WORKING PAPER n.02.01

Marzo 2002

Simulation Based Methods for Financial Time Series

M. Billio[a,b]

a. Università Ca' Foscari, Venice.
b. GRETA, Venice.

Simulation Based Methods for Financial Time Series

(Metodi simulati per serie finanziarie)

Monica Billio
Dipartimento di Scienze Economiche
Università di Venezia
billio@unive.it

Riassunto: Il lavoro presenta una rassegna dei metodi di inferenza fondati sull'uso di simulazioni. Essi sono molto utili nell'ambito dei modelli dinamici a variabili latenti di cui i modelli a volatilità stocastica sono un esempio.

Keywords: Dynamic Latent Variable models, Stochastic Volatility models, Indirect Inference, Efficient Method of Moments, Importance Sampling, Bayesian approach, Markov Chain Monte Carlo methods, Simulated Expectation Maximization algorithm.

1. Introduction

The empirical distributions of financial time series differ substantially from distributions obtained from sampling independent homoskedastic Gaussian variables. Unconditional density functions exhibit leptokurtosis and skewness; time series of stock returns show evidence of volatility clustering; and squared returns exhibit pronounced serial correlation whereas little or no serial dependence can be detected in the return process itself. These empirical regularities suggest that the behaviour of financial time series may be captured by a model which recognizes the time varying nature of return volatility, as follows:

$$y_t = \mu_t + y_t^* \varepsilon_t, \qquad \varepsilon_t \sim IID(0,1), \qquad t = 1, 2, ..., T$$

where y_t denotes the return on an asset. One way of modelling y_t^* is to express it as a deterministic function of lagged residuals. Econometric specifications of this form are known as ARCH models and have achieved widespread popularity in applied empirical research (Bollerslev *et al.* (1992); Bollerslev *et al.* (1993); Bera and Higgins (1993)). Alternatively, volatility may be modelled as an unobserved component following some latent stochastic process, such as an autoregression. Models of this kind are known as stochastic volatility (SV) models (Taylor (1994); Ghysels *et al.* (1996); Shephard (1996)).

SV models turn out to be more appealing for many reasons: broad general features of the data can be reproduced (persistent volatility, volatility clustering effect, leverage effect, asymmetries and leptokurtosis), fewer parameters have to be estimated, and generally they are closed under temporal aggregation. Despite their intuitive appeal, SV models have been used less frequently than ARCH models in empirical applications. This is due to the difficulties associated with their estimation. Unlike ARCH models, where the likelihood function can be evaluated exactly, the likelihood function of a SV model is hard to construct.

SV models are examples of the general class of *parametric dynamic latent variable models* (also called factor models or state space models). These models are becoming increasingly popular because of the flexibility they offer in the modelling of complex

phenomena, such as financial phenomena. Unfortunately, inference in this class of models may be difficult, because the likelihood function appears as a multivariate integral the size of which is equal to the number of observations multiplied by the size of the latent variables. The complexity of these models is such that generic latent variable techniques, like the Expectation Maximisation algorithm, do not apply, and they almost invariably call for simulation based methods.

The aim of the paper is to survey the literature on simulation based methods for the general class of parametric dynamic latent variable models, with particular attention to applications to SV models.

The paper is organised as follows: section 2 presents the setup, i.e. the general class of dynamic latent variable models and the SV models. Section 3 describes the simulation based approaches and section 4 concludes.

2. Dynamic latent variable models

Dynamic latent variable (DLV) models jointly specify a sequence (y_t) of time dependent variables and a sequence (y_t^*) of partially unobserved variables[1]. Typically, y_t and y_t^* satisfy

$$\left\{ \begin{array}{rcl} y_t^* & = & r_t^*(y^{*t-1}, y^{t-1}, \varepsilon_t^*; \theta)\,, \\ y_t & = & r_t(y^{*t}, y^{t-1}, \varepsilon_t; \theta)\,, \end{array} \right. \tag{1}$$

where the ε_t^*'s and ε_t's are independent white noises[2], with marginal distributions which may depend on θ. Let y^{*t} and y^t denote $(y_1^*, ..., y_t^*)$ and $(y_1, ..., y_t)$, respectively.

There are serious difficulties in computing the likelihood function, in fact we have[3]:

$$f(y^T, y^{*T}; \theta) \;\; = \prod_{t=1}^{T} f(y_t|y^{t-1}, y^{*t}; \theta) f(y_t^*|y^{t-1}, y^{*t-1}; \theta)$$

and the likelihood function is:

$$\ell_T(\theta) \equiv f(y^T; \theta) = \int \prod_{t=1}^{T} f(y_t|y^{t-1}, y^{*t}; \theta) f(y_t^*|y^{t-1}, y^{*t-1}; \theta) \prod_{t=1}^{T} dy_t^* \tag{2}$$

which is an integral whose size is equal to the number of observations multiplied by the dimension of the unobserved variables, and thus it is practically unfeasible.

In the literature of the last forty years, several solutions have been proposed for the estimation of this class of models. They can be subdivided into three groups: exact recursive methods; approximated solutions; simulation based methods. The first category is certainly the more attractive, but it is the narrower. In fact, only in the Gaussian and linear case, with the Kalman filter, and in the Markovian and discrete case, with the Hamilton filter, it is possible to obtain an exact recursive algorithm to compute the likelihood function. Examples of the second type are the extended Kalman filter (Anderson and Moore (1979), Harvey (1989)), Fridman and Harris (1998)), the Gaussian sum filter (Sorenson

[1] For sake of simplicity, exogenous (explanatory) variables are not mentioned explicitly, although they can be introduced at little cost.

[2] In this class of models there is no well-defined reduced form and therefore it is, in general, impossible to compute the conditional p.d.f. $f(y_t/y_{t-1}; \theta)$.

[3] If (y^T, y^{*T}) has p.d.f. with respect to some measure $(\mu \otimes \mu^*)^{\otimes^T}$.

and Alspach (1971)), the numerical integration (Kitagawa (1987)), the Monte Carlo integration (Tanizaki and Mariano (1994), Tanizaki and Mariano (1998)) or the particle filter (Gordon *et al.* (1993)), Kitagawa (1996)), Pitt and Shephard (1999)). The last category of methods is the more time consuming and computing demanding, but it is definitely the more general. In the following we will focus on this type of approach.

21 Stochastic volatility models

For sake of simplicity, in the presentation we focus on stochastic volatility models, which are defined by:

$$\begin{cases} y_t^* & = m_t(y^{*t-1}, y^{t-1}; \theta) + \sigma_t(y^{*t-1}, y^{t-1}; \theta)\varepsilon_t^*, \\ y_t & = \exp(0.5\ y_t^*)\varepsilon_t, \end{cases} \tag{3}$$

where $\{\varepsilon_t^*\}$ and $\{\varepsilon_t\}$ are independent scalar white noises (see Taylor (1994), Ghysels *et al.* (1996), Shephard (1996)). The most classical example of (3) is the basic model:

$$\begin{cases} y_t^* = a + by_{t-1}^* + c\varepsilon_t^*, \\ y_t = \exp(0.5\ y_t^*)\varepsilon_t, \end{cases} \tag{4}$$

where ε_t^* and ε_t are normal independent white noises; the inequality constraints $-1 < b < 1$ and $c > 0$ ensure that y_t is stationary and ergodic.

An attractive feature of specification (4) is the possibility of linearizing the model. By taking logarithms of the squared mean adjusted data we obtain:

$$\begin{cases} y_t^* & = a + by_{t-1}^* + c\varepsilon_t^*, \\ \ln(y_t^2) & = y_t^* + \eta_t, \qquad \eta_t = \ln(\varepsilon_t^2) \end{cases} \tag{5}$$

If the original mean equation disturbance, ε_t, is standard normal, η_t follows the logarithm of a χ_1^2 distribution, whose mean and variance are known to be -1.27 and $\pi^2/2$. Harvey *et al.* (1994) suggested a Quasi-Maximum Likelihood (QML) method of estimating the model based on the Kalman filter. In fact, assuming joint conditional normality of $(\varepsilon_t^*, \eta_t)$, equation (5) represents the measurement and transition equations of a general linear state space model.

3. Simulation based methods

In the last ten years, simulation[4] based methods propose several ways of resolving the inference problem for this class of models (see Billio (1999)). In fact, it is clear that one can easily recursively simulate (path simulations) from the system (2) for any given value of parameters, θ.

A first approach relies on simulation based methods which are relatively simple to implement, but which are less efficient than the maximum likelihood (ML) approach: see, for example, the Simulated Method of Moments (Duffie and Singleton (1993)), the Simulated Pseudo-Maximum Likelihood Method (Laroque and Salanié (1993)), the Indirect Inference Method (Gouriéroux *et al.* (1993)) or the Efficient Method of Moments (Gallant

[4]Simulation techniques make use of sequences of pseudo-random numbers which are generated by a computer procedure.

and Tauchen (1996), Gallant *et al.* (1997)). A second approach considers the problem of the computation (or of the approximation) of the likelihood and then of the ML estimator through importance sampling methods (Danielsson and Richard (1993), Durbin and Koopman (1997)). In a Bayesian framework, a third approach considers Markov Chain Monte Carlo (MCMC) techniques based on the data augmentation principle, which yields samples out of the joint posterior distribution of the latent variable and all model parameters, and allows the parameter estimates and the latent factors dynamics to be obtained (Jacquier *et al.* (1994), Kim *et al.* (1998), Chib *et al.* (2002)). Finally, a fourth approach utilizes MCMC methods in order to compute (or approximate) the maximum likelihood estimator (see the Simulated EM (Shephard (1993) and Geyer (1994), Geyer (1996), Billio *et al.* (1998)).

Let us introduce part of these methods and their application to stochastic volatility models.

31 Non efficient simulation based methods

The so-called Indirect Inference methodology was recently introduced in the literature by Smith (1993), Gouriéroux *et al.* (1993), Gallant and Tauchen (1996), for a simulation based inference on generally intractable structural models through an auxiliary model, conceived as easier to handle. This methodology allows the use of somewhat misspecified auxiliary models, since the simulation process in the well-specified structural model and the calibration of the simulated paths against the observed one through the same auxiliary model will provide an automatic misspecification bias correction. There are several ways of implementing this idea[5].

The original approach is the Indirect Inference (II) Method of Gouriéroux *et al.* (1993). Consider an auxiliary model $f_a(y_t|y^{t-1};\beta)$ for the observed data (for example[6] the general linear state space model represented by equations (5)). Let $\hat{\beta}_T = \mathbf{B}_T(y^T)$ denote the QML estimator of β based on f_a as a function $\mathbf{B}_T(\cdot)$ of the observed data set y^T. The II estimator of structural parameters θ minimizes $[\hat{\beta}_T - \tilde{\beta}_{ST}(\theta)]' W [\hat{\beta} - \tilde{\beta}_{ST}(\theta)]$ where W is a weight matrix and $\tilde{\beta}_{ST}(\theta)$ is the β estimator obtained on a simulated path for a given value of θ (i.e. is given by the binding function $\tilde{\beta}_{ST}(\theta) = \lim_{S\to\infty} \mathbf{B}_{ST}(\tilde{y}^{ST})$, which is approximated by $\mathbf{B}_{ST}(\tilde{y}^{ST})$ for large S). This approach may be very computationally demanding as one needs to evaluate the binding function $\tilde{\beta}_{ST}(\theta)$ for each value of θ appearing in the numerical optimisation algorithm.

The estimator of Gallant and Tauchen (1996) circumvents the need to evaluate the binding function by using the score vector $(\partial/\partial\beta) f_a(y_t|y^{t-1};\beta)$ (score generator) to define the matching conditions. If the auxiliary model $f_a(y_t|y^{t-1};\beta)$ is chosen flexibly with a suitable nonparametric interpretation, then the estimator achieves the asymptotic efficiency of maximum likelihood and has good power properties for detecting misspecification (Gallant and Long (1997), Tauchen (1997), hence the term Efficient Method of Moments (EMM). EMM delivers consistent estimates of the structural parameter vector under weak conditions on the choice of the auxiliary model. However, extrapolating from

[5] For all these methods, it is necessary to recycle the random numbers used in the calculation when θ changes, in order to have good numerical and statistical properties of the estimators based on these simulations.

[6] Another possible auxiliary model is an ARMA(p,q) on the logarithms of the squared mean adjusted data (see Monfardini (1998)).

the Generalised Method of Moments evidence, it is natural to conjecture that the quality of inference may hinge on how well the auxiliary model approximates the salient features of the observed data. This intuition is formalized by Gallant and Long (1997), who show that a judicious selection of the auxiliary model, ensuring that the quasi-scores asymptotically span the true score vector, will result in full asymptotic efficiency. In fact, as the score generator approaches the true conditional density, the estimated covariance matrix for the structural parameter approaches that of maximum likelihood[7].

Andersen *et al.* (1999) perform an extensive Monte Carlo study of EMM estimation of a stochastic volatility model. They examine the sensitivity to the choice of auxiliary model using ARCH, GARCH, and EGARCH models for the score as well as nonparametric extensions. EMM efficiency approaches that of maximum likelihood for larger sample sizes, while inference is sensitive to the choice of auxiliary model in small samples, but robust in larger samples[8].

Another possible approach is the Functional Indirect Inference Method proposed by Billio and Monfort (1999). In order to capture the dynamic features of the structural model, the binding functions are conditional expectations of functions of the endogenous variables given their past values and they are estimated by nonparametric kernel techniques. Unlike the II method, no optimisation step is involved in the computation of the binding function and it is useful when no convenient auxiliary model is available. In spite of the non parametric feature of the approach, the estimator is consistent and its convergence rate is arbitrarily close to the classical parametric one. Moreover, since the asymptotic variance-covariance matrix can be chosen arbitrarily small by increasing the number of conditional expectations considered, the method has potentially good finite sample properties.

However, the Indirect Inference theory crucially depends on the correct specification assumption concerning the structural model. There is now an emerging literature (see, for example, Dridi and Renault (2000) and Dridi (2000)) which focuses on procedures more robust to the structural model specification. In particular, Dridi and Renault (2000) propose an extension to the Indirect Inference methodology to semiparametric settings and show how the Semiparametric Indirect Inference works on basic examples using SV models. Finally, Dridi (2000) proposes a general econometric theory, the Simulated Asymptotic Least Squares, which provides a unifying theory for simulation based inference methods and nests all the above mentioned approaches, both in parametric and semiparametric settings.

32 Simulated Maximum Likelihood methods

As previously seen with equation (2), the likelihood function naturally appears as the expectation of the function $\prod_{t=1}^{T} f(y_t|y^{t-1}, y^{*t}; \theta)$ with respect to the p.d.f. P defined by[9]

[7]This result embodies one of the main advantages of EMM, since it prescribes a systematic approach to the derivation of efficient moment conditions for estimation in a general parametric setting.

[8]Care must be taken, however, to avoid overparameterization of the auxiliary model, as convergence problems may arise if the quasi-score is extended to the point where it begins to fit the purely idiosyncratic noise in the data.

[9]It is important to note that this p.d.f. is neither $f(y^{*T}; \theta)$, except if y_t does not cause y_t^*, nor $f(y^{*T}|y^T; \theta)$.

$\prod_{t=1}^{T} f(y_t^*|y^{t-1}, y^{*t-1}; \theta)$, from which it is easy to recursively draw. Therefore, an unbiased simulator of the whole likelihood function $\ell_T(\theta)$ is $\prod_{t=1}^{T} f(y_t|y^{t-1}, \tilde{y}^{*st}(\theta); \theta)$ where $\tilde{y}^{*st}(\theta)$ are drawn from the auxiliary p.d.f. P.

This basic simulator may be very slow, in the sense that the simulator may have a very large variance and then some accelerating technique is needed. One solution is to consider the general method of importance sampling based on a sequence of conditional p.d.f.'s $q(y_t^*|y^T, y^{*t-1})$. Let us denote this probability distribution by Q and the corresponding expectation by E^Q. We have:

$$\ell_T(\theta) = E^P\left[\prod_{t=1}^{T} f(y_t/y^{t-1}, y^{*t})\right] = E^Q\left[\prod_{t=1}^{T} \frac{f(y_t/y^{t-1}, y^{*t}) f(y_t^*/y^{t-1}, y^{*t-1})}{q(y_t^*/y^T, y^{*t-1})}\right] \quad (6)$$

Therefore, an unbiased simulator of $\ell_T(\theta)$ is:

$$\prod_{t=1}^{T} \frac{f(y_t/y^{t-1}, \tilde{y}^{*st}) f(\tilde{y}_t^{*s}/y^{t-1}, \tilde{y}^{*s(t-1)})}{q(\tilde{y}_t^{*s}/y^T, \tilde{y}^{*s(t-1)})},$$

where $\tilde{y}^{*sT}$ is drawn in Q. The problem is then how to choose the importance function: the natural answer is by reducing the Monte Carlo variance. It is easy to calculate the theoretical optimal choice $f(y^{*T}|y^T; \theta) = \prod_{t=1}^{T} f(y_t^*|y^{*t-1}, y^T; \theta)$ (i.e. the smoothing density of the latent variable), for which one simulation is sufficient, but it is clearly not computable. Then it is possible to: consider the smoothing density of an approximating model; fix a parametric family of importance functions and choose the member that minimizes the Monte Carlo variance (which is eventually computed in an approximated way). For the SV model (4), the first solution is proposed by Sandmann and Koopman (1998) by using as approximating model the linearised version (5). In the aim of the second solution, Danielsson and Richard (1993) propose a sequentially optimized importance sampling, which Danielsson (1994) applies to the SV model. In both cases the SML estimates of model parameters are obtained by numerical optimization of the logarithm of the simulated likelihood[10].

33 The Bayesian approach

In the Bayesian setting there are also serious difficulties. In general, the posterior density $f(\theta|y^T)$ and the posterior expectation of θ cannot be computed in a closed form. Again, this complex setting requires a simulation based approach. The data augmentation principle, which considers the latent variable as nuisance parameters, and the utilisation of Gibbs sampling (Gelfand and Smith (1990), by iterating simulations from $f(y^{*T}|y^T, \theta)$ (*data augmentation step*) and $f(\theta|y^T, y^{*T})$ (*parameter simulation step*), allow simulation from the joint posterior distribution $f(y^{*T}, \theta|y^T)$, derivation of the distribution of interest as the marginal distribution of θ and approximation of the posterior expectation

[10] As for non efficient methods, numerical and statistical accuracy is obtained by recycling the random numbers used in the calculation for each parameter value.

by a sample average. When conditional distributions cannot be directly simulated, the corresponding steps in the Gibbs algorithm are replaced by Metropolis-Hastings steps[11]. Moreover, the prior modeling on the parameters is usually quasi non-informative.

The first Bayesian analysis of the basic SV model was provided by Jacquier *et al.* (1994) where the posterior distribution of the parameters was sampled by MCMC methods using a one-move approach (i.e. the latent variables y_t^* was sampled each at time from $(y_t^*|y^T, y_{-t}^*, a, b, c)$, where y_{-t}^* denotes all the elements of y^{*T} excluding y_t^*). Although this algorithm is conceptually simple, it is not particularly efficient from a simulation perspective, as is shown by Kim *et al.* (1998), who develop an alternative, more efficient, multi-move MCMC algorithm. The efficiency gain in the Kim *et al.* (1998) algorithm arises from the joint sampling of y^{*T} in one block conditioned on everything else in the model. Finally, Chib *et al.* (2002) develop efficient Markov Chain Monte Carlo algorithms for estimating generalized models of SV defined by heavy-tailed Student-t distributions, exogenous variables in the observation and volatility equations and a jump component in the observation equation.

34 A MCMC approach to maximum likelihood estimation

Although the Bayesian approach is straightforward to state, it requires the elicitation of a prior, which is often regarded by some econometricians as being difficult in dynamic models. Even if this is not an insurmountable problem, alternatives are available which allow us to perform maximum likelihood estimation using MCMC methods.

The first possibility is the Simulated Expectation Maximisation (SEM) algorithm proposed by Shephard (1993). The EM algorithm exploits the following decomposition of the log likelihood function:

$$\begin{aligned} \log f(y^T;\theta) &= \log f(y^T, y^{*T};\theta) - \log f(y^{*T}|y^T,;\theta) \\ &= E\left[\log f(y^T, y^{*T};\theta)|y^T\right] - E\left[\log f(y^{*T}|y^T,;\theta)|y^T\right] \end{aligned} \tag{7}$$

and iterates:

$$\theta^{i+1} = \arg\max_{\theta} E_{\theta^i}\left[\log f(y^T, y^{*T};\theta)|y^T\right] \tag{8}$$

This is an increasing algorithm such that the sequence θ^i converges to the ML estimator. The problem is that, although $\log f(y^T, y^{*T};\theta)$ has in general a closed form, the same is not true for its conditional expectation. In the SEM algorithm this expectation is replaced by an approximation based on simulations. Thus, the problem is now to be able to draw in the conditional distribution of y^{*T} given y^T and θ. Shephard (1993), in the context of a nonlinear state space model, uses the Hastings-Metropolis algorithm to solve this problem and applies it to the SV model.

Another possible approach is the Simulated Likelihood Ratio (SLR) method proposed by Billio *et al.* (1998). The general principle is:

$$\frac{f(y^T;\theta)}{f(y^T;\bar{\theta})} = E_{\bar{\theta}}\left[\frac{f(y^{*T}, y^T;\theta)}{f(y^{*T}, y^T;\bar{\theta})}\Big|y^T\right] \tag{9}$$

where $\bar{\theta}$ is an arbitrary fixed value of the parameters. Obviously,

$$\arg\max_{\theta} f(y^T;\theta) = \arg\max_{\theta} \frac{f(y^T;\theta)}{f(y^T;\bar{\theta})}$$

[11] Such hybrid algorithms are validated in Tierney (1994).

and with $y^{*T}(s)$, $s = 1, ..., S$, simulated paths in the conditional distribution $f(y^{*T}|y^T, \bar{\theta})$, the SLR method amounts to maximising:

$$\frac{1}{S}\sum_{s=1}^{S}\frac{f(y^{*T}(s), y^T; \theta)}{f(y^{*T}(s), y^T; \bar{\theta})}$$

with respect to θ. The method can be implemented by simulating in the conditional distribution[12] $f(y^{*T}|y^T; \bar{\theta})$. As already noted, it is impossible to simulate directly in this distribution, thus a Hastings-Metropolis approach is suggested.

Contrary to the SEM approach, the SLR method allows for the computation of the likelihood surface and then of likelihood ratio test statistics; it needs only one optimisation run and not a sequence of optimisations; it is possible to store the simulated paths, and then only one simulation run is required. Moreover, as the simulation is made for only one value of the parameter, the objective function will be smooth with respect to θ, even if simulations involves rejection methods.

4. Conclusions

We survey the literature on simulation based methods for the general class of DLV models. In practice, the choice between these different simulation based approaches depends on several criteria, such as efficiency and computing time. Unfortunately, in general there is a trade off between these criteria.

Methods like SML and SLR, have several advantages in the estimation of DLV models. Since they are likelihood method, the classical theory of maximum likelihood carries over to the simulated case and standard likelihood ratio tests can be constructed. MCMC based approaches are certainly more time consuming but also allow estimation of the latent variable dynamics by simulating from the smoothing/posterior distribution of y^{*T}.

Following Billio *et al.* (2001), eventually a two step estimator, which is asymptotically efficient and simple to implement, could be considered: the first step is based on a consistent estimator easy to compute, while the second step is a single iteration of a Newton Raphson optimisation, where the first and second derivatives are obtained by a simulation approach based on relationships between the likelihood function, which is not computable, and the complete likelihood function $f(y^{*T}, y^T; \theta)$, which is in general easily computable. The method is also applied to the SV model (4).

References

Andersen T., Chung H. and Sorensen B. (1999) Efficient method of moments estimation of a stochastic volatility model: A monte carlo study, *Journal of Econometrics*, 91, 61–87.

Anderson B. and Moore J. (1979) *Optimal Filtering*, Prentice Hall, New York.

Bera A. and Higgins M. (1993) Arch models: properties, estimation and testing, *Journal of Economic Surveys*, 7, 305–366.

[12]The resulting Monte Carlo approximation of (9) could be only locally good around $\bar{\theta}$ and so Geyer (1996) suggests updating $\bar{\theta}$ to the maximiser of the Monte Carlo likelihood and repeat the Monte Carlo procedure using the new $\bar{\theta}$. By updating $\bar{\theta}$ a few times, one should obtain better approximations of the relative likelihood function near the true maximum likelihood estimate.

Billio M. (1999) *Méthodes fondées sur des simulations pour l'inférence dans les modèles espace-état non linéaires*, Ph.D. thesis, Paris Dauphine.
Billio M. and Monfort A. (1999) Functional indirect inference, Document de Travail du CREST 9901, Paris.
Billio M., Monfort A. and Robert C. (1998) The simulated likelihood ratio (slr) method, Document de Travail du CREST 9828, Paris.
Billio M., Monfort A. and Robert C. (2001) The simulated newton raphson method, Document de travail, CREST Paris.
Bollerslev T., Chow R. and Kroner K. (1992) Arch modelling in finance: a review of the theory and empirical evidence, *Journal of Econometrics*, 52, 5–59.
Bollerslev T., Engle R. and Nelson D. (1993) Arch models, in: *Handbook of Econometrics*, Engle R. and McFadden D., eds., North-Holland, Amsterdam, volume IV.
Chib S., Nardari F. and Shephard N. (2002) Markov chain monte carlo methods for stochastic volatility models, *Journal of Econometrics*, forthcoming.
Danielsson J. (1994) Stochastic volatility in asset prices: estimation with simulated maximum likelihood, *Journal of Econometrics*, 61, 375–400.
Danielsson J. and Richard J. (1993) Accelerated gaussian importance sampler with application do dynamic latent variable models, *Journal of Applied Econometrics*, 8, 153–173.
Dridi R. (2000) Simulated asymptotic least squares theory, Working Paper 396, London School of Economics - Suntory Toyota, Econometrics.
Dridi R. and Renault E. (2000) Semi-parametric indirect inference, Working Paper 392, London School of Economics - Suntory Toyota, Econometrics.
Duffie J. and Singleton K. (1993) Simulated moments estimation of markov models of asset prices, *Econometrica*, 61, 929–952.
Durbin J. and Koopman S. (1997) Monte carlo maximum likelihood estimation for non-gaussian state space models, *Biometrika*, 84, 669–684.
Fridman M. and Harris L. (1998) A maximum likelihood approach for non-gaussian stochastic volatility models, *Journal of Business and Economic Statistics*, 16, 284–291.
Gallant A., Hsieh D. and Tauchen G. (1997) Estimation of stochastic volatility models with diagnostics, *Journal of Econometrics*, 81, 159–192.
Gallant A. and Long J. (1997) Estimating stochastic differential equations efficiently by minimum chi-square, *Biometrika*, 84, 125–141.
Gallant A. and Tauchen G. (1996) Which moments to match?, *Econometric Theory*, 12, 657–681.
Gelfand A. and Smith A. (1990) Sampling based approaches to calculating marginal densities, *Journal of the American Statistical Association*, 85, 398–409.
Geyer C. (1994) On the convergence of monte carlo maximum likelihood calculations, *Journal of the Royal Statistical Society*, B 65, 261–274.
Geyer C. (1996) Estimation and optimization of functions, in: *Markov Chain Monte Carlo in Practice*, Gilks W., Richardson S. and Spiegelhalter D., eds., Chapman and Hall, London, 241–258.
Ghysels E., Harvey A. and Renault E. (1996) Stochastic volatility, in: *Statistical Methods in Finance*, Rao C. and Maddala G., eds., North-Holland, Amsterdam.
Gordon N., Salmond D. and Smith A. (1993) A novel approach to non-linear and non-gaussian bayesian state estimation, *IEE Proceedings*, F140, 107–133.
Gouriéroux C., Monfort A. and Renault E. (1993) Indirect inference, *Journal of Applied*

Econometrics, 8, 85–118.

Harvey A. (1989) *Forecasting, Structural Time Series Models and the Kalman Filter*, Cambridge University Press.

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

Jacquier E., Polson N. and Rossi P. (1994) Bayesian analysis of stochastic volatility models, *Journal of Business and Economic Statistics*, 12, 371–417.

Kim S., Shephard N. and Chib S. (1998) Stochastic volatility: likelihood inference and comparison with arch models, *Review of Economic Studies*, 65, 361–393.

Kitagawa G. (1987) Non-gaussian state space modeling of nonstationary time series (with discussion), *Journal of the American Statistical Association*, 82, 400, 1032–1063.

Kitagawa G. (1996) Monte carlo filter and smoother for non-gaussian non linear state space models, *Journal of Computational and Graphical Statistics*, 5, 1–25.

Laroque G. and Salanié B. (1993) Simulation based estimation of models with lagged latent variables, *Journal of Applied Econometrics*, 8, 119–133.

Monfardini C. (1998) Estimating stochastic volatility models through indirect inference, *Econometrics Journal*, 1, 113–128.

Pitt M. and Shephard N. (1999) Filtering via simulation: auxiliary particle filter, *Journal of the American Statistical Association*, 94, 590-599.

Sandmann G. and Koopman S. (1998) Estimation of stochastic volatility models via monte carlo maximum likelihood, *Journal of Econometrics*, 87, 271–301.

Shephard N. (1993) Fitting non-linear time series models, with applications to stochastic variance models, *Journal of Applied Econometrics*, 8, 135–152.

Shephard N. (1996) Statistical aspects of arch and stochastic volatility, in: *Time Series Models*, Cox D., Hinkley D. and Barndoff-Nielsen O., eds., Chapman and Hall, London.

Sorenson H. and Alspach D. (1971) Recursive bayesian estimation using gaussian sums, *Automatica*, 7, 465–479.

Tanizaki H. and Mariano R. (1994) Prediction, filtering and smoothing in non-linear and non-normal cases using monte carlo integration, *Journal of Applied Econometrics*, 9, 163–179.

Tanizaki H. and Mariano R. (1998) Nonlinear and nonnormal state-space modeling with monte carlo stochastic simulation, *Journal of Econometrics*, 83, 1-2, 263–290.

Tauchen G. (1997) New minimum chi-square methods in empirical finance, in: *Advances in Economics and Econometrics: Theory and Applications, Seventh World Congress*, Kreps D. and Wallis K., eds., Cambridge University Press, Cambridge.

Taylor S. (1994) Modelling stochastic volatility: a review and comparative study, *Mathematical Finance*, 4, 183–204.

Tierney L. (1994) Markov chains for exploring posterior distributions (with discussion), *Annals of Statistics*, 22, 1701–1786.