only $k > 0$ - there are still some advantages of the proposed classes of priors, since we have exact analytical posterior and predictive results conditionally on η or (η,w) or (η,λ).

6.2. Applications to CES production functions

Let us point at some new possibilities in this case, where the model under consideration takes the form

$$V_t = \left[\delta C_t^{(\varepsilon-1)/\varepsilon} + (1-\delta)L_t^{(\varepsilon-1)/\varepsilon}\right]^{\nu\varepsilon/(\varepsilon-1)} \exp(\beta_2 x_{t2} + \dots + \beta_k x_{tk} + $$

$$+ u_t);$$

$x_{t2}, \dots, x_{tk}$ can be dummy, time or other variables. Taking the logarithms of both sides, we obtain the partly linear regression model

$$y_t = \beta_1 x_{t1}(\eta) + \beta_2 x_{t2} + \dots + \beta_k x_{tk} + u_t, \qquad (6.1)$$

where

$$y_t = \ln V_t,\ \beta_1 = \nu,\ \eta = (\delta\ \varepsilon)' \in (0,1) \times (0, +\infty),$$

$$x_{t1}(\eta) = \begin{cases} \frac{\varepsilon}{\varepsilon-1} \ln\left[\delta\ C_t^{(\varepsilon-1)/\varepsilon} + (1-\delta)\ L_t^{(\varepsilon-1)/\varepsilon}\right], & \varepsilon \neq 1, \\ \delta \ln C_t + (1-\delta) \ln L_t, & \varepsilon = 1; \end{cases}$$

the Cobb-Douglas case ($\varepsilon = 1$) is included in order to preserve continuity with respect to ε, the elasticity-of-substitution parameter. First, it is possible to incorporate prior knowledge in the forms proposed in Section 5 and trivariate numerical integration is sufficient to obtain exact

Bayesian results; previous Bayesian analyses of CES functions used only simple priors, uniform in β.
Second, (6.1) is in the form (4.3), so Jeffreys' (reference) prior is

$$p_J(\beta,\eta,\sigma) \propto \sigma^{-1}\ \beta_1^2\ g_J(\eta) = \sigma^{-1}\ \nu^2\ g_J(\eta),$$

where $g_J(\eta)$ is in the form (4.4). Now it is possible to make comparisons between results corresponding to Jeffreys' prior and to simpler (intuitively noninformative) prior used in the literature.

6.3. Logistic curves

The approach developed in the paper enables the exact Bayesian analysis of the following generalizations of the logistic growth curve:

$$Y_t = \frac{\exp(\beta_1 x_{t1} + \dots + \beta_k x_{tk})}{1 + \eta_2 \exp(-\eta_1 \cdot v_t)} \exp(u_t), \tag{6.2}$$

or

$$Y_t = \beta_0 + \frac{\beta_1 x_{t1} + \dots + \beta_k x_{tk}}{1 + \eta_2 \exp(-\eta_1 \cdot v_t)} + u_t, \tag{6.3}$$

where $\eta_1, \eta_2 > 0$, $u_t \sim iiN(0,\sigma^2)$ and $x_{t1}, \dots, x_{tk}, v_t$ are some explanatory variables. The simplest special cases of (6.2) and (6.3), that is

$$Y_t = \frac{\exp(\beta_1)}{1 + \eta_2 \exp(-\eta_1 \cdot t)} \exp(u_t) \quad , \tag{6.4}$$

$$Y_t = \frac{\beta_1}{1 + \eta_2 \exp(-\eta_1 \cdot t)} + u_t \quad , \tag{6.5}$$

were analyzed under Jeffreys' prior by Osiewalski and Goryl (1986,1988). For (6.4) and several sets of data, two other priors (intuitively "noninformative") were also adopted and led to almost the same posteriors as Jeffreys' prior did; marginal posterior densities were always unimodal, almost symmetric or slightly asymmetric and exact Bayesian results were close to ML (that is, to Bayesian approximate large-sample) results. On the other hand, another example - with (6.5) and Jeffreys' prior - revealed large discrepancies between the exact posterior mean and standard deviation of η_2 and their approximate (ML) counterparts. Since the marginal posterior density of η_2 had a lognormal-like shape, it was easy to propose a "better" parameterization of (6.5), namely in terms of $\eta_0 = \ln \eta_2$, which led to closer exact and approximate (ML) results. For a Bayesian, if only an exact analysis is possible - and it is possible for such models as (6.2) and (6.3) - then there is no fundamental need to seek such a "better" parameterization, even when it is easy to find by inspection of marginal posterior densities. This contrasts with the classical approach to nonlinear regression models, where a "good" parameterization is crucial to rely on ML results in a small sample, and it proves difficult to find such a parameterization; see Ratkowsky (1983), where the classical estimation of logistic functions and many other nonlinear models is presented.

6.4. A generalization to nonscalar error covariance matrices

Let us consider the case when the disturbances of our partly linear model have a nonscalar covariance matrix.
We assume the following model:

$$\begin{bmatrix} y \\ \tilde{y} \end{bmatrix} = \begin{bmatrix} w_\eta \\ \tilde{w}_\eta \end{bmatrix} + \begin{bmatrix} X_\eta \\ \tilde{X}_\eta \end{bmatrix} \beta + \begin{bmatrix} u \\ \tilde{u} \end{bmatrix},$$

$$\begin{bmatrix} u \\ \tilde{u} \end{bmatrix} \sim N\left(0, \frac{1}{\omega} \begin{bmatrix} V_\varphi & \bar{V}_\varphi \\ \bar{V}'_\varphi & \tilde{V}_\varphi \end{bmatrix}\right),$$

where $\varphi \in \Phi$ is an additional unknown parameter (vector) and V_φ, $\bar{V}_\varphi$, $\tilde{V}_\varphi$ are known functions of φ. For example, when the disturbances are described by the normal stationary AR(1) process

$$u_t = \varphi\, u_{t-1} + \varepsilon_t \quad , \quad \varepsilon_t \sim iiN(0, \omega^{-1}),$$

then φ is a scalar parameter, $\varphi \in (-1,1)$ and the covariance matrix of $(u'\tilde{u}')'$ takes the well-known form, namely

$$\operatorname{cov}(u_t, u_{t'}) = \omega^{-1}(1-\varphi^2)^{-1}\, \varphi^{|t-t'|},$$

$$t,\ t' = 1,\ \dots,\ n,\ n+1,\ \dots,\ n+m.$$

Now, under a nonscalar covariance matrix, $\tilde{y}$ and y may be stochastically dependent (if $\bar{V}_\varphi \neq 0$) and the following factorization holds

$$p(y,\tilde{y}|Z,\tilde{Z},\beta,\eta,\omega,\varphi) = p(y|Z,\beta,\eta,\omega,\varphi)\ p(\tilde{y}|y,Z,\tilde{Z},\beta,\eta,\omega,\varphi),$$

where

$$p(y|Z,\beta,\eta,\omega,\varphi) = p_N^n(y|w_\eta + X_\eta\beta,\ \omega^{-1}\ V_\varphi) =$$

$$= (2\pi)^{-\frac{n}{2}}\ \omega^{\frac{n}{2}}\ |V_\varphi|^{-\frac{1}{2}}\ \exp\left\{-\frac{\omega}{2}\left[s_{\eta,\varphi} + (\beta - b_{\eta,\varphi})'\ X_\eta' V_\varphi^{-1} X_\eta (\beta - b_{\eta,\varphi})\right]\right\},$$

$$p(\tilde{y}|y,Z,\tilde{Z},\beta,\eta,w,\varphi) = p_N^m\left[\tilde{y}|Q_{\eta,\varphi}\beta + \bar{V}_\varphi' V_\varphi^{-1}(y-w_\eta) + \tilde{w}_\eta,\ \omega^{-1} S_\varphi\right]$$

and

$$b_{\eta,\varphi} = \left[X_\eta' V_\varphi^{-1} X_\eta\right]^{-1}\ X_\eta' V_\varphi^{-1}(y - w_\eta),$$

$$s_{\eta,\varphi} = (y - w_\eta - X_\eta b_{\eta,\varphi})'\ V_\varphi^{-1}(y - w_\eta - X_\eta b_{\eta,\varphi}),$$

$$Q_{n,\varphi} = \tilde{X}_\eta - \bar{V}_\varphi' V_\varphi^{-1} X_\eta\quad,\quad S_\varphi = \tilde{V}_\varphi - \bar{V}_\varphi' V_\varphi^{-1}\bar{V}_\varphi.$$

Assuming the following prior structure

$$p(\beta,\eta,\omega,\varphi) = p(\beta,\eta)\ p(\omega)\ p(\varphi),\ \beta \in R^k,\ \eta \in H,\ \varphi \in \Phi,\ \omega \in R_+,$$

where $p(\varphi)$ is a marginal prior of φ and, as previously,

$$p(\omega) \propto \omega^{\frac{e}{2}-1}\ \exp\left[-\frac{f}{2}.\omega\right],$$

we can proceed in a similar way as in Subsections 4.1, 5.2, 5.3. For example, when the prior is uniform in β, that is when

$$p(\beta,\eta) = p_1(\beta,\eta) \propto g(\eta),$$

we obtain

$$p_1(\eta,\varphi|y,Z) \propto g(\eta)\ p(\varphi)\ \left[|V_\varphi|\,.\,|X_\eta' V_\varphi^{-1} X_\eta|\right]^{-\frac{1}{2}} (f + s_{\eta,\varphi})^{-\frac{1}{2}(e+n-k)}$$

$$p_1(\beta|y,Z,\eta,\varphi) = p_s^k\left[\beta\Big|e + n - k,\ b_{\eta,\varphi},\ \frac{e + n - k}{f + s_{\eta,\varphi}}\ X_\eta' V_\varphi^{-1} X_\eta\right],$$

$$p_1(\tilde{y}|y,Z,\tilde{Z}) = \int_\Phi \int_H p_1(\eta,\varphi|y,Z)\ p_s^m\Bigg[\tilde{y}\Big|e + n - k,\ Q_{\eta,\varphi} b_{\eta,\varphi} +$$

$$+ \bar{V}_\varphi' V_\varphi^{-1}(y-w_\eta) + \tilde{w}_\eta,\ \frac{e + n - k}{f + s_{\eta,\varphi}}\left[Q_{\eta,\varphi}\left[X_\eta' V_\varphi^{-1} X_\eta\right]^{-1} Q_{\eta,\varphi}' + S_\varphi\right]^{-1}\Bigg] d\eta\ d\varphi;$$

numerical integrations with respect to η and φ will usually be needed in order to obtain a normalizing constant, first- and second - order moments and univariate marginal densities of the posterior and predictive distributions. This increase in dimensionality of calculated integrals constitutes the price for unknown φ in the error covariance matrix.[8] Of course, this exact Bayesian approach is applicable when the matrix V_φ^{-1} has known analytical form (as a function of φ), since numerical inversions of the n×n matrix V_φ for every φ seem impractical.

---

8) See also Richard (1977) where the case of the linear model with autoregressive disturbances is considered. Richard's approach (allowing for nonstationary processes) can be easily generalized to partly linear models.

REFERENCES

Bernardo J.M. (1979), "Reference posterior distributions for Bayesian inference" J.R. Stat. Soc. B 41, 113-147.

Bijak W. (1987), chapter 6 in: "Bayesowska analiza modeli jednorownaniowych z autokorelacja lub nieliniowoscia", workshop paper of the Institute of Cybernetics and Management, Central School of Planning and Statistics, Warszawa,

Box G.E.P. and G.C. Tiao (1973), Bayesian Inference in Statistical Analysis, Addison-Wesley, Reading.

Broemeling L.D. (1985), Bayesian Analysis of Linear Models, Marcel Dekker, New York.

Dalal S.R. and W.J. Hall (1983), "Approximating priors by mixtures of natural conjugate priors", J.R. Stat. Soc. B 45, 278-286.

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

Dickey J.M. (1968), "Three multidimensional-integral identities with Bayesian applications", Ann. of Math. Stat. 39, 1615-1628.

Eaves D.M. (1983), "On Bayesian nonlinear regression with an enzyme example", Biometrika 70, 373-379.

Engle R.F., D.F. Hendry and J.-F.Richard (1983), "Exogeneity", Econometrica 51, 277-304.

Harkema R. and S. Schim van der Loeff (1977), "On Bayesian and non-Bayesian estimation of a two-level CES production function for the Dutch manufacturing sector", J. of Econometrics 5, 155-165.

Osiewalski J. (1987), "Wnioskowanie bayesowskie o parametrach krzywych Törnquista", Zeszyty Naukowe Akademii Ekonomicznej w Krakowie nr 246, 5-30.

Osiewalski J. and A. Goryl (1986), "Estymacja bayesowska parametrow trendu logistycznego", Przeglad Statystyczny 33, 267-285.

Osiewalski J. and A Goryl (1988), "Trend Logistyczny z addytywnym skladnikiem losowym -analiza bayesowska", Prace Naukowe Akademii Ekonomicznej w Katowicach (forthcoming).

Ratkowsky D.A. (1983), Nonlinear Regression Modeling, Marcel Dekker, New York.

Richard J.-F. (1977), "Bayesian analysis of the regression model when the disturbances are generated by an autoregressive process" [in:] New Developments in the Application of Bayesian Methods, ed. by A. Aykaç and C. Brumat, North-Holland, Amsterdam.

Sankar U. (1970), "Elasticities of substitution and returns to scale in Indian manufacturing industries", Review of Economic Studies 11, 399-411.

Tsurumi H. and Y. Tsurumi (1976), "A Bayesian estimation of macro and micro CES production functions", J. of Econometrics 4, 1-25.

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

IN 1987 REEDS VERSCHENEN

257 P.H.M. Ruys, G. van der Laan
Computation of an industrial equilibrium

258 W.H. Haemers, A.E. Brouwer
Association schemes

259 G.J.M. van den Boom
Some modifications and applications of Rubinstein's perfect equilibrium model of bargaining

260 A.W.A. Boot, A.V. Thakor, G.F. Udell
Competition, Risk Neutrality and Loan Commitments

261 A.W.A. Boot, A.V. Thakor, G.F. Udell
Collateral and Borrower Risk

262 A. Kapteyn, I. Woittiez
Preference Interdependence and Habit Formation in Family Labor Supply

263 B. Bettonvil
A formal description of discrete event dynamic systems including perturbation analysis

264 Sylvester C.W. Eijffinger
A monthly model for the monetary policy in the Netherlands

265 F. van der Ploeg, A.J. de Zeeuw
Conflict over arms accumulation in market and command economies

266 F. van der Ploeg, A.J. de Zeeuw
Perfect equilibrium in a model of competitive arms accumulation

267 Aart de Zeeuw
Inflation and reputation: comment

268 A.J. de Zeeuw, F. van der Ploeg
Difference games and policy evaluation: a conceptual framework

269 Frederick van der Ploeg
Rationing in open economy and dynamic macroeconomics: a survey

270 G. van der Laan and A.J.J. Talman
Computing economic equilibria by variable dimension algorithms: state of the art

271 C.A.J.M. Dirven and A.J.J. Talman
A simplicial algorithm for finding equilibria in economies with linear production technologies

272 Th.E. Nijman and F.C. Palm
Consistent estimation of regression models with incompletely observed exogenous variables

273 Th.E. Nijman and F.C. Palm
Predictive accuracy gain from disaggregate sampling in arima - models

274 Raymond H.J.M. Gradus
The net present value of governmental policy: a possible way to find the Stackelberg solutions

275 Jack P.C. Kleijnen
A DSS for production planning: a case study including simulation and optimization

276 A.M.H. Gerards
A short proof of Tutte's characterization of totally unimodular matrices

277 Th. van de Klundert and F. van der Ploeg
Wage rigidity and capital mobility in an optimizing model of a small open economy

278 Peter M. Kort
The net present value in dynamic models of the firm

279 Th. van de Klundert
A Macroeconomic Two-Country Model with Price-Discriminating Monopolists

280 Arnoud Boot and Anjan V. Thakor
Dynamic equilibrium in a competitive credit market: intertemporal contracting as insurance against rationing

281 Arnoud Boot and Anjan V. Thakor
Appendix: "Dynamic equilibrium in a competitive credit market: intertemporal contracting as insurance against rationing

282 Arnoud Boot, Anjan V. Thakor and Gregory F. Udell
Credible commitments, contract enforcement problems and banks: intermediation as credibility assurance

283 Eduard Ponds
Wage bargaining and business cycles a Goodwin-Nash model

284 Prof.Dr. hab. Stefan Mynarski
The mechanism of restoring equilibrium and stability in polish market

285 P. Meulendijks
An exercise in welfare economics (II)

286 S. Jørgensen, P.M. Kort, G.J.C.Th. van Schijndel
Optimal investment, financing and dividends: a Stackelberg differential game

287 E. Nijssen, W. Reijnders
Privatisering en commercialisering; een oriëntatie ten aanzien van verzelfstandiging

288 C.B. Mulder
Inefficiency of automatically linking unemployment benefits to private sector wage rates

289 M.H.C. Paardekooper
A Quadratically convergent parallel Jacobi process for almost diagonal matrices with distinct eigenvalues

290 Pieter H.M. Ruys
Industries with private and public enterprises

291 J.J.A. Moors & J.C. van Houwelingen
Estimation of linear models with inequality restrictions

292 Arthur van Soest, Peter Kooreman
Vakantiebestemming en -bestedingen

293 Rob Alessie, Raymond Gradus, Bertrand Melenberg
The problem of not observing small expenditures in a consumer expenditure survey

294 F. Boekema, L. Oerlemans, A.J. Hendriks
Kansrijkheid en economische potentie: Top-down en bottom-up analyses

295 Rob Alessie, Bertrand Melenberg, Guglielmo Weber
Consumption, Leisure and Earnings-Related Liquidity Constraints: A Note

296 Arthur van Soest, Peter Kooreman
Estimation of the indirect translog demand system with binding non-negativity constraints

IN 1988 REEDS VERSCHENEN

297 Bert Bettonvil
Factor screening by sequential bifurcation

298 Robert P. Gilles
On perfect competition in an economy with a coalitional structure

299 Willem Selen, Ruud M. Heuts
Capacitated Lot-Size Production Planning in Process Industry

300 J. Kriens, J.Th. van Lieshout
Notes on the Markowitz portfolio selection method

301 Bert Bettonvil, Jack P.C. Kleijnen
Measurement scales and resolution IV designs: a note

302 Theo Nijman, Marno Verbeek
Estimation of time dependent parameters in lineair models
using cross sections, panels or both

303 Raymond H.J.M. Gradus
A differential game between government and firms: a non-cooperative approach

304 Leo W.G. Strijbosch, Ronald J.M.M. Does
Comparison of bias-reducing methods for estimating the parameter in dilution series

305 Drs. W.J. Reijnders, Drs. W.F. Verstappen
Strategische bespiegelingen betreffende het Nederlandse kwaliteits-concept

306 J.P.C. Kleijnen, J. Kriens, H. Timmermans and H. Van den Wildenberg
Regression sampling in statistical auditing

307 Isolde Woittiez, Arie Kapteyn
A Model of Job Choice, Labour Supply and Wages

308 Jack P.C. Kleijnen
Simulation and optimization in production planning: A case study

309 Robert P. Gilles and Pieter H.M. Ruys
Relational constraints in coalition formation

310 Drs. H. Leo Theuns
Determinanten van de vraag naar vakantiereizen: een verkenning van materiële en immateriële factoren

311 Peter M. Kort
Dynamic Firm Behaviour within an Uncertain Environment

312 J.P.C. Blanc
A numerical approach to cyclic-service queueing models

313 Drs. N.J. de Beer, Drs. A.M. van Nunen, Drs. M.O. Nijkamp
Does Morkmon Matter?

314 Th. van de Klundert
Wage differentials and employment in a two-sector model with a dual labour market

315 Aart de Zeeuw, Fons Groot, Cees Withagen
On Credible Optimal Tax Rate Policies

316 Christian B. Mulder
Wage moderating effects of corporatism
Decentralized versus centralized wage setting in a union, firm, government context

317 Jörg Glombowski, Michael Krüger
A short-period Goodwin growth cycle

318 Theo Nijman, Marno Verbeek, Arthur van Soest
The optimal design of rotating panels in a simple analysis of variance model

319 Drs. S.V. Hannema, Drs. P.A.M. Versteijne
De toepassing en toekomst van public private partnership's bij de grote en middelgrote Nederlandse gemeenten

320 Th. van de Klundert
Wage Rigidity, Capital Accumulation and Unemployment in a Small Open Economy

321 M.H.C. Paardekooper
An upper and a lower bound for the distance of a manifold to a nearby point

322 Th. ten Raa, F. van der Ploeg
A statistical approach to the problem of negatives in input-output analysis

323 P. Kooreman
Household Labor Force Participation as a Cooperative Game; an Empirical Model

324 A.B.T.M. van Schaik
Persistent Unemployment and Long Run Growth

325 Dr. F.W.M. Boekema, Drs. L.A.G. Oerlemans
De lokale produktiestructuur doorgelicht.
Bedrijfstakverkenningen ten behoeve van regionaal-economisch onderzoek

326 J.P.C. Kleijnen, J. Kriens, M.C.H.M. Lafleur, J.H.F. Pardoel
Sampling for quality inspection and correction: AOQL performance criteria