July 15, 1998

Simulation based methods of moments in empirical finance

Roman Liesenfeld

Universität Tübingen,
Wirtschaftswissenschaftliche Fakultät
Abteilung Statitik und Ökonometrie
Mohlstrasse 36
D–72074 Tübingen, Germany

Jörg Breitung

Humboldt Universität zu Berlin,
Institut für Statistik und Ökonometrie,
Spandauer Strasse 1,
D–10178 Berlin, Germany.

We thank Laszlo Matyas for helpful comments. This paper was printed using funds made available by the Deutsche Forschungsgemeinschaft.

SIMULATION BASED METHODS OF MOMENTS IN EMPIRICAL FINANCE

Roman Liesenfeld and Jörg Breitung

1 Introduction

The estimation of unknown parameters generally involves optimizing a criterion function based on the likelihood function or a set of moment restrictions. Unfortunately, for many econometric models the likelihood function and/or the relevant moment restrictions do not have a tractable analytical form in terms of the unkown parameters rendering the estimation by maximum likelihood (ML) or the generalized method of moments (GMM) infeasible. This estimation problem typically arises in situations, where unobservable variables enter the model nonlinearly, leading to multiple integrals in the criterion function, which cannot be evaluated by standard integration methods. Prominent examples for such econometric models in the field of financial econometrics are continous-time models for the evolution of stock prices or interest rates and discrete-time stochastic volatility models for the dynamics in the volatility of financial data.

Until recently, estimation problems due to the lack of some tractable criterion function were often circumvented by using approximations of the model with criterion functions simple enough to evaluate. An alternative solution in such situations that has received increased attention over the last few years, favoured by the permanently growing computer power, are estimation procedures that use Monte Carlo simulation methods to compute an otherwise intractable criterion function[1]. Seminal for the development of this type of estimation procedures were the contributions of McFadden [1989] and Pakes and Pollard [1989] who introduced the Method of Simulated Moments (MSM) for a cross-sectional context. This approach, which was extended to time-series applications by Lee and Ingram [1991] and Duffie and Singleton [1993], modifies the traditional GMM estimator by using moments computed from simulated data of the model rather than the analytical moments. Like the GMM estimator, the MSM estimator is consistent and asymptotically normal when the number of observations tends to infinity, and is asymptotically equivalent to GMM if the number of simulations approaches infinity. However, in a fully parametric

[1]It is worth noting that Monte Carlo simulation methods themselves have already been used for a long time in the Bayesian econometrics for evaluating posterior distributions, see e.g. Kloek and van Dijk [1978].

model one can expect that MSM, just as GMM, is inefficient relative to procedures based on the likelihood due to the arbitrary choice of moment restrictions. This issue is addressed by the indirect inference estimators proposed by Gouriéroux, Monfort and Renault [1993], Bansal, Gallant, Hussey and Tauchen [1993, 1995] and Gallant and Tauchen [1996a]. These approaches which represent extensions of MSM introduce an auxiliary model in order to estimate the parameters of the model of interest. The first version of indirect inference as proposed by Gouriéroux, Monfort and Renault [1993] employs the parameters of the auxiliary model to define the GMM criterion function, whereas in the second version as suggested by Bansal, Gallant, Hussey and Tauchen [1993, 1995] and Gallant and Tauchen [1996a] the scores of the auxiliary model generate the moment restrictions used in the GMM criterion function. Since in both procedures the GMM criterion is an intractable function in terms of the parameters of interest, simulations are used to evaluate it. Both indirect inference estimators are consistent and asymptotically normal as the number of observations tends to infinity and approach the fully efficient estimator if the auxiliary model is appropriately chosen. Specifically, if the auxiliary model is based on the semi-nonparametric model of Gallant and Nychka [1987], as proposed by Gallant and Tauchen [1996a], one may hope that the loss of efficiency of the indirect inference estimator is small.

The purpose of this chapter is to give a selective review of MSM and indirect inference which represent simulation based methods of moments, and to discuss their applications to models for financial data. Besides these moment based simulation approaches, a variety of other simulation estimators are proposed in financial econometrics including simulated maximum likelihood (Danielsson and Richard [1993] and Richard and Zhang [1997]), and Markov-Chain Monte-Carlo procedures (Jacquier, Polson and Rossi [1994] and Kim, Shephard and Chip [1996]). Surveys on these likelihood-oriented simulation methods are given by Ghysels, Harvey and Renault [1996] and Shephard [1996].

This chapter is organized as follows. In section 2 we outline the estimation context and give some examples. The MSM and the indirect inference estimator are discussed in sections 3 and 4, respectively. Section 5 reviews the semi-nonparametric auxiliary model and in section 6 we address selected practical issues concerning the application of these estimators. Section 7 concludes.

2 General Setup and Applications

Let y_t, $t = 1, \ldots, T$ denote an n-dimensional vector of observable dependent variables and x_t is a k-dimensional vector of observable strongly exogenous variables. For expositional convenience it is assumed that y_t and x_t are stationary. The nonlinear dynamic model is characterized by the conditional density $h_0(y_t|z_t)$, where $z_t = [y'_{t-1}, \ldots, y'_1, y'_0, x'_t, \ldots, x'_1]'$ is the vector of conditioning variables and the initial conditions are represented by y_0. We want to estimate the p-dimensional parameter vector θ from the model

$\mathcal{M} := \{h(y_t|z_t;\theta), \theta \in \Theta\}$, where Θ denotes the parameter space. The true value θ_0 is a unique value of θ such that $h_0(y_t|z_t) = h(y_t|z_t;\theta_0)$. In the following, we use $h(\cdot)$ as a generic notation for all density functions.

The estimation of θ_0 is generally based on the likelihood function $L_T(\theta) = \prod_{t=1}^T h(y_t|z_t;\theta)$ or on moment restrictions based on a set of moments such as $E[y_t|z_t]$ or $E[y_t y_t'|z_t]$. Here we are interested in cases where the likelihood function or the relevant moments have an intractable form, rendering ML estimation or method of moments estimation infeasible. Nevertheless, we assume that the model allows us to simulate values of the process $\{y_t\}$ given some value of the parameter vector θ and the initial conditions y_0.

For dynamic models with lagged endogenous variables two different simulation schemes may be possible (see, Gouriéroux and Monfort [1996, p. 17]). If the model admits a *reduced form* $y_t = \varrho(z_t, \varepsilon_t; \theta)$, where ε_t is an error term stochastically independent of z_t and with a known distribution independent of θ, simulated random variables $y_t^{(r)}(\theta)$, $(r = 1, \ldots, R)$ from the distribution $h(y_t|z_t;\theta)$ can be generated as follows. Artificial random variables $\varepsilon_t^{(r)}$ from the distribution of ε_t are generated and used to calculate

$$y_t^{(r)}(\theta) = \varrho(z_t, \varepsilon_t^{(r)}; \theta)$$

for the observed values of $z_t = [y_{t-1}', \ldots, y_1', y_0', x_t', \ldots, x_1']'$ and some value of the parameter vector θ. For a large number of replications R, the empirical distribution of the simulated values $y_t^{(r)}(\theta)$, $(r = 1, \ldots, R)$ approximates the conditional distribution $h(y_t|z_t;\theta)$ for every t. Since the simulations are performed conditionally on the observed lagged endogenous variables, this simulation scheme is called *conditional simulations*. The second approach, termed *path simulations*, is to generate simulated values of y_t conditionally on simulated lagged endogenous variables, i.e. conditionally on $z_t^{(r)}(\theta) = [y_{t-1}^{(r)}(\theta)', \ldots, y_1^{(r)}(\theta)', y_0', x_t', \ldots, x_1']'$, using some kind of recursion. For large R, the empirical joint distribution of $y_1^{(r)}(\theta), \ldots, y_T^{(r)}(\theta)$, $(r = 1, \ldots, R)$ approximates the distribution $h(y_1, \ldots, y_T|x_1, \ldots, x_T;\theta)$.

In order to motivate the estimation context addressed here, we discuss in the following some examples from financial econometrics[2].

Example 1: **Discrete-time stochastic volatility model**

The standard discrete time stochastic volatility (SV) model proposed by Taylor [1986, 1994] and others is given by

$$y_t = \exp\{w_t^*/2\} u_t \tag{1}$$

$$w_t^* = \gamma + \delta w_{t-1}^* + \nu \eta_t \,, \qquad t = 1, \ldots, T \,, \tag{2}$$

where y_t is the observable return of a financial asset and w_t^* is the unobservable log volatility. The error processes u_t and η_t are mutually and serially independent with

[2] As in most applications in financial econometrics, a time series framework is used. Examples for cross-sectional applications are given by Gouriéroux and Monfort [1993, 1996] and Stern [1997].

known distributions. In accounting for the observed autocorrelation in the variance of financial time series, this SV model represents an alternative to the ARCH and GARCH specifications proposed by Engle [1982] and Bollerslev [1986]. Since the latent log volatility w_t^* enters the model in a nonlinear fashion, the conditional density $h(y_t|z_t;\theta)$ with $\theta = [\gamma, \delta, \nu]'$ and $z_t = [y_{t-1}, \dots, y_1, y_0]'$ does not have an explicit analytical form. To obtain the (marginal) likelihood function associated with the observable variables, the latent variables are "integrated out" from the joint distribution of $y_1, \dots, y_T, w_1^*, \dots, w_T^*$ denoted by $h(y_1, \dots, y_T, w_1^*, \dots, w_T^*|\theta)$. This distribution can be factorized as $h(y_1, \dots, y_T, w_1^*, \dots, w_T^*|\theta) = \prod_{t=1}^T h(y_t|w_t^*;\theta)h(w_t^*|w_{t-1}^*;\theta)$, where $h(y_t|w_t^*;\theta)$ is the conditional density of the returns given the log volatility and $h(w_t^*|w_{t-1}^*;\theta)$ denotes the conditional density of the log volatility given its past value. Hence, for a given initial value of the log volatility w_0^* the marginal likelihood has the following form

$$L_T(\theta) = \int \cdots \int \prod_{t=1}^{T} h(y_t|w_t^*;\theta)h(w_t^*|w_{t-1}^*;\theta)\, dw_1^* \dots dw_T^* \, .$$

For this T-dimensional integral no closed-form solution exists, nor can standard numerical methods be applied to evaluate it making ML estimation infeasible. Furthermore, even if the standard SV model can be estimated by GMM using unconditional moments such as $E[|y_t|]$, $E[y_t^2]$ or $E[y_t^2 y_{t-1}^2]$, GMM is relatively inefficient, especially, if the persistence parameter δ is close to one (see, e.g. Jacquier, Polson and Rossi [1994] and Andersen and Sørensen [1996]). However, the SV model given by (1) and (2) defines a simple data generating process which allows to generate values from the joint distribution $h(y_1, \dots, y_T|\theta)$ implied by the model using path simulations. Note though that conditional simulations from $h(y_t|y_{t-1}, \dots, y_0;\theta)$ appear to be infeasible since the SV model does not admit an explicit expression of the reduced form in terms of lagged endogenous variables $y_t = \varrho(y_{t-1}, \dots, y_0, \varepsilon_t;\theta)$.

Example 2: **Stochastic differential equations**

Consider the following scalar stochastic differential equation:

$$dv_t = a(v_t, \theta)dt + b(v_t, \theta)dW_t \, , \qquad 0 \le t \le N \, , \tag{3}$$

where $a(v_t, \theta)$ and $b(v_t, \theta)$ are the drift and the diffusion function, respectively, and W_t is a Brownian motion. Such continuous-time processes are often used to model stock prices and interest rates. However, in practice the variables are observable only at some discrete (possibly equispaced) points. Hence, the observable variables y_t, $(t = 1, \dots, T)$ are given by $y_t = v_{t \cdot \Delta}$ for some $\Delta > 0$, where the time interval between two observations is $[t, t + \Delta)$. For arbitrary drift and diffusion functions, the distribution of the observable variables generally does not have a closed form expression. A closed-form can be obtained only for some special drift and diffusion functions. As an example, consider the square root process proposed by Cox, Ingersoll and Ross [1985] to model the evolution of interest rates:

$$dv_t = (\alpha_0 + \alpha_1 v_t)dt + \beta_0 \sqrt{v_t} dW_t \ .$$

This stochastic differential equation implies a joint distribution of the observable variables $y_1, \ldots, y_T$ given by $\prod_{t=1}^T h(y_t|y_{t-1};\theta)$, where $h(y_t|y_{t-1};\theta)$ is a non-central χ^2-distribution. However, for more complicated specifications the conditional density $h(y_t|y_{t-1};\theta)$ and, in general its moments, do not have a tractable form since $h(y_t|y_{t-1};\theta)$ appears as a multiple integral (see, e.g. Gouriéroux and Monfort [1996, p. 10f]). This motivates the use of alternatives to standard ML and GMM estimators. An example for a specification with an intractable density $h(y_t|y_{t-1};\theta)$ is the following generalisation of the Cox-Ingersoll-Ross model:

$$dv_t = (\alpha_0 + \alpha_1 v_t)dt + \beta_0 {v_t}^{\beta_1} dW_t \ ,$$

which is proposed by Chan, Karolyi, Longstaff and Sanders [1992].

To simulate values of the observable discrete-time variables according to a continous-time model, one can use a discrete-time approximation, for example, the Euler approximation. If the time interval between two observations $[t, t+\Delta)$ is divided into subintervals of length τ, the corresponding Euler approximation of (3) becomes

$$v_{t+k\tau} = v_{t+(k-1)\tau} + \tau \, a(v_{t+(k-1)\tau}, \theta) + \sqrt{\tau}\, b(v_{t+(k-1)\tau}, \theta)\eta_{t,k} \ , \qquad k = 1, 2, \ldots \ ,$$

where $\eta_{t,k}$ is an $i.i.d.N(0,1)$ random variable. If the time interval τ is sufficiently small, this approximation can be used to simulate values from $h(y_1, \ldots, y_T|\theta)$ according to $y_t = \varrho(y_{t-1}, \varepsilon_t; \theta)$, where $\varepsilon_t = [\eta_{t,1}, \ldots, \eta_{t,1/\tau}]'$ is the vector of error terms.

The common feature of Examples (1) and (2) is that (partially) unobservable processes enter the model nonlinearly, making criterion functions commonly used for estimation intractable. Further examples for this estimation context in financial econometrics are the continous-time stochastic volatility models of Hull and White [1987] and Chesney and Scott [1989], the market microstructure model proposed by Forster and Viswanathan [1995], the dynamic equilibrium model for asset prices estimated by Bansal, Gallant, Hussey and Tauchen [1995] and the multifactor latent ARCH models of Diebold and Nerlove [1989] and Engle, Ng and Rothschild [1990].

3 The Method of Simulated Moments (MSM)

Consider a dynamic model with a well defined reduced form $y_t = \varrho(z_t, \varepsilon_t; \theta)$ allowing us to simulate values of y_t from $h(y_t|z_t, \theta)$ for observed values of the conditioning variables $z_t =$

$[y'_{t-1}, \ldots, y'_1, y'_0, x'_t, x'_{t-1}, \ldots, x'_1]'$. We will focus on the m-dimensional moment function of the form

$$\varphi(y_t, z_t; \theta) = s(y_t, z_t) - \sigma(z_t; \theta) \ , \tag{4}$$

with $m \geq p$ and where $s(y_t, z_t)$ is a function on the data and $\sigma(z_t; \theta)$ is the theoretical counterpart defined as

$$\sigma(z_t; \theta) = E_\theta[s(y_t, z_t)|z_t].$$

Here $E_\theta(\cdot|z_t)$ indicates that the expectation is computed with respect to the density $h(y_t|z_t; \theta)$ and $\sigma(z_t; \theta)$ represents conditional moments as, for example, $E_\theta(y_t|z_t)$ or $E_\theta(y_t y'_t|z_t)$. The index is dropped if the expectation is taken with respect to the true process, i.e., $E \equiv E_{\theta_0}$. We assume that for θ_0 the empirical moment condition

$$E[\varphi(y_t, z_t; \theta_0)|z_t] = 0 \quad \text{ for all } t$$

is satisfied. Let $f(y_t, z_t; \theta_0) = B(z_t)'\varphi(y_t, z_t; \theta_0)$, where $B(z_t)$ is some nonlinear matrix function on z_t, then the corresponding set of unconditional moment restrictions is given by (see, e.g. Newey [1993])

$$E[f(y_t, z_t; \theta_0)] = 0 \quad \text{ for all } t \ .$$

If the expression $\sigma(z_t; \theta)$ cannot be computed analytically, it may be approximated using simulation methods. Since $\sigma(z_t; \theta)$ is the expectation value of $s(y_t, z_t)$ evaluated with respect to $h(y_t|z_t; \theta)$, a natural unbiased estimator for $\sigma(z_t; \theta)$ is given by

$$\hat{\sigma}_R(z_t; \theta) = \frac{1}{R} \sum_{r=1}^{R} s[y_t^{(r)}(\theta), z_t] \ , \tag{5}$$

where $y_t^{(r)}(\theta)$, $(r = 1, \ldots, R)$ are simulated random variables drawn from the distribution $h(y_t|z_t; \theta)$ for the observed values of z_t. The natural estimator of $\sigma(z_t; \theta)$ given in equation (5) results from sampling data using $h(y_t|z_t, \theta)$. However, this estimator may have undesirable properties. For example, it may not be differentiable with respect to θ or it may have a large variance. Therefore, alternative methods of estimating $\sigma(z_t, \theta)$ such as importance sampling procedures were proposed to obtain an estimator with improved properties (see, e.g. Gouriéroux and Monfort [1993] and Stern [1997]).

If the natural Monte Carlo estimator (5) is used to estimate the moment restrictions the method of simulated moments (MSM) estimator for θ_0 is obtained by minimizing the criterion function

$$\hat{\theta}^R_{MSM} = \underset{\theta}{\operatorname{argmin}} \left[\sum_{t=1}^{T} f_R(y_t, z_t; \theta) \right]' A \left[\sum_{t=1}^{T} f_R(y_t, z_t; \theta) \right] \tag{6}$$

where

$$f_R(y_t, z_t; \theta) = B(z_t)'[s(y_t, z_t) - \hat{\sigma}_R(z_t; \theta)]$$

and A denotes an appropriately chosen positive definite weight matrix. If the simulation sample size R tends to infinity, $\hat{\sigma}_R(z_t;\theta)$ converges almost surely to $E_\theta[s(y_t,z_t)|z_t]$ and the MSM estimator equals the corresponding GMM estimator. However, as the sample size T tends to infinity, the MSM estimator is consistent for any fixed $R \geq 1$ as long as different random draws are used across t (cf. McFadden [1989]). The reason for this is that for the estimator $\hat{\theta}^R_{MSM}$ the simulation error is "averaged out" by using the *mean* of $\hat{\sigma}_R(z_t;\theta)$, $(t=1,\ldots,T)$.

The fact that the MSM estimator is consistent for any $R \geq 1$ should not be taken as an indication that R is irrelevant for the asymptotic properties of $\hat{\theta}^R_{MSM}$ as $T \to \infty$. This becomes clear from considering the asymptotic distribution of the MSM estimator, which results as $T^{1/2}(\hat{\theta}^R_{MSM} - \theta_0) \xrightarrow{d} N(0, \text{avar}(\hat{\theta}^R_{MSM}))$. The asymptotic covariance matrix of $\hat{\theta}^R_{MSM}$, as it results from the fact that $\{f(y_t,z_t;\theta_0)\}$ is by construction serially uncorrelated with identical distributions, has the form (see, Gouriéroux and Monfort [1996, p. 29])

$$\text{avar}(\hat{\theta}^R_{MSM}) = \Sigma_1^{-1}\Sigma_2\Sigma_1^{-1} + \frac{1}{R}\Sigma_1^{-1}D'A\,\text{var}[f(y_t^{(r)}(\theta_0),z_t;\theta_0)]\,AD\Sigma_1^{-1}\ , \qquad (7)$$

where

$$\begin{aligned} D &= E\left[B(z_t)'\frac{\partial\sigma(z_t;\theta_0)}{\partial\theta'}\right] \\ \Sigma_1 &= D'AD \\ \Sigma_2 &= D'A\,\text{var}[f(y_t,z_t;\theta_0)]\,AD\ . \end{aligned}$$

The lower bound of the asymptotic covariance matrix obtained for $R \to \infty$ is given by the asymptotic covariance of the corresponding GMM estimator $\Sigma_1^{-1}\Sigma_2\Sigma_1^{-1}$. However, the asymptotic covariance matrix of the MSM estimator contains, compared to that of the GMM estimator, an additional component which is due to the variation in the Monte Carlo estimates of the moment restrictions. This additional Monte Carlo sampling variance vanishes as the simulation sample size increases and the MSM estimator attains the efficiency of the corresponding GMM estimator.

The asymptotic optimal weight matrix which minimizes the asymptotic covariance of $\hat{\theta}^R_{MSM}$ for a given set of moment restrictions is:

$$A_0 = \left(\text{var}[f(y_t,z_t;\theta_0)] + \frac{1}{R}\text{var}[f(y_t^{(r)}(\theta_0),z_t;\theta_0)]\right)^{-1}\ .$$

For this optimal choice of the weight matrix the asymptotic covariance matrix of the MSM estimator is $\text{avar}(\hat{\theta}^R_{MSM}) = [D'A_0D]^{-1}$.

The MSM estimator given above is based on conditional moments of the function $s(y_t,z_t)$ given $z_t = [y'_{t-1},\ldots,y'_1,y'_0,x'_t,x'_{t-1},\ldots,x'_1]'$. A necessary requirement for using such conditional moments for MSM estimation, is that the model admits a well defined reduced form $y_t = \varrho(z_t,\varepsilon_t;\theta)$ in terms of exogenous and lagged endogenous variables in order to perform conditional simulations from $h(y_t|z_t;\theta)$. These conditional simulations are necessary to obtain unbiased estimates for $\sigma(z_t, \theta)$ based on estimators such as that given in equation (5). However, for models which include unobservable variables nonlinearly as, for instance, the SV model in Example 1, a reduced form in terms of lagged endogenous variables is generally not available. Hence, in such cases the MSM estimation based on conditional moments given lagged endogeneous variables is infeasible. In those situations, we may use restrictions based on moments conditional only on exogenous variables or for pure time series models restrictions derived from unconditional moments. Such an MSM approach for pure time series applications has been proposed by Duffie and Singleton [1993], and has been applied by Forster and Viswanathan [1995] and Gennotte and Marsh [1993] for estimating a market microstructure model and a dynamic asset pricing model, respectively.

This unconditional version of the MSM estimator is based on a m-dimensional moment function of the form

$$f(y_t, \ldots, y_{t-l}; \theta) = s(y_t, \ldots, y_{t-l}) - \sigma(\theta) \ , \qquad t = 1, \ldots, T \ , \tag{8}$$

where $\sigma(\theta)$ represents the unconditional expectation value $E_\theta[s(y_t, \ldots, y_{t-l})]$. The corresponding set of moment restrictions is given by $E[f(y_t, \ldots, y_{t-l}; \theta_0)] = 0$. These restrictions include moments such as $E_\theta(y_t)$ and $E_\theta(y_t y_t')$ as well as cross order moments of the form $E_\theta(y_t y_{t-i}')$. If $y_t(\theta)$, $(t = 1, \ldots, R)$ denotes a simulated path from the distribution $h(y_1, \ldots, y_R|\theta)$ implied by the model, the MSM estimator based on these unconditional moments is obtained by

$$\hat{\theta}^R_{MSM} = \underset{\theta}{\operatorname{argmin}} \left[\frac{1}{T} \sum_{t=1}^{T} s(y_t, \ldots, y_{t-l}) - \hat{\sigma}_R(\theta) \right]' A \left[\frac{1}{T} \sum_{t=1}^{T} s(y_t, \ldots, y_{t-l}) - \hat{\sigma}_R(\theta) \right] \ ,$$

where

$$\hat{\sigma}_R(\theta) = \frac{1}{R} \sum_{t=1}^{R} s[y_t(\theta), \ldots, y_{t-l}(\theta)] \ .$$

The matrix A denotes the weight matrix and $\hat{\sigma}_R(\theta)$ is an unbiased Monte Carlo estimator for $\sigma(\theta)$. As the moment function (8) derived from the dynamic model $h(y_t|z_t; \theta)$ is expected to be serially correlated the asymptotic optimal weight matrix is given by

$$A_0 = (\lim_{T \to \infty} [\mathrm{var}\{\frac{1}{\sqrt{T}} \sum_{t=1}^{T} s(y_t, \ldots, y_{t-l})\}])^{-1} \ .$$

Since $s(y_t, \ldots, y_{t-l})$ is independent of the parameter θ and independent of the simulated values $y_t(\theta)$, the matrix A_0 can be estimated by procedures discussed in chapter 3. However, like the MSM estimator based on conditional moments, the MSM estimator using unconditional moments is consistent and asymptotically normally distributed as T tends to infinity. Specifically, the asymptotic distribution for the optimal weight matrix A_0 results as $T^{1/2}(\hat{\theta}^R_{MSM} - \theta_0) \xrightarrow{d} N(0, \mathrm{avar}(\hat{\theta}^R_{MSM}))$, with $\mathrm{avar}(\hat{\theta}^R_{MSM}) = [1 + (1/R)][D' A_0 D]^{-1}$ and $D = E[\partial \sigma(\theta_0)/\partial \theta']$ (see, Duffie and Singleton [1993]). The factor $[1 + (1/R)]$ in the asymptotic variance accounts for the additional variation of the estimator due to the Monte Carlo sampling variance which vanishes as R tends to infinity.

4 Indirect Inference Estimator

The MSM approach is used to optimize a GMM criterion function, which is too complicated to be computed analytically. Another possible approach as proposed by Gouriéroux, Monfort and Renault [1993], is to use a criterion function derived from an *auxiliary*, possibly misspecified model and to recover the *structural* parameters of the original model from the parameter estimates of the misspecified model. Unfortunately, the relationship between the auxiliary and the structural model is too complicated to admit an explicit solution. Therefore, simulation techniques are employed to determine the final estimates. Another view of the indirect inference estimator as followed by Gallant and Tauchen [1996a] is that the derivatives of the criterion function for the auxiliary model (usually the log-likelihood function) can be used as a moment function for a GMM procedure. Thus, the scores of the Quasi-ML procedure of the possibly misspecified auxiliary model are the moments to be matched by a GMM approach. Hence, in this context the auxiliary model is also termed *score generator*. However, if the indirect inference estimator is combined with some flexible data dependent choice of the auxiliary model, the resulting estimator can be expected to be more efficient than a GMM procedure based on an ad-hoc selection of the moments. For this reason, an indirect inference estimator based on such a flexible auxiliary model is called *Efficient Method of Moments* (EMM).

Consider a dynamic model characterized by $h(y_t|z_t;\theta)$ which allows us to simulate values of y_t using path simulations but with intractable criterion functions commonly used for estimation. Furthermore, let $\mathcal{M}^* = \{h^*(y_t|z_t;\lambda), \lambda \in \Lambda\}$ denote the auxiliary model with the q-dimensional vector of auxiliary parameters λ, where $q \geq p$, that is, the auxiliary model has at least as many parameters as θ. The model is misspecified, if there exists no parameter vector λ^* such that $h_0(y_t|z_t) = h^*(y_t|z_t;\lambda^*)$. However, it is assumed that the auxiliary model has some tractable criterion function – here: the log-likelihood – allowing us to estimate λ. For example, if we are interested in estimating the SV model in Example 1, a possible auxiliary model may be a GARCH model which is relatively easy to estimate by ML compared to the SV model.

The Quasi-ML estimate of λ is computed by maximizing the criterion function $Q(Y,X;\lambda) = T^{-1}\sum_{t=1}^{T}\log h^*(y_t|z_t;\lambda)$ with $Y = [y_1,\ldots,y_T]$ and $X = [x_1,\ldots,x_T]$, that is

$$\widetilde{\lambda}_T = \underset{\lambda}{\text{argmax}}\ Q(Y,X;\lambda)\ .$$

The first order condition is that the score vector:

$$g(Y,X;\lambda) = \frac{\partial Q(Y,X;\lambda)}{\partial\lambda} = \frac{1}{T}\sum_{t=1}^{T}\frac{\partial \log h^*(y_t|z_t;\lambda)}{\partial\lambda} \tag{9}$$

is equal to zero. An important concept linking the structural parameters θ with the auxiliary parameters λ, is the so-called *binding function* $\lambda = b(\theta)$ (see, Gouriéroux and Monfort [1996, p. 67]). The binding function is obtained from the solution of the equation

$E_\theta g[Y,X;b(\theta)] = 0$, where the expectation value is evaluated with respect to the joint distribution $h(Y,X|\theta)$ implied by the structural model.

¿From White [1994] it is known that the estimates $\widetilde{\lambda}_T$ converge in probability to the *pseudo-true* value given by $\lambda_0 = b(\theta_0)$. Hence, if λ and θ are of the same dimension and if it is assumed that there exists an inverse function $b^{-1}(\cdot)$, it is possible to obtain an indirect inference estimator for θ_0 as $\hat{\theta}_T = b^{-1}(\widetilde{\lambda}_T)$. The practical problem is, however, that usually the function $b(\theta)$ is unknown and must be evaluated using Monte Carlo simulations. Therefore, we generate R simulated paths $y_1^{(r)}(\theta),\ldots,y_T^{(r)}(\theta)$, $(r=1,\ldots,R)$ from the distribution $h(y_1,\ldots,y_T|x_1,\ldots,x_T;\theta)$ for observed values of the exogenous variables. For every of these simulated paths we obtain an estimate of the vector of auxiliary parameters denoted by $\widetilde{\lambda}_T^{(r)}(\theta)$. Then the unkown binding function $b(\theta)$ can be approximated by

$$\hat{b}_R(\theta) = \frac{1}{R}\sum_{r=1}^{R}\widetilde{\lambda}_T^{(r)}(\theta).$$

If $b(\theta)$ is replaced by $\hat{b}_R(\theta)$ we can construct a simulated minimum distance estimator as:

$$\widehat{\theta}_{MD}^R = \underset{\theta}{\operatorname{argmin}}\,[\widetilde{\lambda}_T - \hat{b}_R(\theta)]'A[\widetilde{\lambda}_T - \hat{b}_R(\theta)], \tag{10}$$

where A is a positive definite weight matrix. This indirect inference estimator suggested by Gouriéroux, Monfort and Renault [1993] searches for a value of θ, for which simulated data from the structural model approximate the properties of the observed data summarized by the estimate $\widetilde{\lambda}_T$ as close as possible.

As the sample size T tends to infinity, the indirect inference estimator is consistent and asymptotically normal for any fixed $R \geq 1$ (see, Gouriéroux, Monfort and Renault [1993]). Furthermore, the asymptotic optimal weight matrix is given by

$$A_0 = J_0 I_0^{-1} J_0\ ,$$

where

$$\begin{aligned} J_0 &= \lim_{T\to\infty} E\left\{\frac{\partial^2 Q(Y,X;\lambda_0)}{\partial\lambda\partial\lambda'}\right\} \\ I_0 &= \lim_{T\to\infty} \operatorname{var}\left\{\sqrt{T}g(Y,X;\lambda_0) - E[\sqrt{T}g(Y,X;\lambda_0)|X]\right\}. \end{aligned}$$

For this optimal choice of the weight matrix the asymptotic distribution of the minimum distance estimator (10) is obtained as $T^{1/2}(\widehat{\theta}_{MD}^R - \theta_0)\xrightarrow{d}N(0,\operatorname{avar}(\widehat{\theta}_{MD}^R))$, where the asymptotic variance of $\widehat{\theta}_{MD}^R$ is given by $\operatorname{avar}(\widehat{\theta}_{MD}^R) = [1+(1/R)][B'A_0B]^{-1}$ with $B = \partial b(\theta_0)/\partial\theta'$ (see, Gouriéroux, Monfort and Renault [1993]).

The second approach for deriving an indirect estimate from the auxiliary model suggested by Gallant and Tauchen [1996a] is to use the moment conditions implied by the scores of the auxiliary model

$$E\,g[Y,X;b(\theta_0)] = 0\ . \tag{11}$$

Using path simulations from the structural model to approximate $E_\theta g[Y_T, X_T; b(\theta)]$, the GMM estimation procedure based on the scores of the auxiliary model results as

$$\widehat{\theta}_{GT}^R = \underset{\theta}{\operatorname{argmin}}\ \hat{g}_R(\theta, \widetilde{\lambda}_T)' A\, \hat{g}_R(\theta, \widetilde{\lambda}_T) \ , \tag{12}$$

where

$$\hat{g}_R(\theta, \widetilde{\lambda}_T) = \frac{1}{R} \sum_{r=1}^{R} \frac{1}{T} \sum_{t=1}^{T} \frac{\partial \log h^*[y_t^{(r)}(\theta) \mid z_t^{(r)}(\theta); \widetilde{\lambda}_T]}{\partial \lambda} \tag{13}$$

is the simulated score function which approximates the moment conditions (11) and A is a positive definite weight matrix. For this estimator the asymptotic optimal weight matrix is given by $A_0 = I_0^{-1}$. Notice that the score vector (9) for the observed data and the estimate $\widetilde{\lambda}_T$ is equal to zero as implied by the first order condition. Hence, the estimator $\widehat{\theta}_{GT}^R$ searches for a value of θ, for which simulated data from the structural model mimic this first order condition.

Both estimators $\widehat{\theta}_{MD}^R$ and $\widehat{\theta}_{GT}^R$ are derived from similar principles although the criterion function is different. Indeed Gouriérioux, Monfort and Renault [1993] show that both approaches yield asymptotically equivalent estimators as T tends to infinity. Thus, the choice between these estimators is a matter of computational convenience. As far as this is concerned the following should be considered. As is usual for nonlinear optimization problems, estimations based on $\widehat{\theta}_{MD}^R$ and $\widehat{\theta}_{GT}^R$ are performed with iterative optimization algorithms. However, at every iteration step of the optimization with respect to θ, the parameter based estimator $\widehat{\theta}_{MD}^R$ requires "secondary" optimizations to estimate the auxiliary parameters λ, whereas the score based estimator $\widehat{\theta}_{GT}^R$ requires only one optimization concerning λ. Furthermore, the estimator $\widehat{\theta}_{MD}^R$, using the optimal weight matrix A_0, requires an estimate of J_0 based on the Hessian matrix which is not necessary for the estimator $\widehat{\theta}_{GT}^R$. On the other hand, for the computational efficiency of the score based estimator $\widehat{\theta}_{GT}^R$, it is necessary that the score vector of the auxiliary model (9) is available in an analytical form which is not essential for the parameter based estimator.

The asymptotic efficiency of the indirect inference estimators depends on the potential of the auxiliary model to approximate the true process. In fact, if $h(y_t|z_t;\theta_0) = h^*(y_t|z_t; b(\theta_0))$ in some neighborhood of θ_0, the structural model is "smoothly embedded within the score generator" (see, Gallant and Tauchen [1996a]), and it follows that the indirect inference estimator is asymptotically efficient. However, in principle two different approaches to select an appropriate auxiliary model (or score generator) exists. The first approach is to search for an auxiliary model that is able to mimic the salient features of the structural model, and that is as close as possible to it. For the SV model (see Example 1), for instance, such a candidate model may be a GARCH specification since the predictions concerning the stochastic behavior of the returns resulting from a GARCH model and the SV model are very similar. The second approach as advocated by Gallant and Tauchen [1996a] is a data dependent choice of the auxiliary model. Specifically, they propose to adopt a flexible, possibly nonparametric, score generator which can be expected to capture any dynamic and distributional feature of the observed data. Such a data dependent procedure associated with the term EMM is considered in the following section in greater detail.

5 The SNP approach

To achieve a high level of efficiency for the indirect inference estimator, Gallant and Tauchen [1996a] suggest to use the class of semi-nonparametric (SNP) models of Gallant and Nychka [1987] for constructing the score generator. As shown by Gallant and Long [1997], these SNP models can be expected to capture the probabilistic structure of any stationary and Markovian time-series.

The SNP model as applied by Gallant and Tauchen [1996b], Andersen and Lund [1997], and Gallant, Hsieh and Tauchen [1997] to various financial time series can be represented by the following conditional density:

$$h_q^*(y_t|z_t;\lambda_q) = \frac{[\mathcal{P}(u_t,z_t)]^2\,\phi(u_t)/|\det(S_t)|}{\int[\mathcal{P}(v,z_t)]^2\,\phi(v)dv}\ . \tag{14}$$

Here $z_t = [y'_{t-1},\ldots,y'_{t-l}]'$ and λ_q is a q-dimensional parameter vector. The n-dimensional vector u_t is obtained from a standardization of y_t, i.e., $u_t = S_t^{-1}(y_t-\mu_t)$, where μ_t and S_t are a location and a scale function, respectively. The density function of a multivariate normal distribution with mean zero and unit covariance matrix is denoted by $\phi(\cdot)$, and $\mathcal{P}(u_t,z_t)$ is a polynomial in u_t with coefficients depending on z_t. The integration constant $\int[\mathcal{P}(v,z_t)]^2\,\phi(v)dv$ ensures that $h_q^*(y_t|z_t;\lambda_q)$ integrates to unity.

The parametrizations of the location function, the scale function, and the polynomial are as follows. To accommodate the dynamic structure in the mean, the location function is the conditional mean of vector autoregression given by

$$\mu_t = b_0 + \sum_{i=1}^{l_\mu} B_i y_{t-i}\ . \tag{15}$$

To capture the dynamics in the variance, the following ARCH-type scale function is applied:

$$\mathrm{vech}(S_t) = c_0 + \sum_{i=1}^{l_S} C_i|y_{t-i}-\mu_{t-i}|\ , \tag{16}$$

where $\mathrm{vech}(S_t)$ is the vector containing the $[n(n+1)/2]$ distinct elements of S_t and $|y_{t-i}-\mu_{t-i}|$ indicates the elementwise absolute value. Alternative scale functions applied by Andersen and Lund [1997] and Andersen, Chung and Sørensen [1998] are based on corresponding GARCH-type specifications. In order to account for non-Gaussianity and dynamic dependencies of the standardized process u_t the normal density $\phi(\cdot)$ is expanded using the square of the polynomial

$$\mathcal{P}(u_t,z_t) = \sum_{|\alpha|=0}^{k_u} a_\alpha(z_t)u_t^\alpha\ , \tag{17}$$

where $u^\alpha = \prod_{i=1}^{n} u_i^{\alpha_i}$ and $|\alpha| = \sum_{i=1}^{n} |\alpha_i|$. The parameter k_u denotes the degree of the polynomial and controls the extent to which $h_q^*(y_t|z_t;\lambda_q)$ deviates from the normal density. For $k_u = 0$ the density function $h_q^*(y_t|z_t;\lambda_q)$ reduces to that of a normal distribution. To achieve identification, the constant term of the polynomial is set equal to 1. To allow for deviations from normality to depend on past values of y_t, the coefficients $a_\alpha(z_t)$ are polynomials in z_t given by

$$a_\alpha(z_t) = \sum_{|\beta|=0}^{k_z} a_{\alpha\beta} z_t^\beta \ ,$$

where $z^\beta = \prod_{i=1}^{(n\cdot l)} z_i^{\beta_i}$ and $|\beta| = \sum_{i=1}^{(n\cdot l)} |\beta_i|$. For $k_z = 0$ the deviations from the shape of a normal distribution are independent from z_t.

Summing up, the leading term of the SNP model, obtained for $k_u = k_z = 0$, is a Gaussian VAR-ARCH specification depending on the lag lengths l_μ and l_S. This leading term captures the heterogeneity in the first two moments. The remaining features of the data such as any remaining non-normality and possible heterogeneity in the higher-order moments are accommodated by an expansion of the squared Hermite polynomial $\mathcal{P}(u_t, z_t)^2\, \phi(u_t)$ controlled by k_u and k_z. To estimate the parameter vector λ_q, whose dimension is determined by l_μ, l_S, k_u, and k_z, the ML method can be used. For this purpose, the integration constant of the SNP model (14) can be computed analytically by applying the recursive formulas for the moments of a standard normal distribution (see, e.g. Patel and Read [1982]).

If the dimension of the SNP model q increases with the sample size T, the Quasi-ML estimate of the SNP model $h_q^*(y_t|z_t;\lambda_q)$ is under weak conditions an efficient nonparametric estimate of the true density $h_0(y_t|z_t)$ (see, Fenton and Gallant [1996a,b]). Furthermore, Gallant and Long [1997] show that the indirect inference estimator with the SNP model as the score generator (or EMM estimator) attains the asymptotic efficiency of the ML estimator by increasing the dimension q. However, how to determine the adequate specification of the SNP model, i.e. to select l_μ, l_S, k_u and k_z, remains a difficult problem. In most practical applications (see e.g. Gallant, Rossi and Tauchen [1992], Gallant and Tauchen [1996b] and Tauchen [1997]) the dimension q of the SNP model is successively expanded and the model selection criteria of Akaike [1974] or Schwarz [1978] are used to determine a preferred specification. Then, in order to prove the adequacy of the Schwarz- or Akaike-preferred specification, diagnostic tests based on the standardized residuals are conducted.

6 Some Practical Issues

In many cases the application of simulation techniques require an immense amount of computer power and thus some care is necessary when implementing the simulation procedures. In this section we therefore address some practical problems and report implications of recent Monte Carlo studies concerning the properties of simulation based estimators.

6.1 Drawing Random Numbers and Variance Reduction

In most applications the simulation based estimator is obtained by optimizing the criterion function using an iterative algorithm. At every iteration step the criterion function must be estimated via simulations given the current parameter values. For the convergence of such an algorithm, it is important to use *common random numbers* at every iteration step for evaluating the criterion function. With regard to the reduced form $y_t = \varrho(z_t, \varepsilon_t; \theta)$ of the model to be estimated, the use of common random numbers means that for every value of θ during the iterative optimization procedure, the same set of simulated random variables $\{\varepsilon_t^{(r)}\}$ is used to generate simulated values of y_t which enter the criterion function. If at each iteration step new values of ε_t were drawn, some extra randomness would be introduced and the algorithm would fail to converge (see, e.g. Hendry [1984]).

As shown above, the overall variance of simulation based estimators consists of two components. The first component represents the variance of the estimator if it was based on the exact criterion function and the second component is the Monte Carlo sampling variance arising because the criterion function is evaluated by simulations. The first component is irreducible whereas the second component can be made arbitrarily small by increasing the simulation sample size. Unfortunately, this often leads to an enormous increase in computing costs. However, there exists a number of techniques developed for reducing the Monte Carlo sampling variance without increasing the computing costs, for instance, the *antithetic variates* and *control variates* procedures.

The idea of the antithetic variates procedure as applied, for example, by Andersen and Lund [1997] for an indirect inference estimator is as follows. If we want to estimate a quantity ω by simulations, here for example, the moment conditions (11), we construct two estimates for these moment conditions according to the estimator (13), say $\hat{\omega}_1$ and $\hat{\omega}_2$, that are negatively correlated. Then the average $\frac{1}{2}(\hat{\omega}_1 + \hat{\omega}_2)$ has lower variance than either of the two individual estimates. Assuming that the error term ε_t in the reduced form of the model has a symmetric distribution around zero, negatively correlated estimates of moment conditions ω can be produced by using a set of simulated values $\{\varepsilon_t^{(r)}\}$ for $\hat{\omega}_1$ and the same set of simulated values but with the opposite sign, i.e. $\{-\varepsilon_t^{(r)}\}$, for $\hat{\omega}_2$. The additional computing costs of these procedure are negligible and the reduction of the Monte Carlo sampling variance may be considerable as reported by Andersen and Lund [1997].

The control variates technique, as applied by Calzolari, Di Iorio and Fiorentini [1998] for indirect inference, uses two components for the final Monte Carlo estimate of the quantity of interest ω. The first component is the natural Monte Carlo estimate for ω denoted by $\hat{\omega}^\star$, and the second component is an estimate $\tilde{\omega}$ created from the same set of simulated random numbers as $\hat{\omega}^\star$ with known expectation and a positive correlation with $\hat{\omega}^\star$. Then the final estimate of ω based on the control variate $\tilde{\omega}$ is given by $\hat{\omega} = (\hat{\omega}^\star - \tilde{\omega}) + E(\tilde{\omega})$. Under suitable conditions, the variance of $\hat{\omega}$ is considerably smaller than that of the natural estimator $\hat{\omega}^\star$. Specifically, Calzolari, Di Iorio and Fiorentini [1998] adjust the parameter based indirect inference estimator by control variates created from the difference $(\hat{\lambda} - \widetilde{\lambda}_T)$, where $\widetilde{\lambda}_T$ is the estimate of the auxiliary parameter λ based on the observed data and $\hat{\lambda}$ is an estimate of λ using simulated data from the auxiliary model. These simulated data are generated using $\widetilde{\lambda}_T$ as the parameter vector and the same set of simulated random numbers as for the indirect inference procedure ifself. Based on Monte Carlo experiments, they show that the indirect inference estimator combined with control variates and applied to continuous-time models (see Example 2) reduces the Monte Carlo sampling variance substantially compared to the simple indirect inference estimator.

6.2 The Selection of the Auxiliary Model

Indirect inference has been applied to a variety of models for financial time series. In the following we discuss strategies used to select an auxiliary model (or score generator).

A data dependent choice of the auxiliary model based on an expansion of the SNP model (14) has been followed by Gallant and Tauchen [1996b], Tauchen [1997] and Andersen and Lund [1997] to estimate continuous-time models for interest rates, as the Cox-Ingersoll-Ross and Chan-Karolyi-Longstaff-Sanders specification (see Example 2). The same approach is used by Gallant, Hsieh and Tauchen [1997] for the estimation of discrete-time SV models (see Example 1) for interest rates, stock returns and exchange rates. In these applications the dimension q of the SNP auxiliary model determined by model selection criteria as those from Akaike [1974] and Schwarz [1978] is typically quite high, resulting in a multitude of auxiliary parameters and hence in a large number of moments. Specifically, it turns out, that an expansion of the scale function as that in equation (16) is necessary to accomodate for the typically observed conditional heteroscedasticity of financial time series and that the expansion of the polynomial (17) is important to capture, for instance, the typically leptokurtic distribution of financial time series not accomodated by a time varying scale function and possible asymmetries of this distribution.

More simple auxiliary models which are close to the structural model, resulting in a comparable number of auxiliary parameters as structural parameters, are chosen in the applications of Broze, Scaillet and Zakoian [1995] and Engle and Lee [1996]. To estimate the Cox-Ingersoll-Ross and Chan-Karolyi-Longstaff-Sanders specification for interest rates Broze, Scaillet and Zakoian [1995] use auxiliary models based on simple discrete-time Euler approximations of the corresponding continuous-time model. Engle and Lee [1996] apply GARCH specifications as auxiliary models to estimate continuous-time SV models for exchange rates, interest rates and stock returns.

However, the data dependent SNP approach to select an auxiliary model is motivated by asymptotic arguments indicating that this approach ensures a high level of efficiency of the indirect inference estimator when the maintained structural model is true. Clearly, if the structural model is true, a simple auxiliary model very close to it in the sense that it reflects all salient features of the structural model can also be expected to ensure a high level of efficiency. Nevertheless, the data dependent SNP approach seems to be more adequate if we are interested in detecting possible misspecifications of the structural model based on corresponding specification tests, which are not discussed here[3].

6.3 Small Sample Properties of Indirect Inference

The theory of the indirect inference estimator, as developed by Gouriéroux, Monfort and Renault [1993], Gallant and Tauchen [1996a] and Gallant and Long [1997], is based on asymptotic arguments. This raises the question on the finite sample properties of the indirect inference estimator. A comprehensive Monte-Carlo study of the performance of EMM in finite samples is conducted by Andersen, Chung and Sørensen [1998]. Specifically, they use the stochastic volatility model (see Example 1) to compare EMM with GMM and likelihood-based estimators and to address the adequate parametrization of the auxiliary model. Their key findings are that EMM provides, independent of the sample size, a substantial efficiency gain relative to the standard GMM procedure. Furthermore, the likelihood-based estimators are generally more efficient than the EMM procedure, but EMM approaches the efficiency of the likelihood-based estimators with increasing sample size, as it is consistent with the asymptotic theory of the EMM estimator. Finally, they find evidence that score generators based on an over-parametrized SNP model lead, especially in smaller samples, to a substantial loss of efficiency. Specifically, they show that the substitution of an ARCH-type scale function in the SNP model as given in equation (16) by a GARCH-type specification, improve the efficiency of the EMM estimator. In fact, this substitution reduces the number of parameters which are necessary to capture the autocorrelation in the variance as implied by the SV model.

7 Conclusion

In recent years, simulation-based inference procedures have become popular in particular in empirical finance. This is due to the complexity of standard models implied by latent factors or continuous-time processes, for example. This chapter reviews different approaches based on a GMM criterion function for the estimation of the parameters. The MSM approach is the simulated counterpart of the traditional GMM procedure and is applicable if the theoretical moments cannot be computed analytically. However, in

[3]For specification tests based on indirect inference, see e.g. Gouriéroux, Monfort and Renault [1993], Tauchen [1997] and Gallant, Hsieh and Tauchen [1997].

many applications it is not clear how to choose the moment conditions. In nonlinear models the structure implies restrictions on a wide range of moments and, therefore, it is difficult to represent the main features of the model using a few moment conditions. In such cases it seems attractive to employ a simple auxiliary model which approximates the main features of the structural model. However, in most cases, the relationship between the parameters of the auxiliary model and the parameters of interest is too complicated to admit an explicit solution. Hence, simulation techniques are applied to evaluate the binding function linking the parameters of interest with the parameters of the auxiliary model. Two asymptotically equivalent approaches for such an indirect infereence framework are available. Gouriéroux, Monfort and Renault [1993] employ a minimum distance procedure whereas Gallant and Tauchen [1996a] use the scores of the auxiliary model as the moment condition to be matched by a (simulation-based) GMM procedure.

Since the efficiency of an indirect inference procedure crucially depends on the potential of the auxiliary model to approximate the model of interest, it seems attractive to use flexible nonparametric models as score generators. Such estimation procedures are known as EMM estimators in the literature and seem to be a fruitful and a promising field of future research.

References

Akaike, H. [1974], "A New Look at the Statistical Model Identification", *IEEE Transactions on Automatic Control 19*, 716-723.

Andersen T.G., Chung, H.J. and B.E. Sørensen [1998], "Efficient Method of Moments Estimation of a Stochastic Volatility Model: A Monte Carlo Study", Working Paper No. 97-12, Brown University.

Andersen T.G. and J. Lund [1997], "Estimating Continuous-Time Stochastic Volatility Models of the Short-Term Interest Rate, *Journal of Econometrics 77*, 343–377.

Andersen, T.G., and B.E. Sørensen [1996], "GMM Estimation of a Stochastic Volatility Model: A Monte Carlo Study", *Journal of Business & Economic Statistics 14*, 328-352.

Bansal, R., Gallant, A.R., Hussey, R. and G. Tauchen [1993], "Computational Aspects of Nonparametric Simulation Estimation", in: D.A. Belsley, ed., *Computational Techniques for Econometrics and Economic Analysis*, Kluwer Academic Publishers, Boston, Massachusetts.

Bansal, R., Gallant, A.R., Hussey, R. and G. Tauchen [1995], "Nonparametric Estimation of Structural Models for High-Frequency Currency Market Data", *Journal of Econometrics 66*, 251-287.

Bollerslev, T. [1986], "Generalized Autoregressive Conditional Heteroskedasticity", *Journal of Econometrics 31*, 307-327.

Broze, L., Scaillet, O. and J. Zakoian [1995], "Testing for Continuous-time Models of the Short-term Interest Rate", *Journal of Empirical Finance 2*, 199-223.

Calzolari, G., Di Iorio, F. and G. Fiorentini [1998], "Control Variates for Variance Reduction in Indirect Inference: Interest Rate Models in Continuous Time", *The Econometrics Journal*, forthcoming.

Chesney, M. and L. Scott [1989], "Pricing European Currency Options: A Comparison of the Modified Black-Scholes Model and a Random Variance Model", *Journal of Financial and Quantitative Analysis 24*, 267-284.

Chan, K.C., Karolyi, G.A., Longstaff, F.A. and A.B. Sanders [1992], "An Empirical Comparison of Alternative Models of the Short-Term Interest Rate", *Journal of Finance 47*, 1209-1227.

Cox, J.C., Ingersoll, J.E. and S.A. Ross [1985], "A Theory of the Term Structure of Interest Rates", *Econometrica 53*, 385-407.

Danielsson, J. and J.F. Richard [1993], "Accelerated Gaussian Importance Sampler with Application to Dynamic Latent Variable Models", *Journal of Applied Econometrics 8*, S153-S173.

Duffie, D. and K.J. Singleton [1993], "Simulated Moments Estimation of Markov Models of Asset Prices", *Econometrica 61*, 929-952.

Diebold, F.X. and M. Nerlove [1989], "The Dynamics of Exchange Rate Volatility: A Multivariate Latent Factor ARCH Model", *Journal of Applied Econometrics 4*, 1-21.

Engle, R.F. [1982], "Autoregressive Conditional Heteroscedasticity with Estimates of the Variance of United Kingdom Inflation", *Econometrica 50*, 987-1007.

Engle, R.F., and G.G.J. Lee [1996], "Estimating Diffusion Models of Stochastic Volatility", in: Rossi, P.E., ed., *Modelling Stock Market Volatility: Bridging the Gap to Continuous Time*, Academic Press, San Diego.

Engle, R.F., Ng, V.K. and M. Rothschild [1990], "Asset Pricing with a Factor-ARCH Covariance Structure: Empirical Estimates for Treasury Bills", *Journal of Econometrics 45*, 213-237.

Fenton, V.M., and A.R. Gallant [1996a], "Convergence Rates of SNP Density Estimators", *Econometrica 64*, 719-727.

Fenton, V.M., and A.R. Gallant [1996b], "Qualitative and Asymptotic Performance of SNP Density Estimators", *Journal of Econometrics 74*, 77-118.

Forster, F.D., and S. Viswanathan [1995], "Can Speculative Trading Explain Volume-Volatility Relation?", *Journal of Business & Economic Statistics 13*, 379-396.

Gallant, A.R., Hsieh, D.A. and G.E. Tauchen [1997], "Estimation of Stochastic Volatility Models with Diagnostics", *Journal of Econometrics 81*, 159-192.

Gallant, A.R., and J.R. Long [1997], "Estimating Stochastic Differential Equations Efficiently by Minimum Chi-Squared", *Biometrica 84*, 125-141.

Gallant, A.R., and D.W. Nychka [1987], "Semi-Nonparametric Maximum Likelihood Estimation", *Econometrica 55*, 363-390.

Gallant, A.R., Rossi, P.E., and G.E. Tauchen [1992], "Stock Prices and Volume", *The Review of Financial Studies 5*, 199-242.

Gallant, A.R., and G.E. Tauchen [1996a], "Which Moments to Match?", *Econometric Theory 12*, 657-681.

Gallant, A.R., and G.E. Tauchen [1996b], "Specification Analysis of Continuous Time Models in Finance", in: Rossi, P.E., ed., *Modelling Stock Market Volatility: Bridging the Gap to Continuous Time*, Academic Press, San Diego.

Gennotte, G., and T.A. Marsh [1993], "Variations in Economic Uncertainty and Risk Premiums on Capital Assets", *European Economic Review 37*, 1021-1041.

Ghysels, E., Harvey, A.C., and E. Renault [1996], "Stochastic Volatility", in: Maddala, G.S. and C.R. Rao, ed., *Handbook of Statistics, Vol.14*, Elsevier Science B.V., Amsterdam.

Gouriéroux, C., and A. Monfort [1993], "Simulation Based Inference: A survey with special reference to panel data models", *Journal of Econometrics 59*, 5-33.

Gouriéroux, C., and A. Monfort [1996], *Simulation Based Econometric Methods*, CORE Lectures, Oxford University Press, New York.

Gouriéroux, C., Monfort. A. and E. Renault [1993], "Indirect Inference", *Journal of Applied Econometrics 8*, S85-S118.

Hendry, D.F. [1984], "Monte Carlo Experimentation in Econometrics", in: Griliches, Z. and M.D. Intriligator, ed., *Handbook of Econometrics, Vol 2*, Elsevier Science B.V., Amsterdam.

Hull, J., and A. White [1987], "The Pricing of Options on Assets with Stochastic Volatilities", *Journal of Finance 42*, 281-300.

Jacquier, E., Polson, N.G. and P.E. Rossi [1994], "Bayesian Analysis of Stochastic Volatility Models", *Journal of Business & Economic Statistics 12*, 371-389.

Kim, S., Shephard, N. and S. Chib. [1996], "Stochastic Volatility: Likelihood Inference and Comparison with ARCH Models", working paper, Nuffield College, Oxford University.

Kloek, T., and H.K. van Dijk [1978], "Bayesian Estimates of Equation System Parameters: An Application of Integration by Monte Carlo", *Econometrica 46*, 1-19.

Lee, B. and B.F. Ingram [1991], "Simulation Estimation of Time-Series Models", *Journal of Econometrics 47*, 197-205.

McFadden, D. [1989], "A Method of Simulated Moments for Estimation of Discrete Response Models Without Numerical Integration", *Econometrica 57*, 995-1026.

Newey, W.K. [1993]," Efficient Estimation of Models with Conditional Moment Restrictions", in: Maddala, G.S., Rao, C.R. and H.D. Vinod, ed., *Handbook of Statistics, Vol. 11*, North-Holland, Amsterdam.

Pakes, A. and D. Pollard [1989], "Simulation and the Asymptotics of Optimization Estimators", *Econometrica 57*, 1027-1057.

Patel, J.K. and C.B. Read [1982], *Handbook of the Normal Distribution,* Marcel Dekker, New York.

Richard, J.F. and W. Zhang [1997], "Accelerated Importance Sampling", working paper, University of Pittsburgh.

Schwarz, G. [1978], "Estimating the Dimension of a Model", *The Annals of Statistics 6*, 461-464.

Shephard, N. [1996], "Statistical Aspects of ARCH and Stochastic Volatility", in: Cox, D.R., Hinkley, D.V. and O.E. Barndorff-Nielsen, ed., *Time Series Models In Econometrics, Finance and Other Fields*, Chapman & Hall, London.

Stern, S. [1997], "Simulation-Based Estimation", *Journal of Economic Literature 35*, 2006-2039.

Tauchen, G.E. [1997], "New Minimum Chi-Square Methods in Empirical Finance", in: Kreps, D. and K. Wallis, ed., *Advances in Econometrics: Theory and Applications, Seventh World Congress, Vol.3*, Cambridge University Press, Cambridge.

Taylor, S.J. [1986], *Modelling Financial Time Series,* John Wiley & Sons, Chichester.

Taylor, S.J. [1994], "Modeling Stochastic Volatility: A Review and Comparative Study", *Mathematical Finance 4*, 183-204.

White, H. [1994], *Estimation, Inference and Specification Analysis,* Cambridge University Press, Cambridge.