Adaptive quadrature for likelihood inference on dynamic latent variable models for time-series and panel data

Cagnone, Silvia and Bartolucci, Francesco

Department of Statistical Sciences, University of Bologna, Department of Economics, Finance and Statistics, University of Perugia

29 October 2013

Adaptive quadrature for likelihood inference on dynamic latent variable models for time-series and panel data

Silvia Cagnone
Department of Statistical Sciences
University of Bologna (IT)
email: silvia.cagnone@unibo.it

Francesco Bartolucci
Department of Economics, Finance and Statistics
University of Perugia (IT)
email: bart@stat.unipg.it

October 29, 2013

Abstract

Maximum likelihood estimation of dynamic latent variable models requires to solve integrals that are not analytically tractable. Numerical approximations represent a possible solution to this problem. We propose to use the Adaptive Gaussian-Hermite (AGH) numerical quadrature approximation for a class of dynamic latent variable models for time-series and panel data. These models are based on continuous time-varying latent variables which follow an autoregressive process of order 1, AR(1). Two examples of such models are the stochastic volatility models for the analysis of financial time-series and the limited dependent variable models for the analysis of panel data. A comparison between the performance of AGH methods and alternative approximation methods proposed in the literature is carried out by simulation. Examples on real data are also used to illustrate the proposed approach.

KEYWORDS: AR(1); categorical longitudinal data; Gaussian-Hermite quadrature; limited dependent variable models; stochastic volatility model

JEL CODES: C13, C32, C33

1 Introduction

Statistical models for the analysis of time-series and panel data often involve continuous time-varying latent variables. The dynamics of the latent variables is typically modeled by assuming that these variables follow an autoregressive process of order 1, denoted as usual by AR(1). In the analysis of financial time-series data, this approach is adopted to study the volatility of returns that is heteroscedastic and autocorrelated. In particular, the Stochastic Volatility (SV) model assumes that volatility is a latent variable modeled over time by means of a first-order autoregressive process (Andersen, 1994; Taylor, 1994), as opposed to ARCH-type models according to which the present volatility is affected by past observations through a deterministic function (Bollerslev, 1986).

A different example of data that may be effectively analyzed by means of latent autoregressive models is that of panel data in which repeated observations on the same units are available. Here the aim of the analysis is to account for the non-observable heterogeneity between individuals. Within the variety of dynamic models discussed in the literature for panel data we consider, in particular, Limited Dependent Variable (LDV) models for discrete data, which are very common in the social sciences (Maddala, 1983).

Both SV and LDV models present some attractive features. They allow to properly capture the variability present in the data through an autoregressive latent structure and at the same time they are more parsimonious than other models based on Latent Markov chains (Bartolucci et al., 2013b). Moreover they admit a common general representation in the non-linear state space framework. Nevertheless, the estimation procedure of these models presents some computational difficulties related to the presence of the time-varying latent variables. These variables have to be integrated out from the likelihood function and an analytical solution for this does not exist. In the literature different solutions to this problem have been proposed.

As for SV models, a simple and easy method to be implemented is the generalized method of moments (Taylor, 1986). Harvey et al. (1994) proposed a quasi maximum likelihood approach based on the Kalman filter that has the advantage of not depending on any distribution of the error terms. In the Bayesian context the Monte Carlo Markov Chain techniques have been widely applied. One of the main references is Jacquier et al. (1994). Fridman and Harris (1998) proposed a direct maximum likelihood estimation of SV models using a non-linear Kalman filter algorithm. This algorithm is based on expressing the likelihood function as a nested sequence of uni-dimensional integrals approximated by the Gauss Legendre numerical quadrature. Bartolucci and De Luca (2001, 2003) extended this approach by computing analytical first and second derivatives of the approximated likelihood. They applied a rectangular quadrature to approximate the integrals. A non-linear Kalman filter algorithm has been discussed, among the others, by Tanizaki and Mariano (1998) and Durbin and Koopman (2002) in the general state space model framework. For the same SV model, Junji and Yoshihiko (2005) proposed a solution based on the Laplace approximation. All these studies show that the performance of these approximation methods is highly sensitive to the values of the model parameters.

For LDV models, a numerical integration solution has been proposed by Heiss (2008). He proposed the use of a non-linear filter algorithm and approximated the uni-dimensional integrals through the Gauss-Hermite (GH) quadrature. He showed the higher performance of the estimation based on this quadrature compared with other estimation methods. However, it is known that the GH based methods guarantee accurate parameter estimates when several quadrature points are used per each dimension. We also have to consider that, for certain target functions, instability problems arise in the phase of maximization when the function is approximated through this method.

In this work, we propose to use the Adaptive Gaussian Hermite (AGH) quadrature method (Naylor and Smith, 1982; Liu and Pierce, 1994) to approximate the uni-dimensional integrals involved in the non-linear filtering algorithm used for the estimation of SV and LDV models, when these models are specified within a general state space framework based on an AR(1) latent process. In the literature, AGH have been compared with other methods in a variety of random effect and latent variable models (see, among the others, Pinheiro and Bates, 1995; Rabe-Hesketh et al., 2002; Joe, 2008; Cagnone and Monari, 2013). In all these studies this numerical method appear to be superior to other quadrature methods, as it requires only few quadrature points to get accurate estimates, and it does not risk to miss the maximum, as it well captures the peak of the integrand involved in the likelihood function.

To our knowledge, AGH has not been previously applied for the estimation of dynamic latent variable models within the state space framework. In order to evaluate its performance for this class of models, we perform a wide simulation study under different conditions of study. In particular, for SV models AGH is compared with alternative numerical approximations discussed in the literature as well as with the Laplace approximation because it can be viewed as a particular case of AGH when one quadrature point is used (Pinheiro and Bates, 1995). As for LDV, AGH quadrature is compared with the classical GH quadrature under different conditions.

The paper is organized as follows. In Section 2 we define the class of models of interest and, as particular cases, we describe the SV and LDV models. In Section 3 model estimation is discussed with particular attention to the proposed AGH approximation. Section 4 reports the results of the simulation study for both SV and LDV models. In Section 5 two applications on real data are illustrated. The conclusions are given in Section 6.

2 State space specification of dynamic latent variable models

In the time-series context, we let y_t be the response variable observed at time t with $t = 1, \ldots, T$. For the case of panel data, in which we observe n sample units at T occasions, we extend this notation by denoting the response variable for unit i at occasion t by y_{it}, $i = 1, \ldots, n$, $t = 1, \ldots, T$, and, since in this case covariates are also typically observed, we denote by $\boldsymbol{x}_{it}$ the vector of the covariates corresponding to y_{it}.

In order to formulate dynamic latent variable models in the state space framework, we start from the more general case of panel data considering that time-series is a particular case with $n = 1$. The proposed formulation is based on the following equations for $i = 1, \ldots, n$ and $t = 1, \ldots, T$:

$$
\begin{aligned}
y_{it} &= G(y^*_{it}), \\
y^*_{it} &= h(\alpha_{it}, \boldsymbol{x}_{it}, \varepsilon_{it}), &\quad (1)\\
\alpha_{it} &= m(\alpha_{i,t-1}, \eta_{it}), &\quad (2)
\end{aligned}
$$

where y^*_{it} is a continuous unobservable variable underlying y_{it} and $G(\cdot)$ is a parametric function, the specification of which depends on the nature of the observed variable. Moreover, α_{it} is a time dependent latent variable and ε_{it} and η_{it} are error terms assumed to be mutually independent. We refer to (1) as measurement equation and to (2) as transition equation of the model. The former specifies the relationship between the manifest variables and the latent variables; the latter specifies the dependence of the latent variables over time. When both equations are linear, we obtain the classical linear state space models and the Kalman filter may be used for model estimation. If one or both the equations are non-linear, we obtain non-linear state space models.

In the following we illustrate two particular cases of non-linear state space models: the Stochastic Volatility (SV) model for analyzing financial time-series (Taylor, 1986) and Limited Dependent Variable (LDV) models for panel data (Maddala, 1983).

2.1 Stochastic volatility models for financial time-series

SV models are widely used in the analysis of the volatility in financial time-series. They assume that the volatility is a latent variable following a first-order autoregressive process and, for this aim, they admit the following non-linear state space representation for $t = 1, \ldots, T$:

$$
\begin{aligned}
y_t &= \exp\left(\frac{1}{2}\alpha_t\right)\varepsilon_t, \quad \varepsilon_t \sim N(0,1), \\
\alpha_t &= \gamma + \alpha_{t-1}\rho + \eta_t, \quad \eta_t \sim N(0, \sigma_\eta^2).
\end{aligned}
$$

In this case, $G(\cdot)$ is the identity function since the response variable is continuous. Moreover, $\exp(\alpha_t)$ is the volatility level underlying y_t; hence, the logarithm of the volatility is assumed to follow an AR(1) process.

2.2 Limited dependent variable models for panel data

An alternative class of models belonging to the general framework of non-linear state space models based on equations (1) and (2) are LDV models for the analysis of panel data. Panel data consist of repeated observations on the same individuals, or more generally statistical units, over time. In this context, the latent variables are time and unit dependent random intercepts that allow us to account for the unobserved heterogeneity between subjects. The non-linear state space representation of these models is given by

$$\begin{aligned} y_{it} &= G(y_{it}^*), \\ y_{it}^* &= \alpha_{it} + \boldsymbol{x}_{it}'\boldsymbol{\beta} + \varepsilon_{it}, \\ \alpha_{it} &= \alpha_{i,t-1}\rho + \eta_{it}, \quad \eta_{it} \sim N(0, \sigma_\eta^2), \end{aligned}$$

where different distributions may be assumed for the error terms ε_{it}.

In LDV models for panel data, the response variable y_{it} is typically discrete so that $G(\cdot)$ is not the identity function. In the binary case $G(y_{it}^*) = I(y_{it}^* > 0)$, where I is the indicator function getting

$$\lambda[p(y_{it} = 1|\alpha_{it}, \boldsymbol{x}_{it})] = \alpha_{it} + \boldsymbol{x}_{it}'\boldsymbol{\beta}, \quad i = 1, \ldots, n, \; t = 1, \ldots, T,$$

where $\lambda(\cdot)$ is the logit function or the inverse standard normal cumulative function, resulting in a logit model or a probit model, respectively.

If the response variables are ordinal with J categories, we define a set of thresholds $\tau_1 \leq \ldots \leq \tau_{J-1}$, such that

$$G(y_{it}^*) = j \Leftrightarrow \tau_{j-1} < y_{it}^* \leq \tau_j, \qquad j = 1, \ldots, J,$$

with $\tau_0 = -\infty$ and $\tau_J = +\infty$. Different parameterizations can be considered for ordinal observed variables. A very common one is given by the following proportional odds model for cumulative probability functions (McCullagh, 1980):

$$\log \frac{p(y_{it} \leq j|\alpha_{it}, \boldsymbol{x}_{it})}{p(y_{it} > j|\alpha_{it}, \boldsymbol{x}_{it})} = \tau_j - \alpha_{it} - \boldsymbol{x}_{it}'\boldsymbol{\beta}, \quad i = 1, \ldots, n, \; j = 1, \ldots, J-1, \; t = 1, \ldots, T. \quad (3)$$

3 Model estimation

Estimation of the models illustrated above may be carried out by the the maximum likelihood method, which is based on the maximization of

$$L(\boldsymbol{\theta}) = \prod_{i=1}^{n} f(y_{i1}, \dots, y_{iT}), \tag{4}$$

where $\boldsymbol{\theta}$ is the vector of all model parameters which affects the manifest joint distribution of the observed variables. The latter can be expressed as

$$f(y_{i1}, \dots, y_{iT}) = \int \cdots \int \prod_{t=1}^{T} f_y(y_{it}|\alpha_{it}) f_\alpha(\alpha_{it}|\alpha_{i,t-1}) d\alpha_{it} \cdots d\alpha_{i1}. \tag{5}$$

The above expression is based on the conditional independence assumption, according to which $f(y_{it}|y_{i1}, \dots, y_{i,t-1}, \alpha_{i1}, \dots, \dots, \alpha_{iT}) = f(y_{it}|\alpha_{it})$ and on the first-order Markov assumption on the latent variables, that is $f_\alpha(\alpha_{it}|y_{i1}, \dots, y_{i,t-1}, \alpha_{i1}, \dots, \dots, \alpha_{i,t-1}) = f_\alpha(\alpha_{it}|\alpha_{i,t-1})$. Notice that $f_\alpha(\alpha_{i1}|\alpha_{i0}) = f_\alpha(\alpha_{i1})$. Moreover, the form of the densities $f_y(y_{it}|\alpha_{it})$ and $f_\alpha(\alpha_{it}|\alpha_{i,t-1})$ depends on how the measurement equation (1) and the transition equation (2) are formulated.

Typically, computation and maximization of the likelihood (4) do not admit an analytical solution. An effective way to solve this problem is to apply non-linear filter techniques that allow us to formulate the multidimensional integral involved in expression (5) into a sequence of uni-dimensional integrals using the rules of conditioning as follows:

$$f(y_{i1}, \dots, y_{iT}) = f(y_{i1}) \prod_{t=2}^{T} f(y_{it}|y_{i1}, \dots, y_{i,t-1}),$$

where

$$f(y_{i1}) = \int f_y(y_{i1}|\alpha_{i1}) f_\alpha(\alpha_{i1}) d\alpha_{i1}$$

and

$$f(y_{it}|y_{i1}, \dots, y_{i,t-1}) = \int f_y(y_{it}|\alpha_{it}) f(\alpha_{it}|y_1, \dots, y_{i,t-1}) d\alpha_{it}. \tag{6}$$

The density $f(\alpha_{it}|y_{i1},\ldots,y_{i,t-1})$ in expression (6) can be obtained as follows:

$$f(\alpha_{it}|y_{i1},\ldots,y_{it}) = \frac{f_y(y_{it}|\alpha_{it})f(\alpha_{it}|y_{i1},\ldots,y_{i,t-1})}{f(y_{it}|y_{i1},\ldots,y_{i,t-1})}, \tag{7}$$

where

$$f(\alpha_{it}|y_{i1},\ldots,y_{i,t-1}) = \int f_\alpha(\alpha_{it}|\alpha_{i,t-1})f(\alpha_{i,t-1}|y_{i1},\ldots,y_{i,t-1})d\alpha_{i,t-1}, \tag{8}$$

The filtering algorithm consists of evaluating recursively formulas (6), (7), and (8).

Formulas (6) and (8) involve uni-dimensional integrals that cannot be computed analytically and hence have to be approximated. A widely used method is represented by the Gauss-Hermite (GH) quadrature that allows to evaluate numerically all the integrals of the form

$$\int e^{-z^2}f(z)dz \simeq \sum_{k=1}^{q} w_k f(z_k),$$

where z_k are the zeros of the Hermite orthogonal polynomial H_k, w_k are the correspondent weights, and q is number of quadrature points (Davis and Rabinowitz, 1975). The approximation is exact if $f(z)$ is a polynomial of degree equal to $2q-1$.

In order to apply the GH to the integral in (6), we rewrite it in the following form

$$f(y_{it}|y_{i1},\ldots,y_{i,t-1}) = \int \frac{f_y(y_{it}|\alpha_{it})f(\alpha_t|y_{i1},\ldots,y_{i,t-1})f_\alpha(\alpha_{it})}{f_\alpha(\alpha_{it})}d\alpha_{it}, \tag{9}$$

where $f_\alpha(\cdot)$ is the marginal distribution of α_{it} and it is a normal density distribution with mean μ and variance σ^2 whose expressions depend on the adopted model. For SV models we have $\mu = \gamma/(1-\rho)$ and $\sigma^2 = \sigma_\eta^2/(1-\rho^2)$; for LDV models we have $\mu = 0$ and $\sigma^2 = \sigma_\eta^2/(1-\rho^2)$. Denoting with $g(\alpha_{it}) = f(\alpha_{it}|y_{i1},\ldots,y_{it-1})/f_\alpha(\alpha_{it})$ and integrating over the standardized $\tilde{\alpha} = \frac{1}{\sigma}(\alpha-\mu)$, we obtain

$$\begin{aligned} f(y_{it}|y_{i1},\ldots,y_{i,t-1}) &= \frac{1}{\sqrt{2\pi}}\int f_y(y_{it}|\sigma\tilde{\alpha}_{it}+\mu)g(\sigma\tilde{\alpha}_{it}+\mu)\exp(-\tilde{\alpha_{it}}^2/2)d\tilde{\alpha}_{it} \\ &\simeq \sum_{k=1}^{q} f_y(y_{it}|\sigma z_k^*+\mu)g(\sigma z_k^*+\mu)w_k^*, \end{aligned} \tag{10}$$

where $z_k^* = \sqrt{2}z_k$ and $w_k^* = (1/\sqrt{\pi})w_k$. The same approximation can be applied to solve integral (8).

3.1 Adaptive Gauss-Hermite quadrature

An improved version of the GH approximation is given by the Adaptive Gauss Hermite (AGH) quadrature rule introduced in the Bayesian context with the aim of efficiently computing posterior densities if they are approximately normal (Naylor and Smith, 1982). It essentially consists of adjusting the GH quadrature locations with the mean and the variance of the posterior density, so that the nodes are more concentrated around the peak of the function to be integrand and a better approximation of this function results in comparison with the classical GH.

For the particular class of models defined in Section 2, the AGH quadrature is obtained by multiplying and dividing the integral in (9) by the normal density $\phi(\tilde{\alpha}_{it}, \tilde{\mu}_{it}, \tilde{\sigma}_{it})$. We have that

$$f(y_{it}|y_{i1}, \dots, y_{i,t-1}) = \frac{1}{\sqrt{2\pi}} \int \frac{f_y(y_{it}|\sigma\tilde{\alpha}_{it} + \mu) g(\sigma\tilde{\alpha}_{it} + \mu) \exp(-\tilde{\alpha_{it}}^2/2)}{\phi(\tilde{\alpha}_{it}, \tilde{\mu}_{it}, \tilde{\sigma}_{it})} \phi(\tilde{\alpha}_{it}, \tilde{\mu}_{it}, \tilde{\sigma}_{it}) d\tilde{\alpha}_{it} \tag{11}$$

such that, transforming the $\tilde{\alpha}_{it}$'s to standardized latent variables, the integral (11) is approximated as follows

$$f(y_{it}|y_{i1}, \dots, y_{i,t-1}) \simeq \tilde{\sigma}_{it} \sum_{k=1}^{q} \frac{f_y(y_{it}|\sigma\nu_{itk} + \mu) g(\sigma\nu_{itk} + \mu) \exp(-\nu_{itk}^2/2)}{\exp(-z_k^{*2}/2)} w_k^*, \tag{12}$$

where $\nu_{itk} = \tilde{\sigma}_{it} z_k^* + \tilde{\mu}_{it}$.

Two different procedures can be used to estimate $\tilde{\mu}_{it}$ and $\tilde{\sigma}_{it}$. The first one consists of approximating $\tilde{\mu}_{it}$ with the mode of the integrand and $\tilde{\sigma}_{it}$ with the standard deviation of the integrand at the mode (Liu and Pierce, 1994; Pinheiro and Bates, 1995; Schilling and Bock, 2005). The advantage of this approach lies on the fact that the quadrature points are not involved in these computations. On the other hand, this method is very computationally demanding since it requires numerical optimization routines. Moreover, when parameter estimates are obtained by using iterative algorithms as in this case, the first two moments have to be computed at each step of the algorithm that, for this reason, can be rather slow. An alternative method consists of computing the posterior mean and standard deviation (Naylor and Smith, 1982; Rabe-Hesketh et al., 2005). Although this method requires the use of quadrature points themselves, it is more robust to fat tailed distributions and it is faster when a sequential scheme is used. For these reasons we adopt the second procedure.

In more detail, the computation of $\tilde{\mu}_{it}$ and $\tilde{\sigma}_{it}$ is iteratively obtained as follows:

1. Choose starting values $\tilde{\mu}_{it}^{(0)} = 0$ and $\tilde{\sigma}_{it}^{(0)} = 1$ so that $\nu_{ik}^{(0)} = z_k^*$.

2. Compute the log-likelihood for the l-th iteration

$$\begin{aligned}\log L^{(l)}(\theta) &= \log \prod_{i=1}^{n} f^{(l)}(y_{i1}, \ldots, y_{iT}) \\ &\simeq \log \left[\prod_{i=1}^{n} \left(\prod_{t=1}^{T} \tilde{\sigma}_{it}^{(l-1)} \sum_{k=1}^{q} \frac{f_y(y_{it}|\sigma\nu_{itk}^{(l-1)} + \mu) g(\sigma\nu_{itk}^{(l-1)} + \mu) \exp(-(\nu_{itk}^{(l-1)})^2/2)}{\exp(-z_k^{*2}/2)} w_k^* \right) \right],\end{aligned}$$

   with $g(\sigma\nu_{itk} + \mu) = 1$ for $t = 1$.

3. Update each node $\nu_{itk}^{(l)}$ by computing the posterior mean and standard deviation as follows

$$\tilde{\mu}_{it}^{(l)} = \tilde{\sigma}_{it}^{(l-1)} \sum_{k=1}^{q} \frac{\nu_{itk}^{(l-1)} f_y(y_{it}|\sigma\nu_{itk}^{(l-1)} + \mu) g(\sigma\nu_{itk}^{(l-1)} + \mu) \exp(-(\nu_{itk}^{(l-1)})^2/2)}{f^{(l)}(y_{it}|y_{i1}, \ldots, y_{it-1}) \exp(-z_k^{*2}/2)} w_k^*$$

$$\tilde{\sigma}_{it}^{(l)} = \sqrt{\tilde{\sigma}_{it}^{(l-1)} \sum_{k=1}^{q} \frac{(\nu_{itk}^{(l-1)})^2 f_y(y_{it}|\sigma\nu_{itk}^{(l-1)} + \mu) g(\sigma\nu_{itk}^{(l-1)} + \mu) \exp(-(\nu_{itk}^{(l-1)})^2/2)}{f^{(l)}(y_{it}|y_{i1}, \ldots, y_{it-1}) \exp(-z_k^{*2}/2)} w_k^* - (\tilde{\mu}_{it}^{(l)})^2}$$

Steps 2 and 3 are repeated until convergence. Parameter estimation is obtained using a quasi-Newton method. Thus the estimation algorithm consists in alternating one step of the quasi-Newton procedure to update parameter estimates and the iterative scheme proposed above to update the nodes of the adaptive numerical quadrature.

4 Simulation study

We carried out two simulation studies in order to evaluate the performance of AGH in LDV and SV models. In both cases the results are compared with other approximation methods discussed in the literature.

4.1 Simulation study for the LDV model

The LDV model considered in this simulation study is the proportional odds model illustrated above. For this model, the performance of AGH is compared with the classical GH approximation respectively for increasing values of the correlation parameter $\rho = (0.5, 0.90, 0.95)$, different combinations of time points ($T = 5, 10$), and sample sizes ($n = 500, 1000$). Five categories for the ordinal responses are chosen and the thresholds are fixed equal to $\tau_1 = -1.65, \tau_2 = -0.5, \tau_3 = 0.5, \tau_4 = 1.65$. The remaining parameters β and σ_η are fixed to 1. One covariate is generated from a stationary AR(1) with autocorrelation parameter equal to 0.5. The quadrature points chosen for AGH are $q = 15, 21$ and for GH are $q = 21, 51$. Under each simulation scenario, 500 samples were generated and the model parameters were estimated on the basis of the two approximation methods.

The comparison between the two different numerical approximations is assessed in terms of both accuracy of estimates and computational performance of the algorithms. For $T = 5$ and $n = 500$ the results are reported in Table 1.

[Table 1 about here.]

For $\rho = 0.5$, under both GH and AGH approximation the algorithm converges properly in almost all the samples (*%cv*). Moreover, the average number of function evaluations (*nr feval*) is quite similar for both the methods. As for the rmse of the estimates, the best performances are obtained with GH_{51} and AGH_{21} that produce very close results for all the parameters estimates. In more detail, in terms of bias AGH_{21} is always better than GH_{51}, in terms of rmse adaptive is better than the classical quadrature for β and σ_η and slightly worse for the autocorrelation parameter. Thus, with less than half quadrature points than those used with GH_{51}, AGH_{21} produces very accurate estimates. Moreover, differently from GH, AGH is not affected by the choice of starting values. Nevertheless, the average computational time (*av time* in seconds) to convergence of the latter approximation is about three times that of the former, making GH preferable in this case. This is due to the iterative routine required for the computation of the posterior means and standard deviations at each iteration of the algorithm. However, it is possible to improve the speed of convergence of AGH by using a pseudo version of the AGH that consists in updating the quadrature nodes for each sample unit with the posterior means and standard deviations only at the first step of the algorithm, implying a consistently reduction of the computation time. This method was proposed by Rizopoulus (2012) for a class of linear mixed effect models for longitudinal data. The author showed that, for the class of models they considered, PseudoAGH produces accurate parameter estimates when proper starting values are used. Here we chose starting values by means of a multi-start strategy. The results of the PseudoAGH with $q = 21$ are reported in the last column of Table 1. We observe that the rmse of ρ is the same as AGH_{21}, whereas the rmse of the other parameter estimates are higher than the rmse obtained with AGH_{21}. On the other hand, the computational time of PseudoAGH is noticeably lower than the standard AGH (the ratio is about 1 to 4 and in same cases about 1 to 5) and slightly lower than GH_{51}.

In the cases of $\rho = 0.90$ and $\rho = 0.95$, the superiority of the adaptive procedure is undoubted. Indeed GH does not produce convergent solutions in all the generated samples because of instability problems. On the contrary, AGH performs very well even with $q = 15$, being the rmse of all the estimates equal to those obtained for $q = 21$. Moreover, in this case PseudoAGH presents the same rmse of all the parameter estimates than both AGH_{15} and AGH_{21} and, as before, reveals superior in terms of computational time. Thus, for high values of the autocorrelation parameters, PseudoAGH reveals the best method.

In order to better understand the behavior of GH and AGH under different values of ρ, the integrand reported in formula (6) is represented in Figure 1, for an ordinal response variable in a given time point, in the cases of $\rho = 0.5$ (upper pictures) and $\rho = 0.95$ (lower pictures), together with the GH and AGH nodes and weights for $q = 21$. In the case of $\rho = 0.5$, the integrand has a smooth shape and both methods approximate it very well. When $\rho = 0.95$, the integrand has a very sharped peak and only AGH properly captures it.

[Figure 1 about here.]

The simulation results for $T = 10$ and $n = 500$ are reported in Table 2. As expected, in general the rmse of the estimates improve compared with those obtained in the previous scenario. As for GH, convergent solutions are obtained only for $\rho = 0.5$, whereas AGH performs well for all the values of the autocorrelation parameter even if the computational time is quite heavy. Also in this case, PseudoAGH seems to be the best compromise between GH and AGH, particularly in the case of high values of ρ where, as before, the accuracy of the parameter estimates results similar and in same cases even better than the standard adaptive quadrature. The reduction in the computational time results of the same magnitude as in the previous scenarios.

[Table 2 about here.]

The case of $T = 5$ and $n = 1000$ (Table 3) is the only one in which we get convergent solutions under GH also for high values of ρ.

[Table 3 about here.]

However, for $\rho = 0.90$, even if GH_{51} performs similarly to AGH_{21} in terms of rmse, we can observe that *%cv* is 74% versus a percentage equal to 98% generated sample in which the algorithm properly converges under AGH_{21} and to %100 under PseudoAGH_{21}. Moreover, as in all the previous cases, the latter results to be the fastest. We can observe the same behavior of the different approximation methods also for $\rho = 0.95$. In particular, in this case GH produces a *%cv* equal only to 36%.

4.2 Simulation study for the SV model

The design of the study considered for the SV model is based on the same setting considered by Jacquier et al. (1994) and widely used in the literature for evaluating the performance of different approximation methods. Here the results obtained with the AGH approximation are compared with those obtained by Fridman and Harris (1998), who used the Gauss-Legendre quadrature approximation (GLQ), by Bartolucci and De Luca (2001), who proposed a rectangular type quadrature method (RQ), and by Junji and Yoshihiko (2005), who used the Laplace approximation (LA). The Laplace approximation is considered since it can be viewed as a particular case of AGH when one quadrature point is chosen (Pinheiro and Bates, 1995).

The parameter values of the generating model were chosen so that different values of the squared Coefficient of Variation of the volatility, $\text{CV} = \exp(\sigma_\eta^2/(1-\rho^2)) - 1$, result. In more detail, the most relevant parameter ρ is fixed to 0.90, 0.95 and 0.98, that are values based on the empirical evidence. The other parameter values are consequently determined so that CV assumes the values 0.1, 1, or 10, producing nine different scenarios. In each simulated scenario, we let $T = 500$ and used 500 Monte Carlo replications; for AGH we adopted $q = 21$ quadrature points.

The mean of the parameter estimates and the root mean square error (rmse) in brackets for SV models with different combinations of the parameters are reported in Table 4.

[Table 4 about here.]

The Table is grouped according to the values of the CV. We observe that in the case of $\mathrm{CV} = 10$ all the methods perform quite similarly in terms of rmse. Only in the case of $\rho = 0.98$, GLQ produces better rmse than the other approximations. However, as the CV decreases, all the methods deteriorate sensibly apart from AGH that produces rmse always with the same small magnitude. The superiority of the proposed method is particularly noticeable in the case of $\mathrm{CV} = 0.1$ for the parameters γ and ρ. In terms of bias, for most of the parameter values AGH gives better or similar results than the other methods.

In the table, the average computational time (in seconds) of AGH is also showed for all the scenarios. We observe that in all the cases the algorithm reaches convergence in few seconds. We cannot compare the computational performance of the proposed method with that of the other approximation methods since for the latter we do not have this information.

5 Real data analysis

In the following, we illustrate the proposed approach by two applications in the context of panel data analysis and in that of time-series data.

5.1 Application of limited dependent variable models to Self-reported health status

We consider panel data deriving from the Health and Retirement Study (HRS) [1] conducted by the University of Michigan with the aim of studying retirement and health among elderly people

[1] The RAND HRS Data file is an easy to use longitudinal data set based on the HRS data. It was developed at RAND with funding from the National Institute on Aging and the Social Security Administration. See http://www.rand.org/labor/aging/dataprod.html for more details.

in the United States over time. These data are referred to a sample of $n = 7,074$ individuals who were asked at $T = 8$ time occasions (from 1992 to 2006 every two years) to report the self-rated health status (SHR) by answering to the question "Would you say your health is excellent, very good, good, fair, or poor?", that is an ordinal response variable with five categories. Also the covariates gender, race ("white", "non white"), education ("high school", "some college", "college and above"), and age measured at each time point are available. Some descriptive statistics for the distributions of the covariates are reported in Table 5, whereas Table 6 reports the conditional sample distribution of SHR_t given the previous response SHR_{t-1}.

[Table 5 about here.]

[Table 6 about here.]

We observe that the sample is mainly composed by females (58.1%), white individuals (82.9%) with an average age at the first observed time point equal to 54.8 years. As for the level of education, 60.9% of the interviewed subjects declared to have a high-school diploma, 19.7% of them a college degree, and 19.4% a higher title. The variable education has been recoded in the following way: 1 for "high school", 2 for "some college", and 3 for "college and above".

The high percentage of people (more than %50 percent for all the categories) that respond to the same category at time $t-1$ and at time t indicates that SHR is highly correlated over time. Previous analyses on HRS data (Heiss, 2008) showed that the proportional odds model illustrated above with time dependent random intercepts following a stationary AR(1) well captures the SHR correlation pattern over time. More recently, Bartolucci et al. (2013a) analyzed the HRS data assuming a more flexible model based on a mixture of AR(1) models for the latent process. They approximated the integrals with a rectangular quadrature method. Since here the aim is to evaluate the performance of the AGH discussed above, we fitted the proportional odds model with a standard stationary AR(1).

In Table 7 we report the results of model estimation under AGH_{21} and PseudoAGH_{21} approximations.

[Table 7 about here.]

We can notice that in this example where the sample size is quite large, the computational time of PseudoAGH$_{21}$ is noticeably lower than that of AGH$_{21}$. The parameter estimates are quite similarly under both methods and in both cases all of them are significant apart from the coefficient of the covariate gender. It is worth noting that the autocorrelation parameter estimate is higher than 0.95 indicating a high persistent latent process over time. This result is in agreement with the previous analyses on the HRS data (Heiss, 2008; Bartolucci et al., 2013a).

5.2 Application of stochastic volatility models to daily exchange rates

To illustrate the application of the method to SV models, we use a data set analyzed by Harvey et al. (1994) and later by several other authors. The data consist of a time-series of daily pound/dollar exchange rates from the period October 1st, 1981 to June 28th, 1985. The series of interest is the logarithm of $n = 945$ daily returns (Figure 2).

[Figure 2 about here.]

For these data, we fitted the standard SV model illustrated above and an SV model with error terms ε_{it} following a t-Student distribution with unknown degrees of freedom ν. This choice is motivated by the fact that many financial time-series exhibit densities with fatter tails than the Gaussian distribution. AGH with $q = 21$ quadrature points is used for approximating the integrals involved in the likelihood of both model. The results of parameters estimates are reported in Table 8.

[Table 8 about here.]

Under both models ρ has a high and significant value, indicating a highly persistent volatility process. The higher value of the log-likelihood obtained under the normal assumption (-921.610 for the SV normal model versus -922.289 for the SV t-Student model) suggests that there is no improvement assuming t-Student distributed errors.

The similarity between the two specifications for the error terms can be also seen by examining the estimated filtered volatilities obtained for each model. Figure 3 shows the plot of the estimated filtered volatilities for each time point of the t-Student distribution (y-axis) versus the normal distribution (x-axis). We observe that there is a substantial agreement in terms of estimated volatility between the two specifications.

[Figure 3 about here.]

6 Conclusions

In this work we proposed the Adaptive Gauss Hermite (AGH) quadrature for approximating the integrals involved in the likelihood of a class of dynamic latent variable models based on a latent process following an autoregressive process of order 1, AR(1). In particular, we focused on Stochastic Volatility (SV) models for the analysis of financial time-series and on Limited Dependent Variable (LDV) models for panel data. Both models can be formalized in a non-linear state space framework and maximum likelihood estimation can be obtained by mean of a non-linear filtering algorithm.

The main advantage of AGH compared with other numerical approximation methods is that it better captures the peak of the integrand in those cases in which it appears very sharp, using fewer quadrature points than other methods. This is due to the fact that the nodes of AGH are scaled and translated at each step of the algorithm with the posterior mean and the posterior standard deviation of the latent variables given the manifest variables.

The good behavior of AGH has been highlighted by means of simulation studies. For SV models we found that the advantages of AGH with respect to other quadrature methods are particularly evident in terms of parameter accuracy for those values of the model parameters that give low coefficients of variation of the volatility. For LDV models, we found that the performance of AGH is related to the value of the autocorrelation parameter ρ. For high values of ρ, parameter estimates are very accurate under AGH even with $q = 15$ quadrature points, whereas GH produces convergent solutions only in few cases when the sample size is large. On the other hand, for $\rho = 0.5$, AGH gives results very similar to the classical GH but with a higher computational cost. To solve this problem, we considered a pseudo version of AGH that consists in updating the nodes of the quadrature only at the first step of the algorithm. We found a very good performance of this method in terms of both computational burden of the algorithm and parameter accuracy in almost all the scenarios considered. However, differently from AGH, the performance of PseudoAGH is dependent on the choice of the starting values.

The potential of the proposed method has been showed also with an application to two real data sets, the first one referred to an American longitudinal survey on the health condition of the elderly population, the second one to a time-series of daily pound/dollar exchange rates in a given period of time. In both examples, the latent variables showed high values of autocorrelation, values for which the adaptive quadrature presents the best performance compared with the other approximations. The high persistency of the latent variable is typical of SV models but it is also plausible in panel data for the underlying process of time dependent response variables, as for the health status considered here.

Acknowledgements

The authors acknowledge the financial support from the grant RBFR12SHVV funded by the Italian Government (FIRB project "Mixture and latent variable models for causal inference and analysis of socio-economic data").

References

TG Andersen. Stochastic autoregressive volatility: a framework for volatility modeling. *Mathematical Finance*, 4:75–102, 1994.

F Bartolucci and G De Luca. Maximum likelihood estimation for a latent variable time series model. *Applied Stochastic Models for Business and Industry*, 17:5–17, 2001.

F Bartolucci and G De Luca. Likelihood-based inference for asymmetric stochastic volatility models. *Computational Statistical and Data Analysis*, 42:445–449, 2003.

F Bartolucci, S Bacci, and F Pennoni. Longitudinal analysis of the self-reported health status by mixture latent autoregressive models. *Journal of the Royal Statistical Society-series C*, page in press, 2013a.

F. Bartolucci, A. Farcomeni, and F. Pennoni. *Latent Markov Models for Longitudinal Data*. Chapman & Hall/CRC Press, Boca Raton, FL, 2013b.

T Bollerslev. Generalized autoregresive conditional heteroskedasticity. *Journal of Econometrics*, 31:307–327, 1986.

S Cagnone and P Monari. Latent variable models for ordinal data by using the adaptive quadrature approximation. *Computational Statistics*, 28:597–619, 2013.

J Durbin and SJ Koopman. A simple and efficient simulation smmother for space time series analyis. *Biometrika*, 89:603–616, 2002.

E.W. Frees. *Longitudinal and Panel Data: Analysis and Application in Social Sciences*. 2004.

M Fridman and L Harris. A maximum likelihood approach for non-gaussian stochastic volatility models. *Journal of Business and Economic Statistics*, 16:284–291, 1998.

AC Harvey, E Ruiz, and N Shephard. Multivariate stochastic variance models. *Review of Economic Studies*, 61:247–264, 1994.

F Heiss. Sequential numerical integration in nonlinear state spece models for microeconometric panel. *Journal of Applied Econometrics*, 23:373–389, 2008.

E Jacquier, NG Polson, and PE Rossi. Bayesian analysis of stochastic volatility models. *Journal of Business and Economic Statistics*, 12:371–417, 1994.

H Joe. Accuracy of laplace approximation for discrete response mixed models. *Computational Statistics and Data Analysis*, 52:5066–5074, 2008.

S Junji and T Yoshihiko. Estimation of stochastic volatility models: an approximation to the nonlinear state space representation. *Communications in Statistics- Simulation and Computation*, 34:429–450, 2005.

Q Liu and DA Pierce. A note on gauss-hermite quadrature. *Biometrika*, 81:624–629, 1994.

G.S. Maddala. *Limited Dependent and Qualitative Variables in Econometrics*. Cambridge University Press, 1983.

P McCullagh. Regression models for ordinal data (with discussion). *Journal of the American Statistical Society, Series B*, 42:109–142, 1980.

JC Naylor and AFM Smith. Applications of a method for the efficient computation of posterior distributions. *Applied Statistics*, 51:214–225, 1982.

JC Pinheiro and DM Bates. Approximation to the loglikelihood function in the nonlinear mixed effects model. *Journal of Computational and Graphical Statistics*, 4:12–35, 1995.

S Rabe-Hesketh, A Skrondal, and A Pickles. Reliable estimation of generalized linear mixed models using adaptive quadrature. *The Stata Journal*, 2:1:21, 2002.

S Rabe-Hesketh, A Skrondal, and A Pickles. Maximum likelihood estimation of limited and discrete dependent variable models with nested random effects. *Journal of Econometrics*, 128:301–323, 2005.

D Rizopoulus. Fast fitting of joint models for longitudinal and event time data using a pseudo-adaptive gaussian quadrature rule. *Computational Statistics and Data Analysis*, 56:491–501, 2012.

S Schilling and RD Bock. High-dimensional maximum marginal likelihood item factor analysis by adaptive quadrature. *Psychometrika*, 70:533–555, 2005.

H Tanizaki and RS Mariano. Nonlinear and nonnormal state-space modeling with monte-carlo stochastic simulations. *Journal of Econometrics*, 83:263–290, 1998.

S Taylor. *Modelling Financial Time Series*. New York: Wiley, 1986.

S Taylor. Modeling stochastic volatility. *Mathematical Finance*, pages 183–204, 1994.

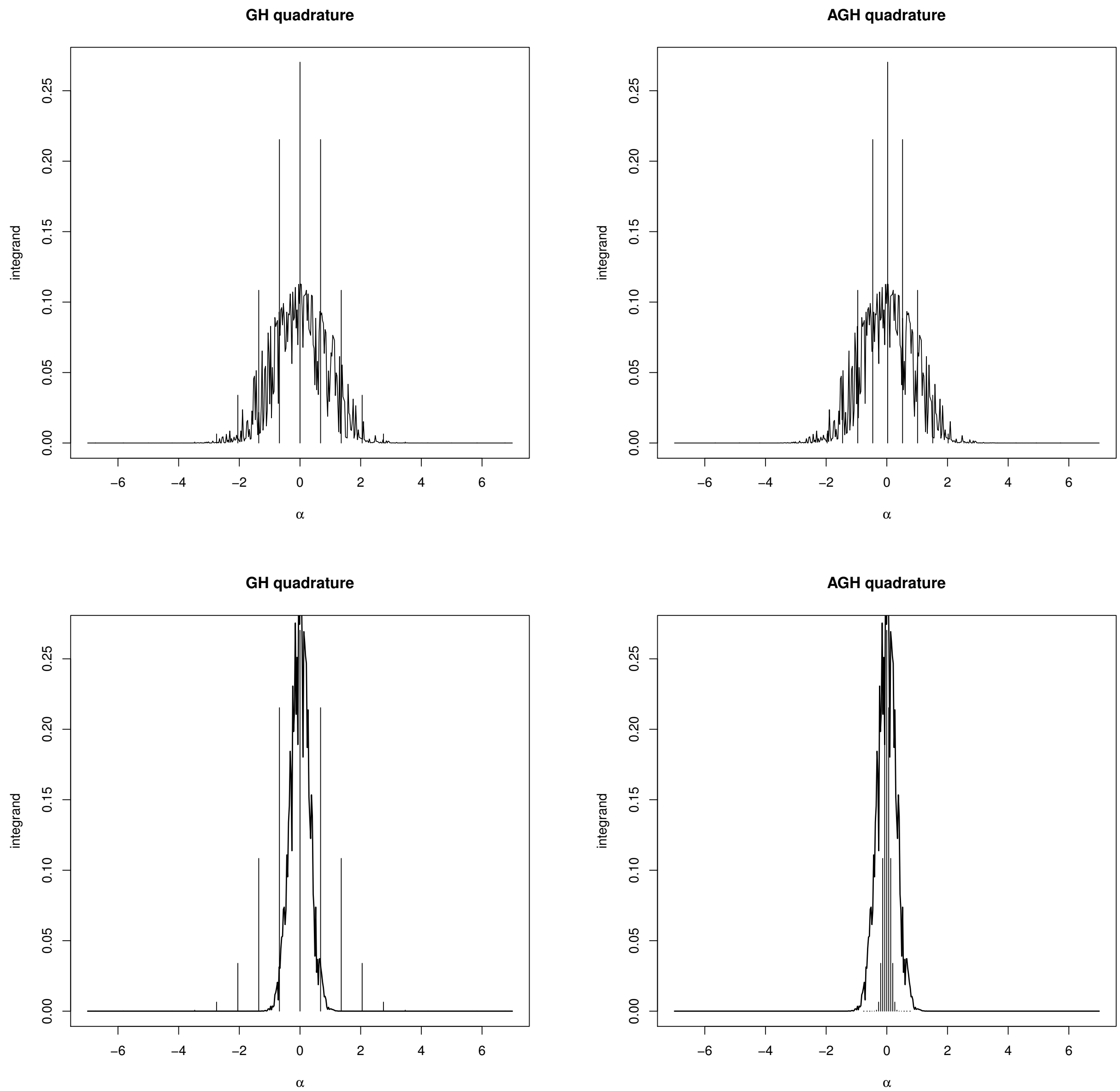

Figure 1: *Performance of the two approximations for* $\rho = 0.5$ *(first row) and* $\rho = 0.95$ *(second row).*

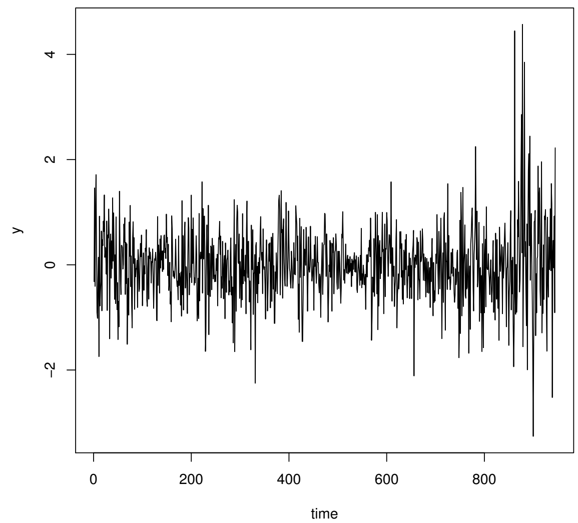

Figure 2: *Log-daily difference of the pound-dollar exchange rate from October 1st, 1981 to June 28th, 1985.*

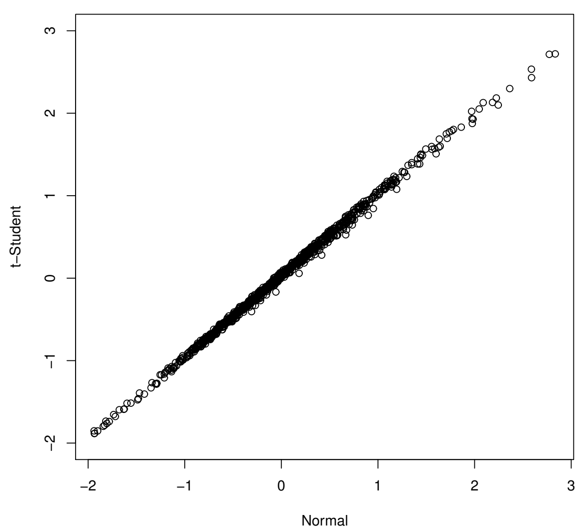

Figure 3: *Filtered estimated volatilities obtained assuming the Normal distribution (x-axis) and the t-Student distribution (y-axis) for the error terms for the daily exchange rates dataset.*

Table 1: *Estimated mean and rmse (in brackets) for the LDV model parameters under* GH *and* AGH *for* $T = 5$, $n = 500$*; results based on* 500 *replications.*

True value	GH_{21}	GH_{51}	AGH_{15}	AGH_{21}	$Pseudo AGH_{21}$
$\rho = 0.5$	0.478 (0.17)	0.487 (0.18)	0.487 (0.18)	0.492 (0.19)	0.493 (0.19)
$\beta = 1$	1.109 (0.35)	1.094 (0.28)	1.118 (0.40)	1.087 (0.25)	1.100 (0.34)
$\sigma_\eta = 1$	1.271 (0.86)	1.238 (0.72)	1.287 (0.97)	1.216 (0.67)	1.244 (0.84)
%cv	92	95	99	100	100
nr feval	48	48	41	41	40
av time (sec)	11.14	40.36	78.36	125.54	32.26
$\rho = 0.90$	-	-	0.871 (0.05)	0.884(0.05)	0.882 (0.05)
$\beta = 1$	-	-	1.016 (0.05)	1.011(0.05)	1.012 (0.05)
$\sigma_\eta = 1$	-	-	1.045 (0.10)	1.028(0.11)	1.030 (0.10)
%cv	-	-	100	100	100
nr feval	-	-	28	35	31
av time (sec)	-	-	64.91	129.56	22.68
$\rho = 0.95$	-	-	0.898 (0.06)	0.917(0.05)	0.914 (0.03)
$\beta = 1$	-	-	1.025 (0.05)	1.015(0.05)	1.019 (0.05)
$\sigma_\eta = 1$	-	-	1.074 (0.11)	1.050(0.10)	1.053 (0.10)
%cv	-	-	100	100	100
nr feval	-	-	27	32	28
av time (sec)	-	-	65.4	100.18	19.14

Table 2: *Estimated mean and rmse (in brackets) for the LDV model parameters under GH and AGH,* $T = 10$*,* $n = 500$*; results based on 500 replications.*

True value	GH_{21}	GH_{51}	AGH_{15}	AGH_{21}	*Pseudo*AGH_{21}
$\rho = 0.5$	0.455 (0.10)	0.481 (0.10)	0.490 (0.10)	0.490 (0.10)	0.490 (0.10)
$\beta = 1$	1.044 (0.11)	1.028 (0.11)	1.025 (0.14)	1.023 (0.10)	1.025 (0.13)
$\sigma_\eta = 1$	1.137 (0.31)	1.081 (0.30)	1.070 (0.35)	1.065 (0.29)	1.070 (0.34)
%cv	61	81	100	100	100
nr feval	63	53	56	56	56
av time (sec)	24.43	80.12	227.48	288.38	73.23
$\rho = 0.90$	-	-	0.892 (0.02)	0.898(0.02)	0.898 (0.02)
$\beta = 1$	-	-	1.003 (0.03)	1.001(0.03)	1.001 (0.03)
$\sigma_\eta = 1$	-	-	1.011 (0.06)	1.001(0.06)	1.002 (0.06)
%cv	-	-	97	98	100
nr feval	-	-	37	36	36
av time (sec)	-	-	137.64	188.03	37.91
True value	GH_{21}	GH_{51}	AGH_{15}	AGH_{21}	*Pseudo*AGH_{21}
$\rho = 0.95$	-	-	0.923 (0.03)	0.938 (0.02)	0.935 (0.04)
$\beta = 1$	-	-	1.013 (0.04)	1.007 (0.04)	1.010 (0.03)
$\sigma_\eta = 1$	-	-	1.034 (0.15)	1.017 (0.08)	1.030 (0.06)
%cv	-	-	99	97	100
nr feval	-	-	36	36	39
av time(sec)	-	-	164.71	278.73	52.23

Table 3: *Estimated mean and rmse (in brackets) for the LDV model parameters under GH and AGH for $T = 5$, $n = 1000$; results based on 500 replications.*

True value	GH_{21}	GH_{51}	AGH_{15}	AGH_{21}	$PseudoAGH_{21}$
$\rho = 0.5$	0.478 (0.14)	0.482 (0.14)	0.487 (0.14)	0.486 (0.15)	0.488 (0.14)
$\beta = 1$	1.064 (0.19)	1.058 (0.17)	1.058 (0.23)	1.053 (0.16)	1.057 (0.23)
$\sigma_\eta = 1$	1.170 (0.52)	1.150 (0.48)	1.147 (0.58)	1.138 (0.46)	1.145 (0.58)
%cv	77	92	100	100	100
nr feval	50	55	58	57	57
av time(sec)	24.46	83.27	191.66	309.6	72.99
$\rho = 0.90$	0.828 (0.08)	0.876 (0.04)	0.875 (0.04)	0.888 (0.04)	0.887 (0.04)
$\beta = 1$	1.030 (0.05)	1.013 (0.04)	1.013 (0.04)	1.008 (0.03)	1.009 (0.03)
$\sigma_\eta = 1$	1.099 (0.12)	1.038 (0.08)	1.036 (0.07)	1.020 (0.08)	1.021 (0.07)
%cv	18	74	96	98	100
nr feval	55	50	49	50	50
av time(sec)	19.36	70.84	167.64	312.05	59.84
$\rho = 0.95$	-	0.902 (0.05)	0.901 (0.05)	0.923 (0.04)	0.918 (0.04)
$\beta = 1$	-	1.021 (0.04)	1.022 (0.04)	1.015 (0.05)	1.016 (0.03)
$\sigma_\eta = 1$	-	1.067 (0.09)	1.068 (0.12)	1.036 (0.10)	1.046 (0.07)
%cv	-	36	98	100	100
nr feval	-	50	44	48	46
av time(sec)	-	72.49	192.63	269.41	62.11

Table 4: *Comparison between different approximation methods in terms of the estimated means and rmse of the SV model parameters.*

CV=10	γ	ρ	σ_η	γ	ρ	σ_η	γ	ρ	σ_η
TRUE	-0.821	0.90	0.675	-0.411	0.95	0.484	-0.164	0.98	0.308
GLQ	-0.896	0.890	0.685	-0.505	0.940	0.495	-0.100	0.986	0.320
	(0.28)	(0.03)	(0.08)	(0.18)	(0.02)	(0.07)	(0.08)	(0.01)	(0.05)
RQ	-0.859	0.895	0.694	-0.472	0.943	0.503	-0.275	0.967	0.343
	(0.25)	(0.03)	(0.08)	(0.18)	(0.02)	(0.07)	(0.18)	(0.02)	(0.07)
LA	-0.905	0.880	0.727	-0.510	0.931	0.534	-0.259	0.965	0.343
	(0.28)	(0.04)	(0.10)	(0.23)	(0.03)	(0.09)	(0.18)	(0.02)	(0.07)
AGH	-0.613	0.925	0.725	-0.509	0.938	0.527	-0.276	0.966	0.358
	(0.23)	(0.03)	(0.15)	(0.16)	(0.02)	(0.08)	(0.19)	(0.02)	(0.08)
av time(sec)		17.97			12.34			14.86	

CV=1.0	γ	ρ	σ_η	γ	ρ	σ_η	γ	ρ	σ_η
TRUE	-0.736	0.90	0.363	-0.368	0.95	0.26	-0.147	0.98	0.166
GLQ	-0.870	0.880	0.370	-0.510	0.930	0.280	-0.090	0.987	0.180
	(0.43)	(0.05)	(0.08)	(0.31)	(0.04)	(0.07)	(0.06)	(0.02)	(0.04)
RQ	-0.812	0.890	0.375	-0.492	0.933	0.278	-0.308	0.958	0.214
	(0.45)	(0.06)	(0.09)	(0.29)	(0.04)	(0.07)	(0.25)	(0.03)	(0.08)
LA	-0.926	0.872	0.422	-0.526	0.927	0.303	-0.278	0.961	0.200
	(0.42)	(0.06)	(0.11)	(0.39)	(0.05)	(0.10)	(0.25)	(0.03)	(0.07)
AGH	-0.572	0.922	0.359	-0.475	0.935	0.293	-0.341	0.953	0.213
	(0.17)	(0.02)	(0.04)	(0.21)	(0.03)	(0.08)	(0.26)	(0.03)	(0.07)
av time(sec)		13.85			11.59			17.83	

CV=0.1	γ	ρ	σ_η	γ	ρ	σ_η	γ	ρ	σ_η
TRUE	-0.706	0.90	0.135	-0.353	0.95	0.096	-0.141	0.98	0.061
GLQ	-1.360	0.810	0.160	-0.810	0.886	0.120	-0.537	0.924	0.088
	(1.72)	(0.24)	(0.12)	(1.15)	(0.16)	(0.09)	(1.13)	(0.16)	(0.09)
RQ	0.944	0.873	0.159	-0.796	0.888	0.148	-0.515	0.927	0.122
	(1.24)	(0.17)	(0.10)	(0.77)	(0.11)	(0.10)	(0.94)	(0.13)	(0.11)
LA	-1.227	0.827	0.178	-0.763	0.892	0.133	-0.489	0.93	0.099
	(1.55)	(0.22)	(0.14)	(1.16)	(0.16)	(0.12)	(0.98)	(0.14)	(0.11)
AGH	-0.521	0.926	0.137	-0.455	0.935	0.098	-0.568	0.920	0.131
	(0.21)	(0.03)	(0.08)	(0.18)	(0.03)	(0.08)	(0.52)	(0.07)	(0.10)
av time(sec)		17.02			20.03			18.77	

Table 5: *Summary statistics for the covariates in the HRS dataset;* $n = 7074$.

Variable	Mean	Stdev
Gender (female)	0.581	0.490
Race (nonwhite)	0.171	0.377
Education		
(high school)	0.609	0.488
(college degree)	0.197	0.398
(college +)	0.194	0.395
Age_{92}	54.80	5.460

Table 6: *Conditional distribution of* SHR_t *given* SHR_{t-1} *for the HRS dataset;* $n = 7074$.

	poor	fair	good	very good	excellent
poor	54.5	34.1	8.4	2.5	0.7
fair	12.8	51.0	27.4	7.2	1.6
good	2.5	16.5	53.3	23.6	4.1
very good	0.8	4.7	25.9	55.6	13.0
excellent	0.4	1.9	10.6	33.7	53.4

Table 7: *Estimates of the parameters of the LDV model adopted for the analysis of the HRS dataset (standard errors in brackets).*

	AGH	*PseudoAGH*
$\hat{\beta}_1$ female	-0.147 (0.074)	-0.099 (0.073)
$\hat{\beta}_2$ non white	-1.509 (0.096)	-1.394 (0.091)
$\hat{\beta}_3$ education	1.182 (0.046)	1.141 (0.047)
$\hat{\beta}_4$ age	-0.109 (0.003)	-0.089 (0.003)
$\hat{\rho}$	0.953 (0.018)	0.955 (0.015)
$\hat{\sigma}$	3.121 (0.036)	2.860 (0.034)
Log-lik	-63591.17	-63595.50
Time (sec)	9612.34	2588.60

Table 8: *Estimates of the parameters of SV model adopted of the analysis of the daily exchange rates using AGH_{21} (with standard errors in brackets).*

	Model	
	SV Normal	SV t-Student
γ	-0.0906 (0.009)	-0.0914 (0.011)
ρ	0.9093 (0.011)	0.9091 (0.010)
σ_η^2	0.3034 (0.002)	0.2831 (0.031)
ν	-	22.020 (11.923)
Log-lik	-921.610	-922.289
Time (sec)	29.73	34.79