Faculdade de Economia,
Administração e Contabilidade
de Ribeirão Preto
Universidade de São Paulo

Texto para Discussão

Série Economia

TD-E 17 / 2011

Métodos de Verossimilhança Empírica em Finanças

Prof. Dr. Márcio Poletti Laurini

Av. Bandeirantes, 3900 - Monte Alegre - CEP: 14040-900 - Ribeirão Preto - SP
Fone (16) 3602-4331/Fax (16) 3602-3884 - e-mail: cebelima@usp.br site: www.fearp.usp.br

Universidade de São Paulo

Faculdade de Economia, Administração e Contabilidade de Ribeirão Preto

A série TEXTO PARA DISCUSSÃO tem como objetivo divulgar: i) resultados de trabalhos em desenvolvimento na FEA-RP/USP; ii) trabalhos de pesquisadores de outras instituições considerados de relevância dadas as linhas de pesquisa da instituição. Veja o site da Comissão de Pesquisa em **www.cpq.fearp.usp.br**. Informações: e-mail: cpq@fearp.usp.br

GENERALIZED EMPIRICAL LIKELIHOOD/MINIMUM CONTRAST ESTIMATION OF STOCHASTIC DIFFERENTIAL EQUATIONS

MÁRCIO POLETTI LAURINI

LUIZ KOODI HOTTA

Abstract. In this study we approach the semiparametric estimation of Stochastic Differential Equations employing methods of generalized empirical likelihood and generalized minimum contrast. The results obtained demonstrate that the estimators proposed, particularly, the *Exponentially Tilted Empirical Likelihood* (Schennach (2007)) estimator, obtain better results than those of the Generalized Methods of Moment generally used in the estimation of stochastic differential equations. These results are derived from the robustness properties of this method in the presence of problems of incorrect specification, which, in the context of the estimation of stochastic differential equations, occurs by using the process' approximate discretization in the construction of moment conditions. The analyses are carried out by means of Monte Carlo experiments for unconditional and conditional formulations of moment conditions.

Key Words: Stochastic Differential Equations, Empirical Likelihood, Generalized Minimum Contrast.

1. Introduction

The use of stochastic processes in continuous time in the modeling and pricing of financial instruments is one of the basis of the modern theory of Finance, and its origin can be traced back to Bachelier (1900)'s seminal study. The use of stochastic processes in continuous time is justified by the mathematical convenience in relation to the use of processes in discrete time and the possibility of employing the mathematical theory developed for the general class of processes known as continuous semi-martingales, making it possible to perform an application of the whole theory of pricing by no-arbitrage (Harrison and Kreps (1979), Harrison and Pliska (1981) and Delbaen and Schachermayer (1994)) in this context. The basic objects of the modeling of stochastic processes in continuous time are the so-called *Stochastic Differential Equations*, which are objects represented in the general form:

$$dX_t = \mu(t, X_t) + \sigma(t, X_t)dW_t, \tag{1.1}$$

where $\mu(t, X_t)$ represents the deterministic part of the process (instantaneous *drift*), $\sigma(t, X_t)$ represents the stochastic component (volatility) of the process, and W_t is the so-called Wiener process or Brownian Motion. This representation is useful because it makes it possible to define the evolution of the process trajectories X_t by means of a representation given by a stochastic integration (e.g Rogers and Williams (2000), Karatzas and Shreve (1987), Kloeden and Platen (1992)):

$$X_t = X_0 + \int_{t_0}^{t} \mu(t, X_t)dt + \int_{t_0}^{t} \sigma(t, X_t)dW_t. \tag{1.2}$$

Different specifications of the drift $\mu(t, X_t)$ and volatility $\sigma(t, X_t)$ processes in the stochastic differential equation give rise to processes with distinct properties. These properties enable the representation of a wide class of processes used in finance. Focusing on the modeling of short-term interest rates, a series of alternative specifications for the modeling of short-term interest rates have been employed. Table 1 presents some formulations used in the literature, comprising the models of Merton (1973), Vasicek (1977), Cox et al. (1985), Dothan (1978), Black and Scholes (1973), Brennan and Schwartz (1980), Cox et al. (1980) and Cox (1975). Notably, on its last line is defined the model called Generalized Cox-Ingersoll-Ross (CIR), containing all the previous models as particular cases, as demonstrated in Chan et al. (1992), which includes a general discussion on the properties of these models.

Parameter estimation in stochastic differential equations is a well developed theme in the econometric literature[1], and there is a very wide range of techniques available. This range of techniques is related to the difficulties inherent to the estimation of stochastic differential

[1]For a review of the literature on stochastic differential equations see Gourieroux and Monfort (1996), Prakasa Rao (1999), Bishwal (2007) and Zivot and Wang (2006)

Merton (1973)	$dX_t = \alpha dt + \sigma dW$
Vasicek (1977)	$dX_t = (\alpha + \beta X_t)dt + \sigma dW$
CIR SR (1985)	$dX_t = (\alpha + \beta X_t)dt + \sigma X_t^{1/2} dW$
Dothan (1978)	$dX_t = \sigma X_t dW$
GBM (1973)	$dX_t = \beta X_t dt + \sigma X_t dW$
Brennan-Schwartz (1980)	$dX_t = (\alpha + \beta X_t)dt + \sigma X_t dW$
CIR VR (1980)	$dX_t = \sigma X_t^{3/2} dW$
CEV (1975)	$dX_t = \beta dt + \sigma dW$
Generalized CIR	$dX_t = (\alpha + \beta X_t)dt + \sigma X_t^{\gamma} dW$

TABLE 1. Models for short-term interest rates.

equations, particularly in the non-existence of analytical solutions for the stochastic integration in general cases and the problem of using discretely observed data in the estimation of a process formulated in continuous time. As examples of estimation methods in this context, we have maximum likelihood, generalized methods of moments (GMM), methods of simulated moments, Martingale estimating equations, Markov chain Monte Carlo and indirect inference, and non-parametric methods. In principle, the most recommended form of estimation consists in employing the likelihood function, because, under the regularity conditions, the maximum likelihood estimators are consistent, efficient and asymptotically normal. However, in the context of the estimation of stochastic differential equations, the non-existence of general solutions is a general difficulty found in the use of methods based on the likelihood of the process, which is formulated by employing the transition density resulting from the solution of the stochastic differential equation.

In the absence of analytical solutions, it is necessary to use approximations in the construction of the likelihood function, such as the use of quasi-maximum likelihood methods, which generates estimators with minimum mean square error, or the use of simulated maximum likelihood, which uses simulated trajectories by Euler or Milstein discretizations in the likelihood evaluation (Pedersen (1995)), or else approximations using Hermite expansions obtained by Ait-Sahalia (2002). Note that, given the employment of approximations in the evaluation of the likelihood function, the optimality properties of this estimator may not remain, and thus other estimators could become competitive.

Estimators using moment conditions are also often employed in the estimation of stochastic differential equations. The estimation using the (GMM) by Hansen (1982), employing a simple discretization of the process, may be the form most widely employed (e.g. Chan et al. (1992)). Although the generalized method of moments is characterized by properties of consistency and asymptotic efficiency, its properties in finite samples and in the presence of specification problems may not be optimal. In order to tackle these problems we discuss the use of two classes of estimators in the estimation of stochastic differential equations employing discrete data - estimators of generalized empirical likelihood and estimators of generalized minimum contrast, comparing their performance with that of the estimators based on estimation by the Generalized Method of Moments. These estimators are semi-parametric in the sense that the parametric form of the stochastic differential equation is used through moment conditions, but the non-observed density of the process is evaluated in a non-parametric form. We analyze the properties of these estimators using unconditional and conditional moments conditions.

The estimators presented (*generalized empirical likelihood, continuous updating empirical likelihood, exponential tilting and exponentially tilted empirical likelihood)* possess the same properties of consistency and first-order asymptotic efficiency (e.g. Smith (2001), Schennach (2007)) as the compared GMM estimators (two-stage GMM, iterative GMM, continuous updating GMM). However, theoretical results demonstrate that these estimators may have superior properties in terms of bias in finite samples, and asymptotic properties of higher order (e.g. Kitamura (2006)). Furthermore, these estimators are asymptotically efficient in the class of semi-parametric estimators (in Bickel et al. (1993)'s sense), and have optimal properties in terms of hypotheses tests: minimax optimality, optimality in the sense of large deviations, and these tests are uniformly more powerful in the generalized Neyman-Person sense. The class of estimators of generalized minimum contrast (exponential tilting and exponentially tilted empirical likelihood) has characteristics of robustness in the presence of specification problems. These characteristics of robustness of the estimators based on generalized minimum contrast are of the utmost importance in the estimation of stochastic differential equations, and because of the non-existence of exact discretizations, all the estimators of continuous processes employing discretely observed data are characterized by a problem of incorrect specification.

This study discusses the use of these methods in the estimation of stochastic differential equations, and the results obtained demonstrate that these estimators obtain superior results when compared with the techniques generally employed of generalized methods of moments. One result of particular interest is that the estimator of exponentially tilted empirical likelihood (Schennach (2007)) obtains results that are much superior in terms of finite sample bias, a result derived from the properties of this model's robustness in the presence of incorrect specification (e.g. Smith (2001), Schennach (2007), Anatolyev and Gospodinov (2011)).

Another important contribution of this paper is to show that the use of methods based on generalized empirical likelihood/ minimum contrast gives a good performance in finite samples through the use of unconditional moments derived from discretizations, procedure that can usually be derived for any continuous time stochastic process. The performance of estimators based on unconditional moments derived from Euler discretizations is comparable to results obtained using conditional moments, which can only be obtained explicitly for a very restricted class of models. These results indicate the excellent properties of the estimators proposed in this article.

This article is structured as follows: section 2 presents a brief review of the estimation of stochastic differential equations employing the GMM. Section 3 presents generalized empirical likelihood and generalized minimum contrast based estimators, discussing their properties, similarities and potential advantages in the estimation of stochastic differential equations. A series of Monte Carlo's experiments is performed in section 4 for the estimation based on the unconditional moments and in section 5 for the estimation based on conditional moments, aiming at stressing some properties of the estimators discussed in this study. The final conclusions are in section 6, showing concisely that the estimators proposed, which are unprecedented in the context of estimation of stochastic differential equations, obtain results that are superior to the techniques of the Generalized Methods of Moments generally employed in the estimation of stochastic differential equations.

2. Estimation by the Generalized Method of Moments

As the technique of the GMM is widely employed in the econometric literature for the estimation of stochastic differential equations, and as it also has deep connections with the estimation methods of Maximum empirical likelihood and generalized minimum contrast, we will start by reviewing this methodology, giving special attention to the moment conditions employed in the estimation.The estimation by the Generalized Method of Moments was introduced by Hansen (1982). The method is based on population moments conditions:

$$E[g(\theta_0, X_t)] = 0, \tag{2.1}$$

where θ_0 is a vector of true values of the parameters. The analogous sample moments conditions are defined as:

$$\overline{g}\,(\theta) = \frac{1}{T}\sum_{t=1}^{T} g(\theta, x_t). \tag{2.2}$$

GMM estimators are defined as solutions to the system:

$$\widehat{\theta} = \arg_{\theta} \frac{1}{T}\sum_{t=1}^{T} g(\theta, x_t) = 0. \tag{2.3}$$

Note that, except in the case of the number of parameters being equal to the number of moment conditions (exactly identified system), the problem described in 2.3 is not identified when the number of conditions is less than the number of parameters, or in general there is no solution when the number of conditions are greater than the number of parameters. In order to obtain a single solution it is used a number of moment conditions larger or equal the number of parameters and define the following criterion function:

$$J(\theta) = \overline{g}\,(\theta)' \, W \overline{g}\,(\theta) \tag{2.4}$$

and the minimization of this function defines the optimum solution of the problem, where W is a positive definite weighting matrix. Hansen (1982) demonstrates that the efficient asymptotic solution of the GMM estimation is obtained when this matrix is given by:

$$W^* = \left\{ \lim_{t \to \infty} Var\left(\sqrt{T}\overline{g}\left(\theta\right)\right)\right\}^{-1} = \Omega(\theta)^{-1} \tag{2.5}$$

and thus the optimal weight is obtained by employing the inverse of the sample variance-covariance matrix. This matrix is usually estimated employing the class of HAC estimators of Newey and West (1987) given by:

$$\widehat{\Omega} = \sum_{s=-(T-1)}^{T-1} k_h(s)\widehat{\Gamma}_s(\theta^*), \tag{2.6}$$

where k is a kernel function dependent on the choice of a bandwidth h, which can be chosen using the Newey and West (1987)'s or Andrews (1991)'s procedures:

$$\widehat{\Gamma}_s(\theta^*) = \frac{1}{T}\sum_{t=1}^{T} g(\theta^*, x_t)g(\theta^*, x_{t+s})', \tag{2.7}$$

The efficient estimator of the GMM is then obtained as a solution of the problem:

$$\widehat{\theta} = \arg\min_{\theta} \overline{g}\left(\theta\right)' \widehat{\Omega}\left(\theta^*\right) \overline{g}\left(\theta\right). \tag{2.8}$$

There are several forms of performing the implementation of the GMM estimator. The initial form proposed by Hansen (1982) is the estimator known as two-stage GMM. This estimator is obtained by performing a first stage by obtaining an initial estimator $\widehat{\theta^*} = \arg\min \overline{g}\left(\theta\right)' \Omega \overline{g}\left(\theta\right)$, where Ω is an initial weight matrix, normally an identity matrix. From this first stage, a HAC matrix $\widehat{\Omega}\left(\theta^*\right)$ is calculated in function of this initial estimation, and the final estimation of the GMM estimator is obtained as $\widehat{\theta} = \arg\min \overline{g}\left(\theta\right)' \widehat{\Omega}\left(\theta^*\right) \overline{g}\left(\theta\right)$ with the HAC matrix obtained in the first stage.

Note that, in this case, there is a dependence on the results of the second stage with the initial estimation on the first stage, and thus this procedure can create a first-order bias impairing the performance of the estimator in finite samples (Hansen et al. (1996)). In order to solve this problem, two alternative procedures are proposed. The first procedure is known as Iterative GMM, which is a modification of the two-stage procedure. In this procedure, the estimation of the first stage is reinitialized with the result of the second stage estimation, and this iteration continues up to when a variation in the vector of parameters becomes smaller than a chosen epsilon.

Another possible estimator is known as GMM with continuous updating (Hansen et al. (1996)). In this case the estimation of the parameter $\widehat{\theta}$ is not performed in stages, but it is performed simultaneously by employing an algorithm of numerical optimization. Starting from an initial vector θ_0 (generally chosen employing the two-stage GMM method) the estimation is performed by $\widehat{\theta} = \arg\min \overline{g}(\theta)' \,\widehat{\Omega}(\theta^*)\, \overline{g}(\theta)$, but now θ and $\widehat{\Omega}(\theta^*)$ are simultaneously determined. This procedure obtains the same first-order properties as the Iterative GMM estimator, but according to Hansen et al. (1996), it has better properties in terms of bias in finite samples. According to Newey and Smith (2004) and Anatolyev (2005), the three methods are asymptotically equivalent, but the second-order bias of the continuous updating estimator is smaller, and the iterations increase the estimator's efficiency. However, the numerical procedure can be subject to multiple modes in the objective function, which renders this estimator numerically unstable.

In order to perform the estimation of stochastic differential equations by employing the GMM, it is necessary to formulate the moment conditions in terms of some discretized form of the model. The first approach employed is by means of the simple discretization adopted in Chan et al. (1992) for the Generalized CIR model (Table 1) given by:

$$X_{t+1} - X_t = \alpha_0 + \beta_0 X_t + \varepsilon_{t+1} \tag{2.9}$$

with the conditions: $E(\varepsilon_{t+1}) = 0$ and $E(\varepsilon_{t+1}^2) = \sigma_0^2 X_t^{2\gamma}$. In this case, we can formulate the moment conditions necessary for the estimation of parameters $(\alpha, \beta, \gamma, \sigma^2))$, by defining $\varepsilon_{t+1} = X_{t+1} - X_t - \alpha_0 - \beta_0 X_t$, and defining four moment conditions in this form:

$$g(\theta) = \begin{bmatrix} \varepsilon_{t+1} \\ \varepsilon_{t+1} X_t \\ \varepsilon_{t+1}^2 - \sigma_0^2 X_t^{2^\gamma} \\ (\varepsilon_{t+1}^2 - \sigma_0^2 X_t^{2\gamma}) X_t \end{bmatrix}. \tag{2.10}$$

and applying the GMM estimation defined by equation 2.8. Moment conditions for the other submodels of the Generalized CIR family can be obtained by imposing the necessary restrictions, according to Table I in Chan et al. (1992). Note that this simple discretization is not consistent - the discretization does not converge to the true solution of the process, since it ignores the time interval between observations. A simple form of obtaining a consistent discretization for this process is to employ a first-order Euler discretization, which defines moment conditions given by a residual vector in this form: $\varepsilon_{t+\triangle t} = r_{t+\triangle} - r_t - (\alpha_0 + \beta_0 r_t)\triangle_t$, and thus constructing the vector of moment conditions as:

$$g(\theta) = \begin{bmatrix} \varepsilon_{t+\triangle t} \\ \varepsilon_{t+\triangle t} X_t \\ \varepsilon_{t+\triangle t}^2 - \sigma_0^2 X_t^{2^\gamma} \triangle t \\ (\varepsilon_{t+\triangle t}^2 - \sigma_0^2 X_t^{2\gamma} \triangle t) X_t \end{bmatrix}. \tag{2.11}$$

This is the form employed in this study. Note that the use of discretization always represents a specification problem in the inference procedure, since, even employing consistent discretizations, the bias term caused by the discretization employed only tends to zero when $\triangle t \to 0$. Note that the time interval $\triangle t$ employed in the process of discretization depends on the frequency of data observation, and thus it is not under the researcher's control. Therefore, there are two sources of bias problems in the estimation of stochastic differential equations: the first form derived from the use of Generalized Methods of Moments estimators, and an additional form generated by the incorrect specification given by the use of a non-exact discretization of the process. Note that in Chan et al. (1992)'s original study, the estimation employs a simple discretization of the model rather than the Euler discretization, and this represents a bias increase in the estimation due to a specification with a larger approximation error. Consequences of this problem can be seen in Prakasa Rao (1999), and a supplementary discussion of this problem is presented in section 4, which demonstrates that this discretization problem leads to a problem of incorrect specification in the estimation of stochastic differential equations.

3. Generalized Empirical Likelihood and generalized minimum contrast Estimators

In the GMM there is a trade-off between weaker necessity of assumptions for its use and the efficiency of the method in finite samples. Conditions of regularity for estimators of the GMM (Hansen (1982), Newey and McFadden (1994)) involve only conditions for the asymptotic validity of the moment conditions and do not assume stronger conditions such as the knowledge of the process distribution, but, in finite samples, the properties of this estimator is not optimal.

The opposite situation would be the estimation by the maximum likelihood method, which employs not only the conditional moments of the process but all the information in the conditional densities. If the process is correctly specified and meets the regularity conditions, then it is a better asymptotically Gaussian estimator, and it also reaches optimality in measures such as Badahur efficiency (Kitamura (2006), DasGupta (2008)). Nevertheless, employing the maximum likelihood in the estimation of stochastic differential equations is difficult by the non-existence of analytical forms for the solution of stochastic differential equations, and thus it is not possible to employ parametric forms for the maximum likelihood estimation.

An alternative form, not yet explored in the literature of inference in continuous time processes, is the use of a form of *non-parametric maximum likelihood estimation* known as *empirical likelihood (EL)*[2]. According to Kitamura (2006), assuming a sequence of IID data $\{x_i\}_{i=1}^{T}$ of

[2]A detailed review of the generalized empirical likelihood estimators can be found in Anatolyev and Gospodinov (2011).

an unknown density, and defining $\triangle$ as the simplex $\Big\{(p_1, \dots, p_T) : \sum_{t=1}^{T} p_t = 1, 0 \le p_t \le 1, t = 1, \dots T\Big\}$, the non-parametric log-likelihood function is defined as:

$$\ell_{NP}(p_1, \dots, p_T) = \sum_{t=1}^{T} \log p_t, \quad (p_1, \dots, p_t) \in \triangle \tag{3.1}$$

which can be interpreted as a log-likelihood of a multinomial distribution with support given by the sample observations $\{x_t\}_{i=1}^{n}$, even if the density x_t is not a multinomial.

A notable advancement in the literature of empirical likelihood was achieved by Owen (1991), who established connections between the non-parametric likelihood and the estimation employing moment conditions, which is also used in the estimation by the GMM, as shown by Qin and Lawless (1994). Assuming the condition of moments in the form:

$$E\left[g(\theta_0, X_t)\right] = \int g(\theta_0, X) d\mu = 0, \theta_0 \in \Theta \subset \mathbb{R}^k, \tag{3.2}$$

it is possible to transform this estimation problem using conditions of moments in a non-parametric likelihood problem employing implicit probabilities p_i, and thus the log-likelihood function to be maximized becomes:

$$\ell_{NP}(p_1, \dots, p_T) = \sum_{t=1}^{T} \log p_t, \;\; s.t. \sum_{t=1}^{T} g(\theta, x_t) p_t = 0$$

The estimator that maximizes this expression is the *maximum empirical likelihood* estimate. The implicit probabilities are related to the validity of the moment conditions. These implicit probabilities give more weight to the observations where the moment conditions are closer to zero. Note the similarity with the estimation by the GMM, which is a simplified form that assumes that all weights are equal, i.e., $p_t = 1/n$.

The use of empirical likelihood is particularly important in the estimation of stochastic differential equations because, except in a few particular cases, there are no exact solutions for the stochastic differential equations, and thus it is not possible to construct analytically the transition densities of the process, which makes it impossible to construct an exact likelihood function. Empirical likelihood method allow us to assess the likelihood of the process in a non-parametric form, and thus they do not depend on the existence of analytical solutions for the stochastic differential equations. This non-parametric evaluation of the likelihood function is efficient in the semi-parametric sense (e.g. Bickel et al. (1993)), and, at the same time, it employs the parametric specification given by the stochastic differential equation to construct moment conditions.

A difference found with the GMM is that, in the methodology of generalized empirical likelihood, the moment condition can be a process weakly dependent and heteroskedastic. In order to tackle this situation, Kitamura (1997) and Kitamura and Stutzer (1997) proposes replacing $g(\theta, x_t)$ for a smoothed version defined as:

$$g^w(\theta, x_t) = \sum_{s=-m}^{m} w(s) g(\theta, x_{t-s}), \tag{3.3}$$

where $w(s)$ are weights obtained by a kernel function adding one, in the spirit of a HAC estimator (Andrews (1991)) . This modification makes it possible to obtain the same conditions of first-order asymptotic efficiency existing in the GMM methods. In this way the estimate given by the moment conditions is given by:

$$\widehat{\theta} = \arg_\theta \sum_{t=1}^{T} p_t g^w(\theta, x_t) = 0. \tag{3.4}$$

The use of smoothing is especially important in the estimation of discretized models, since the Euler discretization for a discretely observed process involves independent processes only when the interval δ in the discretization converges to zero. For discretely observed data, in general the discretization interval is defined by the frequency of the observed sample. Thus, in general, the process of the observed data with a fixed discretization is dependent, justifying the use of smoothing. Another important property is that the use of smoothing can improve the properties in finite samples even for IID data, as discussed in Anatolyev (2005) and Anatolyev and Gospodinov (2011).

An interpretation of equation 3.4 in relation to the GMM estimator is that, while in over-identified models estimated by GMM the moment conditions are not exactly equal to zero, in the estimators defined by this equation the moment conditions are exactly equal to zero by weighting with the use of the empirical probabilities p_t. Note that, in models exactly identified, all the estimators proposed obtain similar results, because in all these estimators the moment conditions are always valid. In over-identified models with valid moment conditions, all these estimators produce the same asymptotic variance.

An alternative interpretation of the empirical likelihood estimator can be obtained, such as that of a particular case of the generalized minimum contrast (GMC) estimator (e.g. Bickel et al. (1993)), similar to the interpretation of the GMM estimator as an estimator of minimum χ^2, or the interpretation of estimators of quasi-maximum likelihood as estimators of minimum contrast. Defining a general divergence function between two measures of probability P and Q as follows:

$$D(P, Q) = \int \phi\left(\frac{dP}{dQ}\right) dQ, \tag{3.5}$$

where ϕ is a convex function. Define M as the set of all the probability measures in $\mathbb{R}^p$ and

$$\mathcal{P}(\theta) = \left\{P \in M : \int g(\theta, x) dP = 0\right\} \tag{3.6}$$

and $\mathcal{P}$ the statistic model of all the probability measures compatible with 3.6. The problem of minimum contrast optimization is given by

$$\inf_{\theta \in \Theta} \inf_{\mathcal{P} \in \mathcal{P}(\theta)} D(P, \mu) \tag{3.7}$$

where μ denote the dominating measure in this model.

Thus in a correctly specified model, this discrepancy must be minimal in $\theta = \theta_0$. In the case of empirical likelihood estimators, the point estimation $\widehat{\theta}$ is the one that minimizes the discrepancy between $\widehat{p}_t$ and uniform weights.

Some measures of divergence employed in the literature are the Kullback-Leibler divergence and the entropy measure. This problem of minimum contrast can be formulated in the form of moment conditions $E(g(\theta_0, X_t)) = 0$, by employing a modified condition in the form of Eq. 3.4 and the minimum contrast estimator is obtained with the use of some contrast function h_T:

$$\widehat{\theta}_n = \arg\min_{\theta, p_t} \sum_{t=1}^{T} h_T(p_t). \tag{3.8}$$

An important result is that an adequate choice of the discrepancy function can lead to a unified representation of empirical maximum likelihood and minimum contrast estimators. This representation can be obtained when the function $h_T(p_t)$ belongs to the Cressie-Read family of discrepancies given by:

$$h_T(p_t) = \frac{[\gamma(\gamma+1)]^{-1}(Tp_t)^{\gamma+1} - 1]}{T} \tag{3.9}$$

and with restrictions on the definition of the Cressie-Read discrepancy, there are particular cases of several classes of estimators. The empirical likelihood is obtained with the restriction $\gamma \to 0$ in the discrepancy function $h_T(p_t)$; the generalized minimum contrast method, known as exponential tilting (ET) of Kitamura and Stutzer (1997) and Imbens et al. (1998), is obtained by $\gamma \to -1$ and the Continuous Updating estimator employing the empirical likelihood formulation is obtained by $\gamma \to 1$.

Smith (2001) demonstrated that it is possible to define another estimator that also includes these estimators as particular cases. The method of generalized empirical likelihood (GEL) of (Smith (2001)) is obtained as a solution for the following saddle point problem:

$$\widehat{\theta}_n = \arg\min_{\theta} \left[\max_{\lambda} \frac{1}{T} \sum_{t=1}^{T} \rho\left(\lambda' g^w(\theta, x_t)\right) \right], \tag{3.10}$$

where λ defines Lagrange multipliers associated to restriction:

$$\sum_{t=1}^{T} p_t g^w(\theta, x_t) = 0. \tag{3.11}$$

Estimators are obtained solving the previous equation with the first-order condition:

$$\sum_{t=1}^{T} p_t \lambda' \left(\frac{\partial g^w(\theta, x_t)}{\partial \theta} \right) = 0, \tag{3.12}$$

where:

$$p_t = \frac{1}{T} \rho' \left(\lambda' g^w(\theta, x_t) \right). \tag{3.13}$$

This generalized likelihood estimator includes the empirical likelihood estimator, assuming the same conditions on γ of the Cressie-Read divergence function, and modifying functions h and ρ. The empirical likelihood estimator is obtained with $h(p) = -ln\, np$ and $\rho(\xi) = ln(1-\xi)$, the estimator of exponential tilting with $h(p) = np\, ln\, np$ and $\rho(\xi) = -exp(\xi)$, the estimator of continuous updating with $h(p) = (np)^2$ and $\rho(\xi) = -(1+\xi)^2/2^3$. The solution can be obtained by numerical optimization or via quasi-Newton iterative methods, and the solution can be formulated in a problem of a smaller dimension by means of a dual formulation (Kitamura (2006)), which is the method used in this study.

An additional class of estimators can be obtained by combining the empirical likelihood estimator and the exponential tilting estimator, generating the estimator known as *exponentially tilted empirical likelihood (ETEL)* proposed by Schennach (2007). This estimator is defined as:

$$\widehat{\theta} = \arg\min_{\theta} \left(T^{-1} \sum_{t=1}^{T} \widetilde{h}(\hat{p}_t(\theta)) \right), \tag{3.14}$$

where $\widehat{p}_i(\theta)$ is the solution of:

[3]See Table 1 in Smith (2001) for further details

$$\min_{\{p_t\}_{t=1}^{T}} T^{-1}\sum_{t=1}^{T} h(p_t) \tag{3.15}$$

subject to $\sum_{t=1}^{T} p_t g(\theta, x_t) = 0$ and $\sum_{t=1}^{T} p_t = 1$, where $\widetilde{h}(p_t) = -ln(Tp_t)$ and $h(p_t) = Tp_t ln(Tp_t)$.

Note that the ETEL (exponentially tilted empirical likelihood) estimator employs the exponential tilting method to find probabilities $\widehat{w}_i(\theta)$ and the empirical likelihood method to estimate the parameter vector $\widehat{\theta}$. These probabilities are related to multipliers λ through the relation:

$$\widehat{p_t}(\theta) = \frac{\left(\widehat{\lambda}(\theta)' g(\theta, x_t)\right)}{\sum_{i=1}^{T}\left(\widehat{\lambda}(\theta)' g(\theta, x_i)\right)}. \tag{3.16}$$

An important property of the ETEL class of estimators is their behavior in the presence of incorrect specification. Imbens et al. (1998) point out that the empirical likelihood estimator can have inadequate behavior in the presence of incorrect specification, due to the presence of a singularity in its influence function; and theorem 1 in Smith (2001) demonstrates that the asymptotic properties of the empirical likelihood estimator can be severely degraded in the presence of minimum specification problems. This effect also affects the estimations of implicit probabilities $\widehat{p_t}$, because, in the presence of specification problems, the implicit probabilities in likelihood problems tend to concentrate in the extreme observations, in opposition to what is expected from a robust estimator.

The result obtained by Smith (2001) is that, in the class of minimum discrepancy estimators, only the exponential tilting estimator has adequate behavior in the presence of specification problems because its influence function does not present singularities. As the ETEL estimator is a combination of empirical likelihood estimators and of the exponential tilting estimator, it maintains the characteristics of asymptotic efficiency and minimum bias of estimator EL, and, additionally, it is robust in the presence of specification problems, due to the use of the exponential tilting estimator to estimate the implicit probabilities, as shown in theorems

8-10 in Smith (2001), indicating that this estimator is $\sqrt{n}$ consistent even in the presence of specification problems.

However, the interpretation of the results of the ETEL estimator should be literally interpreted as an estimation based on pseudo-true values, and in this case the convergence properties in the misspecified models refer to normal rates of convergence for the pseudo-true value that minimizes the (Kullback-Leibler) distance between the true conditional distribution of the generating process and a conditional distribution in function of these pseudo-true values, e.g. Gourieroux and Monfort (1995).

The usual interpretation of robustness properties for moment conditions estimators when the data are generated by a perturbed version in a infinitesimal neighborhood of the true model can be found in Kitamura et al. (2009) for IID data and Evdokimov et al. (2009) for weakly dependent data, and are based on the use of Hellinger distance in the construction of contrast function 3.5. The use of Hellinger distance in estimation of Stochastic Differential Equations was already explored in Giet and Lubrano (2008), and properties of the Hellinger distance estimator formulated as a generalized minimum contrast to the estimation stochastic differential equations is a possibility to be explored.

We can now sum up some common properties of the estimators discussed in this study. The first property is that all the estimators presented (two-stage GMM, Iterative GMM, continuous updating GMM, generalized empirical likelihood, exponential tilting and exponentially tilted empirical likelihood) have the same properties of consistency and first-order asymptotic efficiency (e.g. Smith (2001), Schennach (2007)), they are efficient in the semi-parametric sense of Bickel et al. (1993), in the validity of specified moment conditions. All the estimators have the same asymptotic variance, but the superior results in terms of bias and asymptotic properties of higher order are valid for the estimators based on generalized empirical likelihood, exponential tilting and exponentially tilted empirical likelihood (e.g. Kitamura (2006)). The class of estimators based on empirical likelihood also presents optimal properties in term of hypotheses tests: these tests are optimum in the minimax and large deviation criteria and are uniformly more powerful in the generalized sense of Neyman-Person, as demonstrated in Kitamura (2006).

However, the performance in finite samples can be rather different. The two-stage GMM estimator can be severely biased in the sizes of the sample employed in economics and finance, and continuous updating estimators are numerically unstable due to the existence of multiple modes in the objective function, for example, Hansen et al. (1996)). Newey and Smith (2004) demonstrate that the empirical likelihood estimator must have a bias in finite samples smaller than the bias of estimators of the exponential tilting and continuous updating classes. In empirical likelihood and exponential tilting estimators, the bias does not grow with the number of moment conditions, as happens with the GMM estimator. Newey and Smith (2004) also demonstrate that estimators based on GEL have good properties in terms of second-order bias. Another interesting property is that estimators based on GMC and GEL are invariant to linear transformation in the moment conditions vector, which does not occur with the two-stage GMM estimator. The finite sample properties of these estimators in in the estimation of stochastic differential equations are explored below.

4. Monte Carlo Evidence - Unconditional Moments

As all these methods are first order and asymptotically equivalent, to perform an analysis of the finite sample properties of these estimators, we performed a Monte Carlo analysis evaluating several properties of these estimators, particularly the bias, mean squared error and mean absolute error, accompanied by a discussion about their validity in the presence of incorrect specification in the context of estimation of stochastic differential equations.

The Monte Carlo procedure consists in simulating Generalized CIR models, Vasicek and CIR SR[4] models, performing the estimation with the proposed estimation methods using the unconditional moment conditions defined in section 2, and, based on these estimations, evaluating the bias, mean square error (MSE) and the mean absolute error (MAE) in relation to each parameter estimated. Figures 4.1, 4.2, 4.3 and 4.4 show MSE and MAE sequentially for each parameter and each method, for a more easy visualization of results.

[4]These experiments were performed for the other models as well and produce similar results, but are not presented here for reasons of space.

The simulation procedure employed for the Generalized CIR process employs a Milstein discretization (e.g. Milstein (1974), Kloeden and Platen (1992)) to generate process trajectories, since in this process there is no exact analytical solution for the transition density. For the Vasicek and CIR SR processes, we employed the exact transition density to generate simulated trajectories (e.g. Ait-Sahalia (2002)).

Note that this detail is of fundamental importance. Before discussing this point, we will introduce the notation of strong convergence of discretizations. Suppose that we want to generate a trajectory of the stochastic differential equation $dX_t = \mu(t, X_t) + \sigma(t, X_t)dW_t$ employing a discretization that generates trajectories $Y_t^{\triangle}$ of this process, and that the trajectories of this approximation converged to the true trajectory. An approximation $Y_t^{\triangle}$ is said to be strong order convergent $\gamma > 0$ if there are positive constants K and γ so that each $\triangle$ is valid:

$$E\left(|X_t - Y_t^{\triangle}|\right) \leq K\triangle^{\gamma},$$

in which K does not depend on the discretization interval $\triangle$. On usual Lipschitz and growth conditions, it is possible to demonstrate (e.g. Kloeden and Platen (1992), Prakasa Rao (1999)) that the Euler discretization converges with strong order $\gamma = 0.5$ and the Milstein discretization (Milstein (1974)) is strong order convergent with $\gamma = 1$.

As the discretization employed in moment conditions is of strong order inferior to that employed in the process simulation, an incorrect specification problem arises generated by the discretization employed. This problem occurs in a more intense form when the exact solution of the stochastic differential equation can be used to generate the process trajectory. The fundamental point is that, in the estimation based on approximated discretizations, there is always a bias generated by the process discretization, and one of the objectives of the Monte Carlo study is to verify whether any method manages to produce a reduction in the bias in relation to this effect, which can be interpreted as a specification problem. Note that in Chan et al. (1992)'s original article, the discretization employed is still simpler than Euler's, and thus the existing bias in the estimators must be even greater.

	GMM2S	GMMITER	GMMCUE	GEL	ET	GELCUE	ETEL	SGEL	SET	SGELCUE	SETEL
mean α	0.0585	0.0585	0.0585	0.0555	0.0591	0.0594	0.0427	0.0529	0.0584	0.0592	0.0419
bias α	0.0177	0.0177	0.0177	0.0147	0.0183	0.0186	0.0019	0.0121	0.0176	0.0184	0.0011
mse α	0.0006	0.0006	0.0006	0.0005	0.0006	0.0006	0.0001	0.0004	0.0006	0.0006	0.0001
mae α	0.0190	0.0190	0.0190	0.0171	0.0198	0.0199	0.0055	0.0147	0.0194	0.0198	0.0063
mean β	−0.8595	−0.8595	−0.8595	−0.8095	−0.8696	−0.8736	−0.5588	−0.7687	−0.8596	−0.8720	−0.5719
bias β	−0.2674	−0.2674	−0.2674	−0.2174	−0.2775	−0.2815	0.0333	−0.1766	−0.2675	−0.2799	0.0202
mse β	0.1347	0.1347	0.1347	0.1142	0.1478	0.1462	0.0060	0.0874	0.1452	0.1465	0.0058
mae β	0.2884	0.2884	0.2884	0.2599	0.3050	0.3057	0.0501	0.2229	0.3017	0.3059	0.0533
mean σ^2	2.0247	2.0247	2.0247	1.5815	1.3286	1.3440	1.7024	1.6174	1.3095	1.3495	1.7090
bias σ^2	0.3543	0.3543	0.3543	−0.0889	−0.3418	−0.3264	0.0320	−0.0530	−0.3609	−0.3209	0.0386
mse σ^2	2.9527	2.9527	2.9527	1.4709	1.9360	2.1132	0.0045	0.9256	1.4653	2.3160	0.0143
mae σ^2	0.7768	0.7768	0.7768	0.7532	0.3050	1.0843	0.0447	0.6088	0.9348	1.0786	0.0597
mean γ	1.4939	1.4939	1.4939	1.4426	1.3792	1.3749	1.5450	1.4612	1.3880	1.3790	1.5450
bias γ	−0.0060	−0.0060	−0.0060	−0.0573	−0.1207	−0.1250	0.0451	−0.0387	−0.1119	−0.1209	0.0451
mse γ	0.0263	0.0263	0.0263	0.0206	0.0413	0.0445	0.0049	0.0146	0.0367	0.0477	0.0050
mae γ	0.0996	0.0996	0.0996	0.1079	0.1667	0.1746	0.0479	0.0876	0.1547	0.1747	0.0507

TABLE 2. Monte Carlo - Generalized CIR Model - α =0.0408, β =-0.5921, $\sigma^2 = 1.6704$, γ =1.4999.

The first Monte Carlo experiment corresponds to the simulation of 1,000 trajectories of size 474 of a Generalized CIR process with a parameter vector given by α =0.0408, β =-0.5921, σ^2 =1.6704 and γ =1.4999. This set of parameters, and all other parameters used in Monte Carlo analysis, are based on estimated values in the article by Chan et al. (1992) for the series of Treasury Bills, and thus reflect values consistent with real data. The results of this experiment are displayed in Table 2 and Figure 4.1. Each figure shows respectively the bias and MSE obtained by each estimator. The results obtained demonstrate that there is a relevant bias in the estimation of all the parameters, and particularly of parameter σ^2. The results in terms of the size of the bias and of the mean square error are quite similar for almost all the estimators, except for estimators ETEL and SETEL, which present far superior results in terms of bias, MSE and MAE in relation to the other methods for all the parameters estimated, which is evident in Figure 4.1.

In the Monte Carlo experiment for the Vasicek process (Table 3 and Figure 4.2), we simulated again 1,000 trajectories with a parameter vector given by $\alpha = 0.0154$, $\beta = -0.1779$, $\sigma^2 = 0.0004$ and $\gamma = 0$. The results indicate again that the ETEL estimators' performance is superior, and it is also noticeable that, in this experiment, the estimator with the worst performance was the estimator GMMCUE. For the CIR SR process (Table 4 and Figure 4.3), we simulated trajectories of the process with $\alpha = 0.0189, \beta = -0.2339, \sigma^2 = 0.0073$ and $\gamma = 0.5$. The same behavior of better performance of the ETEL class of estimators was observed, as well as a similar performance of the other estimators.

	GMM2S	GMMITER	GMMCUE	GEL	ET	GELCUE	ETEL	SGEL	SET	SGELCUE	SETEL
mean α	0.0263	0.0263	0.0330	0.0235	0.0262	0.0262	0.0181	0.0220	0.0262	0.0263	0.0174
bias α	0.0109	0.0109	0.0176	0.0081	0.0108	0.0108	0.0027	0.0066	0.0108	0.0109	0.0020
mse α	0.0003	0.0003	0.0271	0.0002	0.0003	0.0003	<1e-4	0.0002	0.0003	0.0004	<1e-4
mae α	0.0129	0.0129	0.0195	0.0101	0.0128	0.0129	0.0035	0.0086	0.0128	0.0130	0.0036
mean β	−0.3031	−0.3034	−0.3096	−0.2668	−0.3015	−0.3018	−0.1705	−0.2470	−0.3022	−0.3035	−0.1701
bias β	−0.1252	−0.1255	−0.1317	−0.0889	−0.1236	−0.1239	0.0074	−0.0691	−0.1243	−0.1256	0.0078
mse β	0.0408	0.0412	0.0850	0.0257	0.0405	0.0410	<1e-4	0.0188	0.0405	0.0422	0.0001
mae β	0.1442	0.1446	0.1517	0.1087	0.1429	0.1437	0.0075	0.0897	0.1434	0.1451	0.0079
mean σ^2	0.0004	0.0004	0.0042	0.0002	0.0004	0.0004	−0.0007	0.0002	0.0004	0.0004	−0.0004
bias σ^2	<1e-4	<1e-4	0.0038	−0.0002	<-1e-4	<-1e-4	−0.0011	−0.0002	<-1e-4	<-1e-4	−0.0008
mse σ^2	<1e-4	<1e-4	0.0089	<1e-4	<1e-4	<1e-4	<1e-4	<1e-4	<1e-4	<1e-4	<1e-4
mae σ^2	<1e-4	<1e-4	0.0038	0.0002	0.1429	<1e-4	0.0017	0.0003	<1e-4	<1e-4	0.0022

TABLE 3. Monte Carlo - Vasicek Model - $\alpha = 0.0154$, $\beta = -0.1779$, $\sigma^2 = 0.0004$, $\gamma = 0$.

	GMM2S	GMMITER	GMMCUE	GEL	ET	GELCUE	ETEL	SGEL	SET	SGELCUE	SETEL
mean α	0.0273	0.0274	0.0273	0.0263	0.0268	0.0272	0.0354	0.0241	0.0269	0.0273	0.0322
bias α	−0.0116	−0.0115	−0.0116	−0.0126	−0.0121	−0.0117	−0.0035	−0.0148	−0.0120	−0.0116	−0.0067
mse α	0.0003	0.0003	0.0003	0.0003	0.0003	0.0003	0.0002	0.0003	0.0003	0.0003	0.0002
mae α	0.0139	0.0139	0.0139	0.0143	0.0143	0.0140	0.0095	0.0161	0.0143	0.0139	0.0114
mean β	−0.3536	−0.3549	−0.3547	−0.2539	−0.3438	−0.3520	−0.2243	−0.2501	−0.3441	−0.3543	−0.2222
bias β	−0.1197	−0.1210	−0.1208	−0.0200	−0.1099	−0.1181	0.0096	−0.0162	−0.1102	−0.1204	0.0117
mse β	0.0374	0.0382	0.0382	0.0086	0.0337	0.0373	0.0003	0.0077	0.0346	0.0381	0.0003
mae β	0.1468	0.1482	0.1480	0.0497	0.1377	0.1458	0.0110	0.0457	0.1387	0.1478	0.0128
mean σ^2	0.0072	0.0072	0.0072	0.0077	0.0073	0.0072	0.0103	0.0088	0.0073	0.0072	0.0115
bias σ^2	<-1e-4	<-1e-4	<-1e-4	0.0004	<-1e-4	<-1e-4	0.0030	0.0015	<-1e-4	<-1e-4	0.0042
mse σ^2	<1e-4	<1e-4	<1e-4	<1e-4	<1e-4	<1e-4	<1e-4	<1e-4	<1e-4	<1e-4	<1e-4
mae σ^2	0.0004	0.0004	0.0004	0.0009	0.1377	0.0004	0.0041	0.0019	0.0004	0.0004	0.0054

TABLE 4. Monte Carlo - CIR-SR Model - $\alpha = 0.0189, \beta = -0.2339, \sigma^2 = 0.0073$, $\gamma = 0.5$.

Note that, so far, the problem of incorrect specification was caused only by the use of approximated discretization in the construction of the process' moment conditions. In order to verify whether the better performance properties of the ETEL class of estimators are valid in more general situations of incorrect specification, we employed, as data generating process, trajectories of the Generalized CIR process with parameter vector α =0.0408, β =-0.5921, σ^2 =1.6704 and γ =1.4999. However, as specification of the estimated model, we now employed a CIR SR model assuming that $\gamma = .5$.

The results of this experiment (Table 5 and Figure 4.4) indicate that, in this general case, a better performance of ETEL and ET estimators also occurs, but the other estimators have a much worse performance in relation to the estimation of parameter σ^2. Note that the problem of incorrect specification is expected, in this situation, to affect mainly the estimation of the process variance, because, in the classes of CIR models, the volatility is a function of the level of process with parameter γ.

	GMM2S	GMMITER	GMMCUE	GEL	ET	GELCUE	ETEL	SGEL	SET	SGELCUE	SETEL
mean α	0.0589	0.0688	0.0461	0.0198	0.0312	0.0426	0.0187	0.0167	0.0313	0.0548	0.0210
bias α	0.0181	0.0280	0.0053	−0.0210	−0.0096	0.0018	−0.0221	−0.0241	−0.0095	0.0140	−0.0198
mse α	0.0009	0.0015	0.0020	0.0007	0.0024	0.0034	0.0032	0.0007	0.0114	0.0025	0.0119
mae α	0.0217	0.0334	0.0366	0.0236	0.0227	0.0501	0.0230	0.0249	0.0252	0.0434	0.0271
mean β	−0.8801	−1.0498	−0.6860	−0.2530	−0.3515	−0.6374	−0.2211	−0.2307	−0.3347	−0.8366	−0.2279
bias β	−0.2880	−0.4577	−0.0939	0.3391	0.2406	−0.0453	0.3710	0.3614	0.2574	−0.2445	0.3642
mse β	0.1998	0.4041	0.5297	0.1449	0.1839	0.8179	0.1387	0.1368	0.1857	0.6700	0.1675
mae β	0.3408	0.5425	0.5972	0.3739	0.3916	0.8046	0.3710	0.3685	0.3848	0.7148	0.3753
mean σ^2	0.0081	0.0081	0.0082	0.0093	0.0269	0.0107	0.0147	0.0136	0.0140	0.0085	0.0188
bias σ^2	−1.6623	−1.6623	−1.6622	−1.6611	−1.6435	−1.6597	−1.6557	−1.6568	−1.6564	−1.6619	−1.6516
mse σ^2	2.7632	2.7633	2.7628	2.7593	2.9352	2.7588	2.7417	2.7452	2.7466	2.7621	2.7358
mae σ^2	1.6623	1.6623	1.6622	1.6611	0.3916	1.6604	1.6557	1.6568	1.6565	1.6619	1.6539

TABLE 5. Monte Carlo - Misspecified Model - α =0.0408, β =-0.5921,σ^2 =1.6704, γ =1.4999.

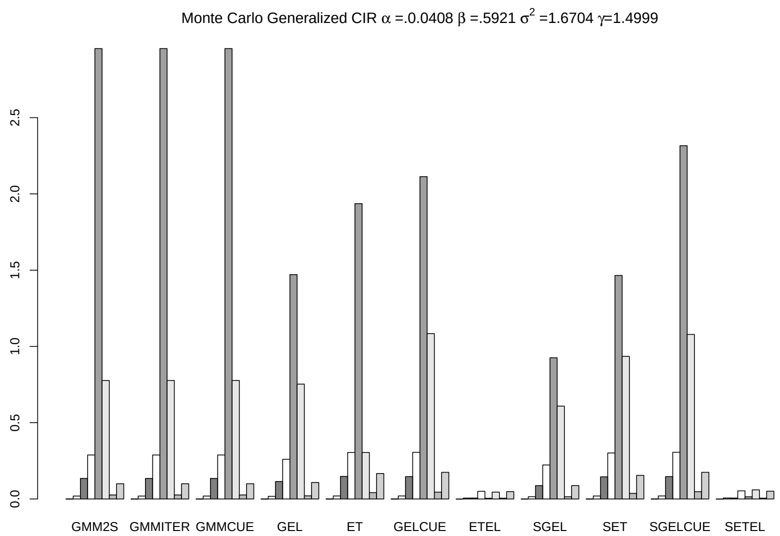


FIGURE 4.1. Monte Carlo Generalized CIR Model

To show the effect of sample size in the estimation, we performed the same Monte Carlo procedure for the CIR SR model[5], now using a sample size of 2000, shown in Table 6. As expected, the results indicate that all the estimators show substantial reductions in the bias, MSE and MAE, but still are dominated by ETEL estimator. The results of this table also show that GMM based estimators on need a larger sample size to achieve a performance close to the estimators based on GEL/GMC, as is evident in this table.

Although it is interesting to analyze the effect of discretization on the properties of estimators, there is a simple way to perform this analysis, since it is impossible to separate the

[5]We perform this same study for other models studied and the results are similar.

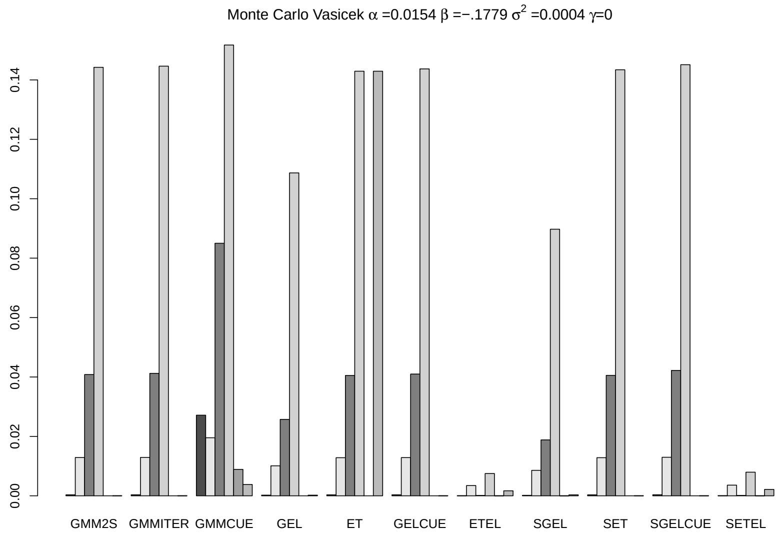


Figure 4.2. Monte Carlo Vasicek Model

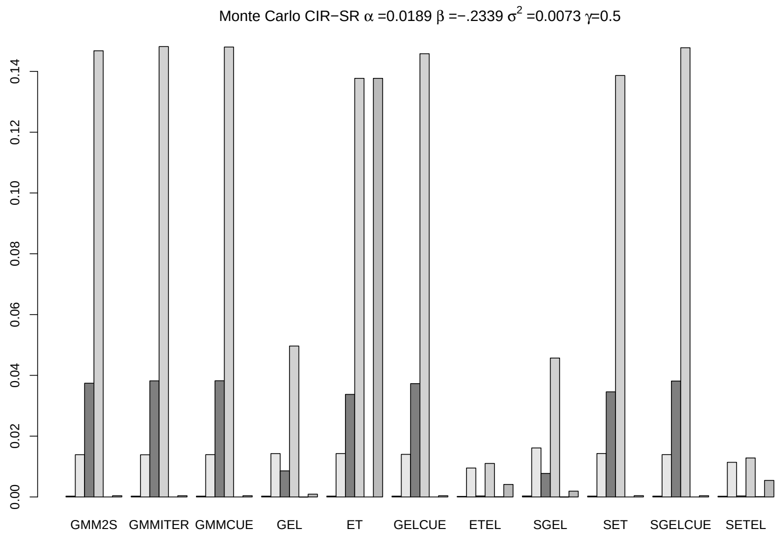


Figure 4.3. Monte Carlo CIR-SR Model

effect of sample size in this analysis. For example a study with a sample size of 500 with discretization interval of 1/12, as studied in this article, would be equivalent to a sample of 41.66 years. A sample discretization of 1/365 (daily data) with 500 observations covers a period of 1.36 years, a very limited time span to analyze a series of interest rates, and very sensitive to

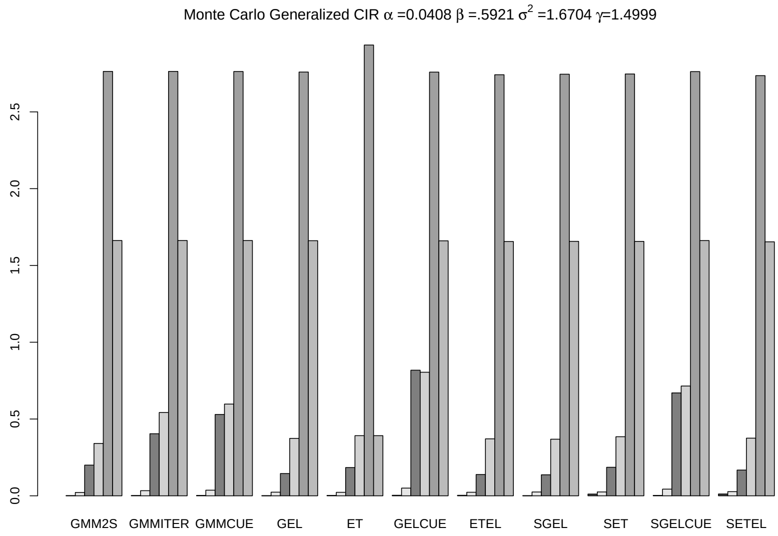


FIGURE 4.4. Monte Carlo Misspecified CIR Model

	GMM2S	GMMITER	GMMCUE	GEL	ET	GELCUE	ETEL	SGEL	SET	SGELCUE	SETEL
mean α	0.0209	0.0209	0.0209	0.0209	0.0209	0.0209	0.0194	0.0206	0.0209	0.0209	0.0194
bias α	0.0020	0.0020	0.0020	0.0020	0.0020	0.0020	0.0005	0.0017	0.0020	0.0020	0.0005
mse α	<1e-4	<1e-4	<1e-4	<1e-4	<1e-4	<1e-4	<1e-4	<1e-4	<1e-4	<1e-4	<1e-4
mae α	0.0040	0.0040	0.0040	0.0039	0.0040	0.0040	0.0022	0.0036	0.0040	0.0040	0.0019
mean β	−0.2621	−0.2623	−0.2623	−0.2611	−0.2620	−0.2619	−0.2257	−0.2573	−0.2622	−0.2624	−0.2260
bias β	−0.0282	−0.0284	−0.0284	−0.0272	−0.0281	−0.0280	0.0082	−0.0234	−0.0283	−0.0285	0.0079
mse β	0.0049	0.0049	0.0049	0.0047	0.0049	0.0049	0.0002	0.0041	0.0049	0.0050	0.0002
mae β	0.0528	0.0529	0.0529	0.0508	0.0527	0.0527	0.0094	0.0474	0.0527	0.0530	0.0096
mean σ^2	0.0073	0.0073	0.0073	0.0073	0.0073	0.0073	0.0073	0.0073	0.0073	0.0073	0.0075
bias σ^2	<-1e-4	<-1e-4	<-1e-4	<-1e-4	<-1e-4	<-1e-4	<1e-4	<1e-4	<-1e-4	<-1e-4	0.0002
mse σ^2	<1e-4	<1e-4	<1e-4	<1e-4	<1e-4	<1e-4	<1e-4	<1e-4	<1e-4	<1e-4	<1e-4
mae σ^2	0.0002	0.0002	0.0002	0.0002	0.0527	0.0002	0.0006	0.0002	0.0002	0.0002	0.0010

TABLE 6. Monte Carlo - CIR-SR Model , Sample Size 2000 - $\alpha = 0.0189, \beta = -.2339, \sigma^2 = 0.0073,\ \gamma = .5$.

initial conditions, since for the usual values of the persistence parameter estimated for models of interest rates (e.g. Chan et al. (1992)) the half-life for the error correction process is much longer than this period, and so the results are very dependent on the initial conditions and not representative.

5. MONTE CARLO EVIDENCE - CONDITIONAL MOMENTS

In the simulations presented so far, the properties of the estimators were studied using simple Euler discretizations of stochastic differential equations. Although this methodology is applicable to any stochastic differential equation without requiring the existence of analytical solutions, for the models with known analytical solutions the use of simple discretizations of the process may represent as inefficient use of the available information when it is possible to derive explicit conditional moments of the process. To analyze the properties of the estimators based on conditional moments we performed two analysis. In the first analysis we build conditional moment conditions obtained by analytical solution of stochastic differential equation for the CIR SR process. In the second we use the Itô Conditional Moment Generator Methodology proposed in Zhou (2003) which gives conditional moments through the use of Generalized Ito's lemma (Merton (1971) and Lo (1988))[6]. In these two procedures we performed a Monte Carlo analysis for the CIR-SR process with the same parameters used in the previous section.

Rewriting the CIR SR model as:

$$dr(t) = \kappa(\theta - r(t))dt + \sigma\sqrt{r(t)}dW(t) \tag{5.1}$$

the transition density of this process is given by the following non-central χ^2 density:

$$p_{r(t)}(x) = p_{\chi^2(\eta,\lambda_t)/c_t}(x) = c_t p_{\chi^2(\eta,\lambda_t)}(c_t x) \tag{5.2}$$

$$with : c_t = \frac{4\kappa}{\sigma^2(1 - exp(-\kappa t))} \tag{5.3}$$

$$\eta = 4\kappa\theta/\sigma^2 \tag{5.4}$$

$$\lambda_t = c_t r_0 exp(-\kappa t). \tag{5.5}$$

Thus we obtain the conditional moments as:

$$E[r_t|r_s] = r_s + r_s e^{-\kappa(t-s)} + \theta\left(1 - e^{-\kappa(t-s)}\right) \tag{5.6}$$

$$Var[r_t|r_s] = r_s\frac{\sigma^2}{\kappa}(e^{-\kappa(t-s)} - e^{-2\kappa(t-s)}) + \theta\frac{\sigma^2}{2\kappa}\left(1 - e^{-\kappa(t-s)}\right)^2 \tag{5.7}$$

[6]This approach was recently generalized in Cuchiero et al. (2010) and Filipovic et al. (2011) to a class known as Polynomial Process, which also include Levy processes in addition to affine diffusion models.

and from these moments create the array of four moment conditions by multiplying these two conditions by r_{t-1}. The Monte Carlo analysis from this specification is shown in Table 7. Compared to the results obtained for the estimation using non-conditional moments (Table 4) results indicate that with this specification we obtain a bias generally smaller (but positive) for the parameter α, but at the cost of a higher bias for the parameter β, while for α parameter the results are equivalent. In terms of mse and mae for all parameters the results are basically equivalent to the estimation using unconditional moments.

	GMM2S	GMMITER	GMMCUE	GEL	ET	GELCUE	ETEL	SGEL	SET	SGELCUE	SETEL
mean α	0.0284	0.0284	0.0284	0.0284	0.0284	0.0284	0.0212	0.0284	0.0284	0.0284	0.0212
bias α	0.0095	0.0095	0.0095	0.0095	0.0095	0.0095	0.0023	0.0095	0.0095	0.0095	0.0023
mse α	0.0002	0.0002	0.0002	0.0002	0.0002	0.0002	0.0001	0.0002	0.0002	0.0002	<1e-4
mae α	0.0113	0.0113	0.0112	0.0112	0.0112	0.0112	0.0047	0.0112	0.0112	0.0113	0.0045
mean β	−0.3619	−0.3619	−0.3619	−0.3612	−0.3608	−0.3617	−0.2452	−0.3605	−0.3617	−0.3620	−0.2452
bias β	−0.1280	−0.1280	−0.1280	−0.1273	−0.1269	−0.1278	−0.0113	−0.1266	−0.1278	−0.1281	−0.0113
mse β	0.0414	0.0414	0.0414	0.0411	0.0412	0.0413	0.0048	0.0409	0.0413	0.0414	0.0047
mae β	0.1486	0.1486	0.1486	0.1479	0.1475	0.1485	0.0308	0.1471	0.1480	0.1486	0.0300
mean σ^2	0.0073	0.0073	0.0073	0.0073	0.0073	0.0073	0.0074	0.0073	0.0073	0.0073	0.0075
bias σ^2	<-1e-4	<-1e-4	<-1e-4	<-1e-4	<-1e-4	<-1e-4	<1e-4	<1e-4	<-1e-4	<-1e-4	0.0002
mse σ^2	<1e-4	<1e-4	<1e-4	<1e-4	<1e-4	<1e-4	<1e-4	<1e-4	<1e-4	<1e-4	<1e-4
mae σ^2	<1e-4	<1e-4	<1e-4	<1e-4	0.1475	<1e-4	<1e-4	<1e-4	<1e-4	<1e-4	0.0003

TABLE 7. Monte Carlo - CIR-SR Model - Conditional moments - $\alpha = 0.0189, \beta = -.2339, \sigma^2 = 0.0073,\ \gamma = .5$.

The second methodology for the construction of conditional moments is based on the Itô Conditional Moment Generator Methodology proposed in Zhou (2003). The methodology considers all the conditional moments of K'th order simultaneously applying the Generalized Itô's lemma to each r_T^k and then takes the conditional expectation:

$$E(r_T^k) = r_t^k + E_t\left[\int_t^T (\mu_u k r_u^{k-1} + \frac{1}{2}\sigma^2 k(k-1) r_u^{k-2}) du\right].$$

Taking the derivative with respect to time T, and interchanging the expectation and integration operators is possible to obtain a system of differential equations of the form:

$$\frac{dE_t(r_s^k)}{ds} = E_t\left[\mu_s k r_s^{k-1} + \frac{1}{2}\sigma^2 k(k-1) r_u^{k-2})\right]$$

subject to boundary condition $E_t(r_t^k) = r_t^k$.

The class of systems that have analytical solutions is the set of processes with solutions in the form:

$$E_t(R_T) = e^{(T-t)A(\beta)} R_t + A^{-1}(\beta) \left(e^{(T-t)A(\beta)} - I \right) g(\beta)$$

$$\frac{dE_t(R_s)}{ds} = A(\beta) + E_t(R_s) + g(\beta)$$

where I is the KxK identity matrix and e denotes the matrix exponential.

For the CIR SR process for the solution of the system consists of the matrices $A(\beta)$ and $g(\beta)$ in the form:

$$A(\beta) = \begin{bmatrix} -\kappa & 0 & 0 & 0 \\ 2\kappa\theta + \sigma^2 & -2\kappa & 0 & 0 \\ 0 & 3\kappa\theta + 3\sigma^2 & -3\kappa & 0 \\ 0 & 0 & 4\kappa\theta + 6\sigma^2 & -4\kappa \end{bmatrix}$$

$$g(\beta) = \begin{bmatrix} \kappa\theta \\ 0 \\ 0 \\ 0 \end{bmatrix}$$

In this example we follow the example in Zhou (2003), who used as moment conditions the moments $E(r_T^k)$ of order K = 1,2,3,4, multiplied by r_{t-1}^k, again with k = 1,2,3,4 with a total of 16 conditional moment conditions. The results of this experiment are shown in Table 8. The overall results compared to results obtained earlier for the CIR SR model, indicate that this method can reduce the bias, mse and mae for the parameters α, but most notably for the methods based on GEL/GMC . However for the parameter β the results are different, showing an increase in bias for the estimators based on GMM and the GEL estimator. Similarly the mse increase for estimators based on GMM and the GEL estimator, and decrease for the others. For the parameter σ^2 the results indicate that this estimation method is slightly worst in terms of bias and mae.

We can make some considerations on the use of conditional moments. The first important point to note is that the class of models that admit conditional moments in analytic form is rather limited, as opposed to the use of Euler discretizations that can be used generally. Another important aspect to note is that the estimators based on the Generalized Ito's lemma shows complexity and computational cost much higher than the other estimators, because it is necessary to evaluate the matrix exponential. In summary, considering the generality and the gains in performance in finite samples, the estimators based on unconditional moments derived from the Euler discretizations are still competing.

	GMM2S	GMMITER	GMMCUE	GEL	ET	GELCUE	ETEL	SGEL	SET	SGELCUE	SETEL
mean α	0.0291	0.0294	0.0286	0.0231	0.0186	0.0184	0.0206	0.0192	0.0184	0.0183	0.0192
bias α	0.0102	0.0105	0.0097	0.0042	−0.0003	−0.0005	0.0017	0.0003	−0.0005	−0.0006	0.0003
mse α	0.0003	0.0003	0.0003	0.0001	<1e-4	<1e-4	<1e-4	<1e-4	<1e-4	<1e-4	<1e-4
mae α	0.0121	0.0124	0.0121	0.0079	0.0028	0.0027	0.0040	0.0031	0.0028	0.0028	0.0031
mean β	−0.3828	−0.3901	−0.3749	−0.2916	−0.2256	−0.2245	−0.2233	−0.2230	−0.2251	−0.2243	−0.2230
bias β	−0.1489	−0.1562	−0.1410	−0.0577	0.0083	0.0094	0.0106	0.0109	0.0088	0.0096	0.0109
mse β	0.0549	0.0600	0.0582	0.0176	0.0003	0.0002	0.0002	0.0002	0.0003	0.0002	0.0002
mae β	0.1717	0.1789	0.1696	0.0868	0.0119	0.0114	0.0114	0.0117	0.0117	0.0111	0.0117
mean σ^2	0.0070	0.0069	0.0070	0.0069	0.0067	0.0066	0.0071	0.0071	0.0065	0.0065	0.0071
bias σ^2	−0.0003	−0.0004	−0.0003	−0.0004	−0.0006	−0.0007	−0.0002	−0.0002	−0.0008	−0.0008	−0.0002
mse σ^2	<1e-4	<1e-4	<1e-4	<1e-4	<1e-4	<1e-4	<1e-4	<1e-4	<1e-4	<1e-4	<1e-4
mae σ^2	0.0008	0.0008	0.0008	0.0020	0.0119	0.0011	0.0010	0.0010	0.0012	0.0013	0.0010

TABLE 8. Monte Carlo - CIR-SR Model - Itô Conditional Moment Generator - $\alpha = 0.0189, \beta = -.2339, \sigma^2 = 0.0073$, $\gamma = .5$.

6. CONCLUSIONS

In this article we consider semi-parametric methods based on the empirical likelihood/generalized minimum contrast for the estimation of stochastic differential equations. These estimators are characterized by properties of asymptotic efficiency of higher order, properties of optimality in hypotheses testing and robustness of the estimators based on exponential tilting in relation to incorrect specification. These properties are particularly important in the context of estimation of stochastic differential equations, since, in general, it is not possible to construct the exact likelihood function of the process due to the non-existence of analytical solutions (and consequently of exact discretizations) for stochastic differential equations. These methods allow to approximate the density of these processes using a nonparametric approximation of the log-likelihood, allowing for the incorporation of this information into the estimates of the parameters of the stochastic differential equations.

The results indicate that these methods yields good properties in finite samples, achieving an overall performance superior to the generalized methods of moments usually employed in the estimation of stochastic differential equations using moment conditions. The results also indicate that estimators based on the unconditional moments derived from the Euler discretizations have a performance comparable to the same estimators constructed from conditional moments derived from transition densities. However, the last conditional moments can only be derived in analytical form for a limited class of processes. Since the Euler discretizations are easily obtained for almost all stochastic differential equations, the use of methodologies based on empirical likelihood/generalized minimum contrast allows computationally simple estimators with good properties in terms of bias and efficiency for a wide class of processes.

The results obtained also indicate that the exponentially tilted empirical likelihood estimator, in particular the one proposed by Schennach (2007), obtains a performance which is superior to other proposed techniques, due to its properties of robustness in the presence of specification problems. As it is possible to interpret the estimation of the stochastic differential equations by employing discrete data as an incorrect specification problem, due to the use of an approximated discretization of the model, the results of the Monte Carlo experiments demonstrate that the performance of this estimator is quite superior to the other estimation methods employing moment conditions. Also, in general, the estimators based on empirical likelihood/generalized minimum contrast have a better performance in terms of bias and mean square error than the GMM estimators.

Acknowledgments

This research was partially supported by FAPESP and CNPq. The authors would like to thank the Associate editor and the reviewers for their valuable comments and suggestions.

References

Ait-Sahalia, Y. (2002). Maximum-likelihood estimation of discretely-sampled diffusions: A closed-form approximation approach. *Econometrica 70*, 223–262.

Anatolyev, S. (2005). Gmm, gel, serial correlation and asymptotic bias. *Econometrica 73*, 983–1002.

Anatolyev, S. and N. Gospodinov (2011). *Methods for Estimation and Inference in Modern Econometrics*. Chapman and Hall/CRC.

Andrews, D. W. K. (1991). Heteroskedasticity and autocorrelation consistent covariance matrix estimation. *Econometrica 59*, 817–858.

Bachelier, L. (1900). Theorie de la speculation. English translation by A J Boness in *The Random Character of Stock Market Prices*, ed. Paul H Cootner, pg 17–78, Cambridge, Mass, MIT press 1967.

Bickel, P., C. Klassen, Y. Ritov, and J. Wellner (1993). *Efficient and Adaptative Estimation for Semiparametric Models*. Johns Hopkins Press.

Bishwal, J. P. N. (2007). *Parameter Estimation in Stochastic Differential Equations*. Springer.

Black, F. and M. S. Scholes (1973). The pricing of options and corporate liabilities. *Journal of Political Economy 7*, 637–54.

Brennan, M. and E. S. Schwartz (1980). Analyzing convertible bonds. *Journal of Financial and Quantitative Analysis 15*, 907–929.

Chan, K. G., G. Karolyi, F. Longstaff, and A. B. Sanders (1992). An empirical comparasion of alternative models of term structure of interest rates. *Journal of Finance 47*, 1209–1227.

Cox, J. (1975). Notes on option pricing i: constant elasticity of variance diffusions. Working Paper, Stanford University.

Cox, J. C., J. . E. Ingersoll, and S. A. Ross (1985). A theory of the term structure of interest rates. *Econometrica 53*, 385–408.

Cox, J. C., J. E. Ingersoll, and S. A. Ross (1980). An analysis of variable rate loan contracts. *Journal of Finance 35*, 389–403.

Cuchiero, C., M. Keller-Ressel, and J. Teichmann (2010). Polynomial processes and their applications in mathematical finance. *Finance and Stochastics*. forthcoming.

DasGupta, A. (2008). *Asymptotic Theory of Statistics and Probability*. Sprimger.

Delbaen, F. and W. Schachermayer (1994). A general version of the fundamental theory of asset pricing. *Mathematische Annalen 300*, 463–250.

Dothan, U. L. (1978). On the term structure of interest rates. *Journal of Financial Economics 6*, 59–59.

Evdokimov, K., Y. Kitamura, and T. Otsu (2009). Robust estimation of moment condition models with weakly dependente data. Unpublished Working Paper.

Filipovic, D., M. E, and P. Schneider (2011). Density approximations for multivariate affine jump-diffusion processes. Unpublished Working Paper.

Giet, L. and M. Lubrano (2008). A minimum hellinger distance estimator for stochastic differential equations: An application to statistical inference for continuous time interest rate models. *Computational Statistics & Data Analysis 52*(6), 2945 – 2965.

Gourieroux, C. and A. Monfort (1995). *Statistics and Econometric Models.* Cambridge University Press.

Gourieroux, C. and A. Monfort (1996). *Simulation-Based Econometric Models.* Oxford University Press.

Hansen, L. P. (1982, July). Large sample properties of Generalized Method of Moments estimators. *Econometrica 50*, 1029–1054.

Hansen, L. P., J. Heaton, and A. Yaron (1996). Finite sample properties od some alternative gmm estimators. *Journal of Business and Economic Statistics 14*, 262–280.

Harrison, J. M. and D. Kreps (1979). Martingales and arbitrage in multiperiod securities markets. *Journal of Economic Theory 20*, 381–408.

Harrison, J. M. and S. Pliska (1981). Martingales and stochastic integrals in the theory of continous trading. *Stochastic Processes and Their Applications 11*, 215–260.

Imbens, G. W., R. H. Spady, and P. Johnson (1998). Information theoretic approaches to inference in moment conditions models. *Econometrica 66*, 333–357.

Karatzas, I. and S. E. Shreve (1987). *Brownian Motion and Stochastic Calculus.* Springer–Verlag.

Kitamura, Y. (1997). Empirical likelihood methods with weakly dependent processes. *Annals of Statistics 25*, 2084–2102.

Kitamura, Y. (2006). Empirical likelihood methods in econometrics: Theory and practice. Unpublished Working Paper.

Kitamura, Y., T. Otsu, and K. Evdokimov (2009). Robustness, infinitesimal neighborhoods, and moment restrictions. Cowles Foundation Discussion Paper 1720.

Kitamura, Y. and M. Stutzer (1997). An information-theoretic alternative to generalized method of moments estimation. *Econometrica 65(5)*, 861–874.

Kloeden, P. and E. Platen (1992). *Numerical Solution of Stochastic Differential Equations.* Springer–Verlag.

Lo, A. (1988). Maximum likelihood estimation of generalized itô process with discretely sampled data. *Econometric Theory 4*, 231–247.

Merton, R. (1971). Optimum consumption and portofolio rules in a continuos time model. *Journal of Economic Theory 3*, 373–413.

Merton, R. C. (1973). The theory of rational option pricing. *Bell Journal 4*, 141–183.

Milstein, G. N. (1974). Aproximate integration of stochastic diferential equations. *Theory of Probability and Applications 19*, 557–562.

Newey, W. and D. McFadden (1994). *Handbook of Econometrics*, Volume 4, Chapter Large sample estimation and hypothesis testing. Elsevier.

Newey, W. and R. J. Smith (2004). High-order properties of gmm and generalized empirical likelihood estimators. *Econometrica 72*, 219–255.

Newey, W. K. and K. D. West (1987). A simple, positive semi-definite, heteroskedasticity and autocorrelation consistent covariance matrix. *Econometrica 55*, 703–708.

Owen, A. (1991). Empirical likelihood for linear models. *The Annals of Statistics 18(1)*, 1725–1747.

Pedersen, A. R. (1995). A new approach to maximum likelihood estimation for stochastic differential equations based on discrete observations. *Scandinavian Journal of Statistics*.

Prakasa Rao, B. L. S. (1999). *Statistical Inference for Diffusion Type Process.* Arnold.

Qin, J. and J. Lawless (1994). Empirical likelihood and general estimating equations. *The Annals of Statistics 20(1)*, 300–325.

Rogers, L. C. G. and D. Williams (2000). *Diffusions, Markov Process and Martingales: Volume 2 Itô Calculus.* Cambridge.

Schennach, S. (2007). Point estimation with exponentially tilted empirical likelihood. *Annals of Statistics 35(2)*, 634–672.

Smith, R. J. (2001). Gel criteria for moment conditions models. Working Paper, University of Bristol.

Vasicek, O. (1977). An equilibrium characterization of the term structure. *Journal of Financial Economics 5*, 177–88.

Zhou, H. (2003). Ito conditional moment generator and the estimation of short rate processes. *Journal of Financial Econometrics 1*, 250–271.

Zivot, E. and J. Wang (2006). *Modeling Financial Time Series with S-PLUS, second edition.* Springer-Verlag.

ESTIMATION OF STOCHASTIC VOLATILITY MODELS USING METHODS OF GENERALIZED EMPIRICAL LIKELIHOOD/MINIMUM CONTRAST

MÁRCIO POLETTI LAURINI
IBMEC-RJ
LUIZ KOODI HOTTA
IMECC-UNICAMP

ABSTRACT

In this article we discuss the estimation of Stochastic Volatility (SV) models using generalized empirical likelihood/minimum contrast methods. We show via Monte Carlo simulations that the proposed methods have a superior or equivalent performance to the other estimation methods proposed in the literature to estimate SV models, and, additionally, they offer robustness properties in the presence of specification problems such as heavy-tailed distributions and the presence of outliers.
Keywords: Stochastic Volatility, Empirical Likelihood, Minimum Contrast, Robustness.
Keywords: Volatilidade Estocástica, Verossimilhança Empírica, Mínimo Contraste, Robustez.
JEL Codes: C14, C22.

Márcio Poletti Laurini - IBMEC-RJ. Email - marcio.laurini@ibmecrj.br.
Luiz Koodi Hotta - Imecc Unicamp - Email - hotta@ime.unicamp.br.

1. Introduction

Measurement of asset volatility is a fundamental aspect of finance. Precise volatility measurements in financial asset returns are necessary in certain aspects, such as risk management (McNeil et al. (2005)) and asset pricing (Singleton (2006)). Among the available forms for modeling volatility, the class of models known as SV models stands out[1]. In this class of models, volatility is treated as a non-observed latent factor. One of the main reason for its popularity is that SV models can be derived from continuous time diffusions (e.g. Barndorff-Nielsen et al. (2002)), and, thus they become closer to the pricing literature using no-arbitrage and martingale methods. These models are also attractive because, as empirical evidence shows, they are better at capturing stylized facts of financial series, and their predicative performance is superior in comparison to other classes of volatility models (e.g. Koopman et al. (2005)), such as, for example, the class of GARCH models (Engle (1982), Bollerslev (1986)). However, as volatility is treated as a non-observable latent process, the estimation of volatility models is more complicated than the estimation of concurrent models, such as the GARCH class, in which volatility is a deterministic function of the past, which makes the evaluation the likelihood function a simple procedure.

In SV models, the exact evaluation of the likelihood function, due to the presence of the latent volatility factor, requires the calculation of an integral with a dimension equivalent to the sample size. The numeric evaluation of this problem requires methods based on simulation, such as importance sampling methods (e.g. Geweke (1994), Liesenfeld and Richard (2003)) or Markov Chain Monte Carlo (MCMC) (Shephard (1993),Jacquier et al. (1994)). Although these methods are efficient and with the currently available computational power, quite feasible, some problems still remain, such as the determination of a function of importance appropriate or the problem of correlation in the chains in MCMC sampling. It is also possible to work with likelihood function approximations, such as the estimation by quasi-maximum Likelihood (Harvey et al. (1994), Jungbacker and Koopman (2009)), based on a linearization of the SV model. In this methodology, the evaluation of the likelihood functions is made by means of a decomposition of the prediction error using the Kalman filter, which renders a consistent estimator which is asymptotically Gaussian though inefficient and biased in finite samples.

Other ways of evaluating this model employ the estimation by simulation using the methods of indirect inference and the efficient method of moments (Gourieroux et al. (1993), Gallant and Tauchen (1996)). These two methods are asymptotically efficient, and have good properties in finite samples (Monfardini (1998)), but they are less efficient than the MCMC methods of Shephard (1993) and Jacquier et al. (1994). The simplest estimation form for volatility models is the method of moments, the original form of estimation employed in the estimation of the seminal log-normal SV model proposed by Taylor (1986). This methodology was later refined by Melino and Turnbull (1990) through the use of the generalized method of moments (GMM) by Hansen (1982), which generates consistent and asymptotically efficient estimators. These estimators are computationally simple, but their properties in finite samples can be poor and they are inefficient when compared with estimators based on MCMC. A comprehensive study of these estimators' properties can be found in Andersen and Sorensen (1996), and a complete survey about the estimation of SV models using the method of moments can be found in Renault (2009).

[1]For a review of methods for estimating SV models see, for example, Broto and E. (2004), Ghysels et al. (1996), Shepard and Andersen (2009) and Jungbacker and Koopman (2009)

The performance of SV model estimators employing GMM is weakened by the fact that the GMM estimator's bias grows with the number of moment conditions (e.g. Newey and Smith (2004)), and the efficiency in this method depends on an adequate choice of the moment conditions. The GMM estimator manages to reach the efficiency of the maximum likelihood estimator if one of the moments is the score function of the maximum likelihood estimator, or if the moments employed project this function. In practice, the efficient estimation by GMM involves the use of a large number of moment conditions. As the bias in finite samples of the GMM estimator is proportional to the number of moments employed, there is a trade-off between bias and variance in the estimation by GMM when a high number of moment conditions is used. Another problem in the estimation of SV models by GMM is the lack of robustness in the moment conditions employed. The estimation of the log-normal SV model is based on conditions that employ moments of superior orders, and this can be a serious problem in the presence of outliers or processes of heavy-tailed innovation. In this situation, the effects of outliers in the sample are raised to potencies of third or fourth order, which significantly affects the estimation in finite samples.

A further problem lies in the formulation of moment conditions. Although the GMM estimator is semi-parametric, and thus it is not necessary to specify the distribution function of the process, the formulation of moment conditions for SV models generally employs moments derived from the specification of a distribution function for the innovations, as in the case of the so-called log-normal SV model of Taylor (1986). If this assumption is not valid, the properties of the GMM estimator may be degraded.

In this way, the computationally simplest implementation of the generalized method of moments leads to an estimator with poor properties in finite samples (Andersen and Sorensen (1996)), and, on the other hand, the implementation of efficient estimators, such as the methods based on MCMC, are computationally intensive and subject to convergence problems. In this study we propose an alternative form of estimation employing semi-parametric methods of generalized empirical likelihood and generalized minimum contrast. These methods, as will be demonstrated, represent a computationally simpler way of implementation because they can be based on the same moment conditions as the estimators of generalized moment methods, and they produce efficient estimators with good properties in finite samples, as will be demonstrated by a series of Monte Carlo studies. Estimators based on generalized empirical likelihood and generalized minimum contrast derive from a semi-parametric methodology, which permits the estimation of finite dimensional parameters related to the generating process of the parametric part of the process in question - in our case, the parameters of SV process - but accomplishing efficiency (in the semi-parametric sense defined by Bickel et al. (1993)) by means of a non-parametric estimation for the process distribution. This enables us to use the information in the sample in an efficient way. As this methodology uses more information than the estimation by the generalized method of moments - since the latter employs only moments and not the whole information in the sample, it manages to present superior properties in finite samples, comparable or superior to simulation based methods such as MCMC, efficient method of moments, or minimum Hellinger distance (Takada (2009)).

Furthermore, the proposed estimators also address the problem of lack of robustness in the presence of outliers. Two sub-classes of estimators studied (the Exponential Tilting (ET) estimator (Imbens et al. (1998), Kitamura and Stutzer (1997)) and the Exponentially Tilted Empirical Likelihood (ETEL) estimator (Schennach (2007)) have properties of robustness in the presence of incorrect specification problems, and these properties appear to be important in the presence of outliers contaminating the data and in the presence of heavy-tailed distributions in the innovations of the mean and of the process volatility.

This study's analysis methodology is based on Monte Carlo studies for the verification of the properties of the proposed estimators. In order to obtain compatibility in the results obtained, we followed the same designs of Monte Carlo experiments used in the studies by Jacquier et al. (1994), Andersen and Sorensen (1996) and Takada (2009), which facilitates a direct comparison of the results. The Monte Carlo experiments are based on the specifications by Taylor (1986)'s log-normal SV model used in those studies. The simulated series are estimated by methods employing estimators of generalized empirical likelihood, ET and ETEL, as well as the smoothed moments' versions of these models. As reference criterion we will also use the estimation by generalized method of moments, employing two-stage, iterated and continuous updating versions. This benchmarking is useful because the moment conditions are the same.

The objective of these analyses is to verify the properties of the estimators proposed in relation to the size of the sample used, the set of moment conditions, and in relation to the robustness in the presence of heavy-tailed processes of innovation and outliers. To this end, we undertook three classes of experiments. In the first class, we analyzed the effect of sample size and of the set of instruments, analyzing the estimation with sample sizes of 250, 500 and 1,000 observations, using sets of 24 and 14 moment conditions, following Andersen and Sorensen (1996)'s study. In the second class, we verified the estimators' properties in the presence of heavy-tailed innovation processes, and for this we employed two experiment configurations. The first configuration uses a Student t distribution with 4 degrees of freedom as innovation process of the mean; and in the second configuration, we used the same Student t distribution with 4 degrees of freedom, but now as innovation process in the equation that describes the process volatility. The last class of experiments verifies the effect of the outliers on the estimation, and once again, with two kinds of experiments. The first experiment verifies the effects of an outlier on the mean equation (Level Outlier as named by Hotta and Tsay (1998); and the second experiment verifies the effect of an outlier on the volatility equation (Volatility Outlier according to Hotta and Tsay (1998)).

This study is structured as follows: in section 2 we briefly revise the log-normal SV model employed; in section 3, we revise the use of moment conditions in the estimation of SV models; in section 4, we present the estimation methods based on empirical likelihood and generalized minimum contrast; section 5 shows Monte Carlo experiments; and the final conclusions are in section 6.

2. Log-Normal Stochastic Volatility Model

The so-called log-normal volatility model introduced by Taylor (1986) can be described by the following structure:

$$y_t = \sigma_t \varepsilon_t, \tag{1}$$

$$log\sigma_t^2 = \alpha + \beta log\sigma_{t-1}^2 + \sigma u_t, \tag{2}$$

where the equation 1 describes the behavior of the process mean, and equation 2 contains the volatility dynamics. It is usually assumed that the innovation processes in the mean and in volatility are given by independent normal distributions, that is, $(\varepsilon_t, u_t) \sim iidN\,(0, I_2)$ and in this model the parameter vector is given by $\theta = (\alpha,\, \beta,\, \sigma)$. Note that it is possible to interpret this model in a semi-parametric form, as pointed out by Renault (2009), without an a priori specification of the innovation process distributions. Renault (2009) denotes this model as Exponential - SARV because the variance exponential is an autoregressive process.

As demonstrated by Francq and Zakoïan (2006), it is not necessary to assume a distribution for this model's estimation, since, as previously noted by Ruiz (1994), $logy_t^2 = log\sigma_{t-1}^2 + log\varepsilon_t^2$, and this corresponds to an ARMA model (1,1) for the log of the square of the observed process y_t, which makes it possible to derive the representation employed by Francq and Zakoïan (2006) to obtain a consistent estimator by least squares for this model. Francq and Zakoïan (2006) also demonstrate that there is an ARMA(m,m) model for any $logy_t^m$ potency of this process, although it is important to note that the log-normal representation is quite realistic, as indicated by Andersen (1994).

This log-normal specification makes it possible to construct moment conditions of any order, as demonstrated by Taylor (1986) and Melino and Turnbull (1990). The moment conditions of the log-normal SV model can be obtained by initially defining the unconditional mean and variance of the logarithm of the variance:

$$\mu = E\left[\log \sigma_t^2\right] = \frac{\alpha}{1-\beta}, \sigma_y^2 = Var\left[\log \sigma_t^2\right] = \frac{\sigma^2}{1-\beta^2},$$

and the remaining moments as:

$$E\left[|y_t|\right] = (2/\pi)^{1/2}\, E\left[\sigma_t\right],$$

$$E\left[y_t^2\right] = E\left[\sigma_t^2\right],$$

$$E\left[|y_t^3|\right] = 2\sqrt{2/\pi} E\left[\sigma_t^3\right],$$

$$E\left[y_t^4\right] = 3E\left[\sigma_t^4\right],$$

$$E\left[|y_t y_{t-j}|\right] = (2/\pi)\, E\left[\sigma_t \sigma_{t-j}\right],$$

$$E\left[y_t^2 y_{t-j}^2\right] = E\left[\sigma_t^2 \sigma_{t-j}^2\right].$$

Moments of superior order can be written out as:

$$E\left[\sigma_r^r\right] = exp\left(\frac{ru}{2} + \frac{r^2 u^2}{8}\right)$$

for any positive integer j and constants r and s, and in the same way covariances can be obtained by:

$$E\left[\sigma_t^r \sigma_{t-s}^s\right] = E\left[\sigma_t^r\right] E\left[\sigma_t^s\right] exp\left(\frac{rs\beta^j \sigma^2}{4}\right).$$

The moment conditions employed by Andersen and Sorensen (1996) and in our study comprise a set of 24 moment conditions using absolute moments of second to fourth order and lags of first to tenth orders:

(3) $$g^{24}_{t,y_t}(\theta) = \left(|y_t|, y_t^2, |y_t^3|, y_t^4, |y_t y_{t-1}|, ..., |y_t y_{t-10}|, y_t^2 y_{t-1}^2, ..., y_t^2 y_{t-10}^2\right)$$

We also employed a second vector of moment conditions with 14 moment conditions given by:

(4)
$$g^{14}_{t,y_t}(\theta) = \left(|y_t|, y_t^2, |y_t^3|, y_t^4, |y_t y_{t-2}|, |y_t y_{t-4}|, |y_t y_{t-6}|, |y_t y_{t-8}||y_t y_{t-10}|, y_t^2 y_{t-1}^2, y_t^2 y_{t-3}^2, y_t^2 y_{t-5}^2, y_t^2 y_{t-7}^2, y_t^2 y_{t-9}^2\right)$$

With these two vectors of moment conditions we can perform the estimation using the generalized method of moments defined in section 3 and the generalized empirical likelihood and generalized minimum contrast methods in section 4.

3. Estimation of Stochastic Volatility Models using the Method of Moments

The estimation by Hansen (1982)'s generalized method of moments is performed by making the sample moments equal to the population moments , which is equivalent to equalizing the moment conditions vector $g(\theta, Y_t)$ to zero in the form:

(5) $$\overline{g}(\theta, y_t) = \frac{1}{T}\sum_{t=1}^{T} g(\theta, y_t) = 0.$$

This system is generally over-identified (there are more moment conditions than parameters), and so in general there are no solutions. In order to obtain a solution, a criterion function must be employed:

(6) $$J(\theta) = \overline{g}(\theta, y_t)' W \overline{g}(\theta, y_t)$$

and an optimal solution is defined as the minimization of $J(\theta)$, with W being a positive definite weighting matrix. The fundamental result obtained by Hansen (1982) is to demonstrate that the asymptotically efficient solution of the estimation is obtained when this matrix is given by:

(7) $$W^* = \left\{\lim_{t\to\infty} Var\left(\sqrt{T}\overline{w}(\theta)\right)\right\}^{-1} = \Omega(\theta)^{-1}.$$

where $\Omega(\theta)$ denotes the variance-covariance matrix of the model's parameters. In this way, the asymptotically efficient weight is obtained by employing the inverse of the variance-covariance parameter matrix. This matrix is generally unknown, and is usually estimated using the HAC class of estimators by Newey and West (1987):

(8) $$\widehat{\Omega} = \sum_{s=-(T-1)}^{T-1} k_h(s)\widehat{\Gamma}_s(\theta^*),$$

where k denotes a kernel function in relation to a certain parameter of bandwidth h, chosen by means of Newey and West (1987) or Andrews (1991)'s procedures:

$$\widehat{\Gamma}_s(\theta^*) = \frac{1}{T}\sum_{t=1}^{T} g(\theta^*, y_t) g(\theta^*, y_{t+s})'. \tag{9}$$

The efficient estimator of the generalized method of moments is then obtained as a solution to the problem:

$$\widehat{\theta} = \arg\min_{\theta} \overline{g}\left(\theta, y_t\right)' \widehat{\Omega}\left(\theta^*\right) \overline{g}\left(\theta, y_t\right) \tag{10}$$

There are several forms to carry out the implementation of the GMM estimator. The initial form proposed by Hansen (1982) is the estimator known as two-stage GMM. This estimator is obtained by performing a first stage, finding an initial $\widehat{\theta^*} = \arg\min \overline{g}\left(\theta\right)' \Omega \overline{g}\left(\theta\right)$ estimator, where Ω is an initial weighting matrix, usually an identity matrix. Following from this first stage, a HAC matrix $\widehat{\Omega}\left(\theta^*\right)$ is calculated in function of that initial estimation, and the final estimation of the GMM estimator is obtained as $\widehat{\theta} = \arg\min \overline{g}\left(\theta\right)' \widehat{\Omega}\left(\theta^*\right) \overline{g}\left(\theta\right)$ with the HAC matrix that was obtained in the first stage.

A point to be noted is that, in this case, the second stage results depend on the initial estimation in the first stage, and thus this procedure can create a first order bias, weakening the estimator's performance in finite samples (Hansen et al. (1996)). In order to solve this problem, two alternative procedures were proposed. The first procedure is known as iterative GMM, in which the first stage estimation is reinitialized with the result of the second stage estimation, and this iteration continues until the variation in the parameter vector or in the criterion function becomes smaller than an established tolerance.

Another possible estimator is known as GMM with continuous updating (Hansen et al. (1996)). In this case, the estimation of the parameter $\widehat{\theta}$ is not performed in stages, but rather by simultaneously employing a numeric optimization algorithm. Starting from an initial vector θ_0 (usually chosen by employing a two-stage GMM method), the estimation is performed by $\widehat{\theta} = \arg\min \overline{g}\left(\theta\right)' \widehat{\Omega}\left(\theta^*\right) \overline{g}\left(\theta\right)$, but now θ and $\widehat{\Omega}\left(\theta^*\right)$ are simultaneously determined by the numeric optimization procedure. This procedure obtains the same first order properties of the iterative GMM estimator, but, according to Hansen et al. (1996), it has better properties in terms of bias in finite samples, and this estimator is invariant under model reparameterization. According to Newey and Smith (2004) and Anatolyev (2005), the three methods are asymptotically equivalent, but the second order bias in finite samples of the continuous updating estimator is smaller. However, the numeric procedure may be subject to multiple modes in the objective function, which renders this estimator numerically unstable.

The estimation of the SV model by GMM is performed by employing the moment conditions defined by the vector given by Eq. 3. There are, however, some specific points in the estimation of SV. As discussed in Melino and Turnbull (1990) and Hall (2005), the numerical procedure in this problem becomes more difficult due to the presence of non-differentiable moment conditions by using absolute moments. Although these functions are differentiable at almost all the points and the use of absolute moments does not affect the asymptotic properties of the estimators (e.g. Hall (2005)), it is important to discuss how to deal with this problem. Melino and Turnbull (1990) assume that the value of the function is 0 at the non-differentiable points, but this procedure can be problematic because it leads to a discontinuity in the determination of the step size in the numeric optimization algorithm. An alternative form consists in performing a procedure of numerical interpolation at the non-differentiability point, which is the procedure carried out in this study. The properties of this approximation can be seen in Hall (2005).

Properties of the GMM estimator in the estimation of SV models can be found in Andersen and Sorensen (1996)'s study, and a complete revision of the use of methods of moments, including the use of simulated methods of moments, can be found in Renault (2009). The results demonstrate that this estimator, despite being computationally simple, has poor properties in finite samples due to bias and inefficiency problems, although the results are better that those obtained by the quasi-maximum likelihood estimator (e. g. Jacquier et al. (1994)). The problem in finite samples of the GMM estimator is related to the need to use a large number of moments to secure the estimator's efficiency, but the bias of the GMM estimator in finite samples is proportional to the number of moment conditions used. Thus, in finite samples there is a trade-off between bias and efficiency. Note that, although the principal advantage of the GMM estimator lies in its semi-parametric formulation, which does not require assumptions about the sample distribution, the estimator employs only the moments of the process, and it does not employ all the information contained in the sample.

In Andersen and Sorensen (1996)'s article, several details are discussed in the specification of the GMM estimator for SV models, such as the choice of the Kernel function and the bandwidth employed, convergence problems and other subgroups of moment conditions. In this study we employ the quadratic spectral function as kernel function, with the optimum bandwidth chosen by Andrews (1991)'s procedure.

4. Generalized empirical likelihood and generalized minimum contrast Estimators.

The GMM is a method particularly useful in estimating non-linear models when the moments are known. However there is a trade-off between, on the one hand, the weaker need of assumptions for its use, and, on the other, the method's efficiency in finite samples, as discussed in the previous section. The regularity conditions for GMM estimators (Hansen (1982), Newey and McFadden (1994), Hall (2005)) involve only conditions for the asymptotic validity of the moment conditions, and they do not assume stronger conditions such as the knowledge of process distribution, which represents an underutilization of the information presented in the sample.

The opposite situation would be the estimation by the method of maximum likelihood, which uses not only the conditional moments of the process but also all the information present in the conditional densities. If the process is correctly specified and meets the regularity conditions, it is the best asymptotically Gaussian estimator, besides reaching optimality in measures such as Badahur efficiency (Kitamura (2006), DasGupta (2008)). Note that the estimation by maximum likelihood in the context of the estimation of SV models is more complex because the volatility is a latent variable, and the evaluation of the exact likelihood function usually requires simulation methods such as importance sampling or MCMC. Approximations using the quasi-maximum likelihood principle represent a cost in terms of their inferior performance in finite samples.

In this context, an alternative form of formulating estimators that do not need the parametric specification of the process distribution consists in employing semi-parametric estimation methods based on a non-parametric estimation of the likelihood function of the process. These semi-parametric estimators are known as Empirical Likelihood (EL) methods, formulated as generalizations of the non-parametric likelihood methods by Kiefer and Wolfowitz (1956).

According to Kitamura (2006)'s presentation, the non-parametric log-likelihood function of a sequence of IID data $\{y_i\}_{i=1}^n$ of unknown density is defined as:

$$\ell_{NP}(p_1, \ldots, p_n) = \sum_{i=1}^{n} \log p_i, \quad (p_1, \ldots, p_n) \in \triangle, \tag{11}$$

defining $\triangle$ as the simplex $\left\{(p_1, \ldots, p_n) : \sum_{i=1}^n p_i = 1, 0 \leq p_i \leq 1, i = 1, \ldots n\right\}$.

This definition is equivalent to addressing each point of the sample as originating from a multinomial distribution with the support given by the sample $\{y_i\}_{i=1}^n$ observations, even though the y_i density is not multinomial. As this formulation does not involve any model and does not contain the model's parametric structure, it is somehow nonrestrictive when employed in inference problems involving a parametric part with a finite number of parameters. The semi-parametric specification of this process was obtained by Owen (1991), who established the concept of empirical likelihood.

This formulation is important because it allows connections between the non-parametric estimation of the likelihood function and the estimation using moment conditions, formulated with the estimation equation and M-estimators principle - as shown by Qin and Lawless (1994), and these estimation equations can be formulated by using moment conditions in the same way as GMM estimators.

Assuming moment conditions given by:

$$E\left[g(\theta, Y)\right] = \int g(\theta, y) d\mu_Y(y) = 0, \theta \in \Theta \subset \mathbb{R}^k, \tag{12}$$

where μ_Y is the distribution of the random variable Y, the estimation problem using moment conditions can be transformed into a non-parametric likelihood estimation, by the construction of implicit probabilities p_i, and thus the log-likelihood function to be maximized becomes:

$$\ell_{NP}(p_1, \ldots, p_n) = \sum_{i=1}^{n} \log p_i, \quad s.t. \sum_{i=1}^{n} g(\theta, y_i) p_i = 0 \tag{13}$$

The value that maximizes this expression is the maximum empirical likelihood estimative and it maximizes the empirical likelihood function of the process and simultaneously imposes the validity of the moment conditions. These implicit probabilities give more weight to observations where the moment conditions are closer to zero, and less weight to other observations. Note that the generalized method of moments can be obtained as a particular case assuming all weights to be $p_i = 1/n$.

This empirical likelihood formulation is particularly useful in the estimation of models with latent variables where there is no way of evaluating the exact likelihood function of the process. Whereas it is not necessary, when dealing with the GMM estimator, to assume the knowledge of the process likelihood, in the estimators of empirical likelihood the information of the process distribution is used in the estimation by means of its non-parametric estimation. This construction makes it possible to obtain efficiency properties in the semi-parametric sense defined by Bickel et al. (1993).

Note that, when the sample is not an IID process, as time series data present in models of stochastic volatility and denoted by the index of time t, it is necessary to modify the treatment given to the moment conditions. In this situation, the method is modified assuming that the moment conditions originate from a process that is weakly dependent and possibly heterokedastic. Anatolyev (2005) proposes to substitute $g(\theta, y_t)$ for a smoothed version defined as:

$$g^w(\theta, y_t) = \sum_{s=-m}^{m} w(s) g(\theta, y_{t-s}), \tag{14}$$

where $w(s)$ are weights obtained by a kernel function adding one, in the spirit of a HAC estimator (Andrews (1991)). This modification makes it possible to obtain the same conditions of first order asymptotic efficiency present in the GMM methods. The moment conditions are then as follows:

$$\sum_{t=1}^{T} p_t g^w(\theta, y_t) = 0. \tag{15}$$

The GMM estimators is generally defined by the minimization of the quadratic form 10, and in in the overidentified case not all the moment conditions are necessarily equal to zero at the estimated parameter value. In the empirical likelihood estimators formulated by the moment conditions, these conditions are set exactly equal to zero using the ponderation given by the empirical probabilities p_t. Note that in models exactly identified, all the proposed estimators obtain similar results, because in all these estimators the moment conditions are always valid. An important result is that in overidentified models with valid moment conditions all these estimators obtain the same asymptotic variance (e.g. Kitamura (2006)).

It is possible to formulate these empirical likelihood estimators as particular cases of the semi-parametric class of estimators based on the minimization of distances, or, as defined by Bickel et al. (1993), estimators of generalized minimum contrast (GMC)[2]. This formulation makes it possible to obtain the properties of semi-parametric efficiency in this class of estimators. Note that we can also draw a parallel with the interpretation of the GMM estimator as an estimator of minimum χ^2, or the interpretation of quasi-maximum likelihood estimators as estimators of minimum contrast (White (1982)).

In order to show this alternative interpretation of empirical likelihood estimators, we start by defining a general divergence function $D(P, Q)$ between two probability measures P and Q as:

$$D(P, Q) = \int \phi\left(\frac{dP}{dQ}\right) dQ, \tag{16}$$

where ϕ is a convex function. This is an important condition because it allows us to define the conditions of regularity in the process, e.g. Bickel et al. (1993). Define M as the set of all probability measures in $\mathbb{R}^p$ and $\mathcal{P}$, the statistic model defined by measures of probability compatible with 17:

$$\mathcal{P}(\theta) = \left\{ P \in M : \int g(\theta, y) dP = 0 \right\} \tag{17}$$

[2]See Bickel et al. (1993), cap 7, for a general discussion of conditions of regularity, existence and efficiency of generalized minimum contrast estimators.

The estimator of generalized minimum contrast is defined as a solution of:

$$\inf_{\theta\in\Theta}\inf_{\mathcal{P}\in\mathcal{P}(\theta)} D(P,\mu), \tag{18}$$

where μ denote the dominating measure in this model. Thus, in a correctly specified model, this discrepancy must be the unique, and minimum in $\theta = \theta_0$.

In order to establish the connection with empirical likelihood estimators defined by equation 15 and the minimum contrast estimators by means of implicit probabilities, it should be noted that the minimum contrast estimators can be formulated as a problem in the form of moment conditions $E(g(\theta, y_t)) = 0$, turning the minimum contrast estimators into a function of these probabilities, using contrast function h_T:

$$\widehat{\theta}_n = \arg\min_{\theta,p_t}\sum_{t=1}^{T} h_T(p_t). \tag{19}$$

In the case of empirical likelihood estimators, the point estimate $\widehat{\theta}$ is the value which minimizes the discrepancy between $\widehat{p_t}$ and uniform weights. An important result is that an adequate choice of the discrepancy function can lead to a unified representation of empirical likelihood and minimum contrast estimators. This representation can be obtained when the function $h_T(p_t)$ belongs to the Cressie-Read family of discrepancies given by:

$$h_T(p_t) = \frac{[\gamma(\gamma+1)]^{-1}(Tp_t)^{\gamma+1} - 1]}{T} \tag{20}$$

which encompasses cases of several classes of estimators. Empirical likelihood is obtained with the restriction $\gamma \to 0$ in the discrepancy function $h_T(p_t)$; the method of generalized minimum contrast, known as ET of Kitamura and Stutzer (1997) and Imbens et al. (1998), is obtained with $\gamma \to -1$; and the continuous updating estimator using the empirical likelihood formulation is obtained with $\gamma \to 1$.

Note that the problem of estimation involves obtaining estimators not only for the implicit probabilities but also for the parameters of the parametric part of the model, which is, in principle, a high dimension optimization problem. Smith (2001) demonstrated that it is possible to define another estimator that also has these estimators as particular cases, and that makes possible a dual formulation of inferior dimension.

The Smith (2001) Generalized Empirical Likelihood (GEL) estimate is obtained as a solution for the following saddlepoint problem:

$$\widehat{\theta}_n = \arg\min_{\theta}\left[\max_{\lambda}\frac{1}{T}\sum_{t=1}^{T}\rho\left(\lambda' g^w(\theta,y_t)\right)\right], \tag{21}$$

where λ defines Lagrange multipliers imposing a restriction:

$$\sum_{t=1}^{T} p_t g^w(\theta, y_t) = 0. \tag{22}$$

Estimators are obtained by solving the previous equation with the first-order condition:

$$\sum_{t=1}^{T} p_t \lambda' \frac{\partial g^w(\theta, y_t)}{\partial \theta} = 0 \tag{23}$$

with:

$$p_t = \frac{1}{T} \rho' \left(\lambda' g^w(\theta, y_t) \right). \tag{24}$$

This generalized likelihood estimator contains the empirical likelihood estimator, assuming the same conditions of the Cressie-Read divergence function over γ, through modifications of functions h and ρ. The EL estimator is obtained by $h(p) = -ln\,np$ and $\rho(\xi) = ln(1-\xi)$; the ET estimator by (Kitamura and Stutzer (1997) , Imbens et al. (1998)) with $h(p) = np\,ln\,np$ and $\rho(\xi) = -exp(\xi)$; and the continuous updating estimator as $h(p) = (np)^2$ and $\rho(\xi) = -(1+\xi)^2/2$[3].

An additional class of estimators which do not belong directly to the class of EL or minimum contrast estimators, but which is obtained by combining the empirical likelihood estimator and the ET estimator, is the ETEL estimator proposed by Schennach (2007). This estimator is defined as:

$$\widehat{\theta} = \arg\min_{\theta} \left(n^{-1} \sum_{i=1}^{n} \widetilde{h}(p_t(\theta)) \right), \tag{25}$$

where $\widetilde{g}_i(\theta)$ is the solution of:

$$\min_{\{g_i\}_{i=1}^{n}} n^{-1} \sum_{i=1}^{n} h(p_t) \tag{26}$$

subject to $\sum_{i=1}^{n} p_t g(\theta, y_t) = 0$ and $\sum_{i=1}^{n} p_t = 1$, with $\widetilde{h}(\widehat{p_t}) = -ln(np_t)$ and $h(p_t) = np_t ln(np_t)$.

Note that the ETEL estimator employs the ET method to find the probabilities $\widehat{p_i}(\theta)$, and the EL method to estimate the parameter vector $\widehat{\theta}$. These probabilities are related to the multipliers λ by the relation:

$$\widehat{p_t}(\theta) = \frac{\left(\widehat{\lambda}(\theta)' g(\theta, y_t)\right)}{\sum_{i=1}^{n} \left(\widehat{\lambda}(\theta)' g(\theta, y_t)\right)}. \tag{27}$$

An important property of the estimators of ETEL class is their behavior in the presence of incorrect specification. Imbens et al. (1998) point out that the EL estimator can display inadequate behavior in the presence of incorrect specification due to the presence of a singularity in its influence function, and, according to theorem 1 in Smith (2001), the asymptotic properties of the EL estimator can be severely weakened in the presence of minimum specification problems. This also affects the estimations of the implicit probabilities, because, in the presence of specification problems, the implicit probabilities in likelihood problems tend to concentrate on the extreme observations, in opposition to what is expected in a robust estimator in Huber

[3]See Table 1 Smith (2001) for more details.

(1981) and Hampel et al. (1986)'s sense, which should minimize the importance of extreme observations in the construction of an estimator.

We will now summarize some common properties of the estimators discussed in this study. The first property is that all the estimators employed (two-stage GMM, iterative GMM, GMM continuous updating, GEL, ET, and ETEL) have the same properties of consistency and first order asymptotic efficiency (e.g. Smith (2001), Schennach (2007)), and in the validity of moment conditions all the estimators have the same asymptotic variance. However, their performance in finite samples can be quite different. The two-stage GMM estimator can be severely biased in sample sizes employed in economics and finance, and continuous updating estimators are numerically unstable due to the presence of multiple modes in the objective function (e.g. Hansen et al. (1996)). Another interesting property is that estimators based on GMC and GEL are invariant to linear transformations in the vector of moment conditions, which does not occur in the two-stage GMM estimator. Estimators based on generalized empirical likelihood/minimum contrast are efficient in the semi-parametric sense of Bickel et al. (1993), and have superior properties in terms of higher order asymptotic bias. These estimators also present optimum properties in terms of hypotheses testing. As demonstrated by Kitamura (2006), these tests are optimum in the minimax and large deviations criteria, and are uniformly more powerful in the generalized sense of Neyman-Pearson.

A fundamental point is that in the EL and minimum contrast estimators based on the Cressie-Read discrepancy, the bias in finite samples does not grow with the number of moment conditions used. This property makes it possible for the efficiency of the estimators to be obtained with the use of a high number of moment conditions, without implying an increase in the bias in the finite samples as occurs in the use of the GMM estimator, which leads to the problem of the inferior performance of this method in comparison with other forms of estimation.

The result obtained by Smith (2001) is that in the class of minimum contrast/empirical likelihood estimators, the only estimator with adequate behavior in the presence of specification problems is the ET estimator, because its influence function does not present singularities. The ETEL estimator is a combination of the EL estimator and the EL estimator, and it maintains the EL estimator's characteristics of asymptotic efficiency and minimum bias. Additionally, it inherits the robustness in the presence of specification problems, due to the use of the ET estimator to estimate implicit probabilities, as demonstrated by theorems 8-10 in Smith (2001), who proves that this estimator is $\sqrt{n}$ convergent even in the presence of specification problems.

Estimators for the parameters of the parametric part of the model and for the implicit probabilities can be obtained by numeric optimization or via quasi-Newton iterative methods. These methods can be formulated in a problem of smaller dimension using a dual formulation (Kitamura (2006)) through the numeric optimization employing Lagrange multipliers defined by equations 21 and 27, which is the general form used in this study.

Note that in the estimation of SV models we are subject to the same problem of using non-differentiable moment conditions due to the use of absolute moments. This problem impedes the simple use of iterative methods for the estimation of Lagrange multipliers proposed by Kitamura (2006), and thus, in these cases, we need to use the same techniques of numeric optimization with the interpolation in the vicinity of the discontinuity points discussed in the estimation by GMM.

5. Monte Carlo Studies

The performance of the proposed estimators is analyzed through a series of Monte Carlo studies, with the purpose of verifying the performance of each estimator in different parameter configurations, sample sizes, moment conditions employed, and robustness in the presence of specification problems and outliers. In order to analyze these problems, we worked with three parameter configurations for each experiment performed. These configurations follow the same configurations employed in the articles by Jacquier et al. (1994), Andersen and Sorensen (1996) and Takada (2009). The set of simulated models correspond to the parameters (α, β, σ) given by (-0.736, .9, .3629), (-0.368, .95, .26) and (-.1472, .98, .1657). This choice is justified in the study by Jacquier et al. (1994) where these configurations are considered as generating the same unconditional variance but with distinct persistence configurations.

In the first analysis, we performed the estimation of the reference models (Gaussian innovations in the mean and volatility equations and without outlier) using the estimators defined previously. For each parameter vector we carried out 1,000 replications. The sample size will be equal to 500 in all cases except in the analysis of the sample size effects. Each simulated series was estimated by the following methods: two-stage GMM (GMM2S), Iterative GMM (GMMITER), GMM Continuous Updating (GMMCUE), GEL, ET and ETEL, as well as the versions with smoothed moments of these three last estimators (SGEL, SET and SETEL).

Tables 1, 2 and 3 show the estimation results of these reference models with three parameter configuration; each table presenting the mean, the bias, mean squared error (MSE), and mean absolute error (MAE) of each parameter estimator. In order to ease the visualization of the results, we shown in Figure 1 the MSE and MAE of each estimator for each parameter. In terms of mean quadratic error and mean absolute error generally the estimators based on EL and GMC are much superior to those obtained by estimators based on GMM, and this superiority is valid for all the three parameters estimated in all parameter configurations. This result gives support to the use of these methods as competitive methodologies in the estimation of SV models.

Although the straight comparison in this article is performed with estimators using the same moment conditions, due to the use of the same parameter configuration of other studies, it is possible to compare the results obtained with other estimation methodologies. The results obtained are directly comparable with those analyzed in Takada (2009)'s article, who proposed an estimator for SV models employing simulated Minimum Hellinger Distances, comparing this method with other methodologies, such as the efficient method of moments (EMM), MCMC, and maximum likelihood Monte Carlo.

Table 1 in Takada (2009) shows the results for these estimators' MSE for the first parameter vector studied, for a sample of size 500. The results of a direct comparison with the results presented in this table indicate that the estimators based on GEL/GMC are superior to the following methods in terms of MSE: SMHD (Simulated Minimum Hellinger Distance), EMM (Efficient Method of Moments) and MCMC. They also have a superior or equivalent performance to the MCML (Monte Carlo Maximum Likelihood) estimators by the criterion of mean quadratic error. In comparison with the results of that article, we notice that the results of all the estimators based on GEL/GMC are superior to all these methods, except for the estimation of α where the estimators obtain a mean quadratic error equal to the MCML estimator.

In this comparison it is important to notice that the GEL/GMC estimators do not require Monte Carlo simulation procedure, and are computationally simpler than these methods, indicating that the use of EL and MC makes it possible to obtain superior properties in finite samples when compared with the methods so far considered as the most efficient in the SV model estimation, with a noticeably smaller computational and implementation cost.

5.1. **Effect of Sample Size and Set of Instruments.** In order to verify the effect of the sample size in the estimators' performance, we carried out an analysis with the estimation of the parameter vectors studied with samples of size 250 (Tables 4, 5 and 6) and 1,000 (Tables 7, 8 and 9) and employing the 24 moment conditions defined by equation 3. As expected, the increase in the sample size decreases the MSE and MAE of all the estimators, but with different effects for each parameter configuration of and estimation method. Summarizing these results, we show in Figure 2 the relative efficiency, defined as a ratio between the MSE of the sample of size 250 and the MSE of sample size 1,000 for each configuration.

Except for the GEL estimator in parameter configuration 2, with efficience rate inferior to one, there is a real gain in terms of MSE for all the parameters. This particular result for the GEL estimator in this configuration can be explained by the greater convergence difficulty noted in this particular configuration, but it is important to note that, in the version with smoothed moments, this estimator behaves as expected.

As can be seen in Figure 2, the sample size has heterogeneous effects for each estimator, depending on the parameter configuration. The estimators based on GEL/GMC with smoothed moments have greater gain in the configuration with smaller persistence while those based on GMM behave in the opposite way. This result can be interpreted by the fact that the smoothing of moments is more efficient when the volatility persistence is smaller.

As previously discussed, the main theoretical motivation for the use of estimators based on GEL/GMC lies in the possibility of using a larger number of moment conditions to achieve a more efficient estimation, since the finite samples bias in these methods does not grow with the number of moment conditions, as occurs with GMM estimators. In order to verify this property, we employ a new estimation with a subset of the moment conditions vector, now working with 14 moment conditions only, according to Eq. 4, instead of the original 24 moment conditions given by Eq. 3.

The results of this comparison are displayed in Tables 10, 11 and 12, and the comparisons between estimators employing MSE and MAE with the use of 14 moment conditions are placed in Figure 3. We can note that in this configuration the GEL/GMC estimators still display a superior performance in comparison with those based on GMM, but now this performance is not as superior as in the configuration with 24 moment conditions, which gives support to the conjecture of a superior use of the moment conditions in terms of bias and variance for the estimators of GEL/GMC class.

Figure 4 presents the relative efficiency between MSE using 14 moments and the estimator with 24 moments. For the GMM estimators the efficiency presents modest increases or reductions increasing the number of instruments, similarly to the results obtained in the studies by Andersen and Sorensen (1996). However there are, in general, very significant efficiency gains in MSE for the estimators based on GEL/GMC, reaching values over 200 times in the second parameter configuration. Nevertheless, for the third parameter configuration, we can observe that the estimation with a number of moment conditions represents a reduction in the relative efficiency of all the methods for the estimators of α and β.

5.2. **Student-t Distribution (4) in the mean innovations.** As previously discussed, although the SV log-normal model is defined by moments of a log-normal distribution, it can be interpreted in a semi-parametric form as an autoregressive model for the exponential of the volatility process, without the need for a distribution specification for the innovation processes (e.g. Francq and Zakoïan (2006), Renault (2009)). However, as we are employing these theoretical moments assuming the distribution specification of innovations in the construction of the moment conditions, it is important to verify whether alternative specifications significantly alter the properties of the estimators in finite samples. It is particularly interesting to verify if, consistently with what is observed for financial series, heavy-tailed processes affect these estimators.

The first analysis undertaken was to replace the standard Gaussian distribution in the innovations of the mean process for a Student-t distribution with 4 degrees of freedom. This choice was purposely made with the aim of verifying the effect of a distribution with heavier tails on the estimation of SV models. Note that, as we are employing higher moments, the heavy-tail effect can be magnified in the estimation, since now each observation is raised to potencies of second, third and fourth orders. We particularly use this number of 4 degrees of freedom in Student-t to have a distribution with non-finite kurtosis and, consequently, to have a robustness test under extreme conditions.

Tables 13, 14 and 15 show the results of this experiment using 24 moment conditions and Tables 16, 17 and 18, using 14 moment conditions. It can be seen that in this situation the estimators based on GMC/GEL clearly maintain their dominance over the estimators based on GMM, as it becomes more evident in Figures 5 and 6, which show MSE and MAE of each estimator, and once again we have the same result of best performance in this situation of the GEL/GMC-based estimators.

In order to verify whether in this case it is still advantageous to work with a larger set of instruments Figure 7 shows the ratio of the MSEs between the estimators with 14 and 24 moment conditions. The results show that in this situation the increase in the number of instruments can impair the performance of the estimators, and this effect occurs both for the GMM estimators and for the GEL/GMC estimators, although the effect is heterogeneous in terms of the configuration and of the parameter analyzed. In the situation of lower persistence, it is advantageous to work with the larger number of instruments for the GEL/GMC estimators, but this result is not maintained in the other parameter configurations, and particularly in the configuration with high persistence, the use of the larger set of instruments causes almost a general degradation in the performance of all the methods.

5.3. **Student-t Distribution (4) in the volatility innovations.** In the next experiment, we modified the data generating process, assuming now that the innovation process in the volatility equation is given by a Student-t process with 4 degrees of freedom, assuming in this case the usual supposition of Gaussian innovations in the mean equation. Note that, in this configuration, the effects are expected to be more harmful, since now the effect of heavier tails is directly spread by the volatility equation's autoregressive structure, unlike the previous case where the heavy-tailed innovations affected the mean equation, which was a process without correlation.

Tables 19, 20 and 21 show the results obtained with 24 moment conditions, and Tables 22, 23 and 24 show the results obtained with 14 moment conditions. These results are summarized in Figures 8 and 9. We note that these heavier tailed innovations effectively damage the performance of the GMM-based estimators, and moderately damage the GEL-based estimators. In this experiment, the robustness properties of the methods based on ET and ETEL become evident, and these methods generally have a superior performance in comparison with the other methods. For example, the ratio between MSE for α estimated by Iterative GMM and by the smoothed ETEL method has a value of 5102.984 for the first parameter configuration, showing the dominance of these methods in this situation of incorrect specification. As previously discussed, this robustness property is derived from the bounded influence function of the estimators based on ET, and it proves to be quite important in this situation. As financial data is characterized by heavy tails, we have an additional justification for the use of the estimators proposed in this study.

Likewise, we can verify the effects of using a number of moment conditions in this configuration. Figure 10 shows the relative efficiency effects of the estimators obtained with the increase in the number of instruments from 14 to 24. However, in this configuration, we have mixed results because for the first parameter configuration there is a general gain in the estimators - though more noticeable for the estimators based on GEL/GMC -, but for the other configurations there are losses, particularly in the estimation of the volatility parameter σ in the second configuration.

5.4. **Experiment 4 - Level Outlier.** In order to verify the effects of aberrant observations (outliers) in the process of stochastic estimation, we undertook two classes of experiments. In this part of our study we will verify the effect of the so-called level outliers (in Hotta and Tsay (1998)'s nomenclature) in the estimation of SV models. In this experiment the generating process is given by:

$$y_t = \sigma_t \varepsilon_t + LO_t \tag{28}$$

$$log\sigma_t^2 = \alpha + \beta log\sigma_{t-1}^2 + \sigma u_t, \tag{29}$$

where LO_t is a binary variable with positive value of 5 standard deviations of the process if the observation is carried out in the period t=251, and zero in the other observations. Note that in this experiment the outlier do not affect the persistence in the volatility process. The results of this experiment are displayed in Tables 25, 26 and 27 for the set of 24 moments; and in Tables 28, 29 and 30 for the set of 14 moments; and the results of MSE and MAE are summarized in Figures 11 and 12. We observe a better performance of the estimators based on GEL/GMC, particularly those employing the ET method for the calculation of the Lagrange multipliers. For example, the ratio of 260.8 between the MSE of the GMM Iterative estimator and the smoothed ET estimator of α in the parameter configuration 3 supports the evidence that the robustness properties of this class of estimators have advantages in the estimation of SV models. The performance of these estimators is more noticeable in the situation of longer volatility persistence, given by parameter vector 3.

It is not possible, however, to identify a clear effect of the number of moment conditions in this experiment, since the effects are similar to those occurred in the previous experiments with heavy-tailed innovations. As per Figure 13, the relative efficiency between 14 and 24 moment conditions, for parameter vectors 2 and 3 indicates that the increased number of instruments represent a loss in performance in most cases, particularly for the estimation of parameter σ.

5.5. **Experiment 5 - Volatility Outlier.** In the last specification tested, we verified the effect of a so-called volatility outlier (as named by Hotta and Tsay (1998)) in the estimation. In this experiment, the data generating process is given by:

$$y_t = \sigma_t \varepsilon_t \tag{30}$$

$$log\sigma_t^2 = \alpha + \beta log\sigma_{t-1}^2 + \sigma u_t + VO_t, \tag{31}$$

where VO_t is a binary variable with positive value of 5 standard deviations in observation 251 of the volatility equation and zero in the other observations. In this situation, there is a direct propagation of the effects of the outlier in the volatility because now the effect is directly transmitted by the autoregressive structure in the volatility equation, whereas the effect was indirect in the case of a level outlier.

Tables 31, 32 and 33 (estimation with 24 moments) and 34, 35 and 36 (estimation with 14 moments) show the results of estimations which can be summarized by Figures 14 and 15 with the MSE and MAE results. As per previous experiments, the GEL/GMC-based estimators have in general a superior performance in comparison with the GMM-based methods, and show that the same properties of robustness remain valid in this volatility outlier situation, which would be potentially more serious for the estimation of volatility parameters.

The effect of the larger number of instruments in this situation can be seen in Figure 16, which indicates that there is an efficiency gain with a higher number of instruments in the situation with low persistence; however, for situations with higher volatility persistence, the additional instruments generally present noticeable deterioration in the estimators' MSE.

6. Conclusions

In this study we discussed the estimation of SV models using estimators based on generalizations of the empirical likelihood and minimum contrast methods. The performance of these estimators, as shown by a set of Monte Carlo experiments, proved to be superior to the estimation methods based on generalized method of moments, and also superior to the methods based on simulation such as MCMC and Monte Carlo maximum likelihood as studied in Takada (2009).

The results obtained in this study are consistent with those obtained by other studies (e.g. Newey and Smith (2004)), which demonstrate that alternative estimators based on moments, formulated as GEL/GMC-based estimators, display superior performance, nullifying the bias problems occurring in the usual GMM estimators. The proposed estimators manage to obtain superior properties in finite samples by a better use of the informational content present in the moment conditions, since the higher efficiency is obtained not only by means of weighting by the estimators' variance - as in the case of GMM estimators - but also by the non-parametric estimation of the likelihood function of the process, as discussed in Antoine et al. (2007). Another related property lies in the fact that the bias of these estimators does not grow with the number of moment conditions, as happens in the case of GMM estimators. Thus, it is possible to obtain efficiency properties by using an adequate number of moment conditions. This characteristic can be particularly important in the estimation of multivariate SV models, in which the number of moment conditions is proportional to the number of series studied. As the estimation of multivariate SV models still represents a great computational challenge, (e.g. Chib et al. (2009)), the estimation by methods based on empirical likelihood/minimum contrast can be an efficient alternative to be explored.

These results are particularly interesting because the implementation of the methods discussed in this study is computationally simpler than the implementation of methods based on simulation, requiring just one specification of the moment conditions of stochastic volatility processes. Although this study is based on the specification of the log-normal SV model, it is important to note that this procedure can be generalized by using the methodology proposed by Meddahi (2001), which makes possible the automatic generation of moment conditions for processes that belong to the so-called SV-eigenfunctions family.

Another important characteristic is related to robustness properties and specification problems, particularly of the methods based on ET, which, due to properties in their influence function, manage to be $\sqrt{n}$ consistent even in the presence of specification problems. This property is particularly important in the presence of processes of heavy-tail innovations, as verified in this study by the use of a Student-t distribution with non-finite kurtosis, or else in the presence of level or volatility outliers.

References

Anatolyev, S.: 2005, Gmm, gel, serial correlation and asymptotic bias, *Econometrica* **73**, 983–1002.

Andersen, T.: 1994, Stochastic autoregressive volatility: A framework for volatility modelling, *Mathematical Finance* **4**, 75–102.

Andersen, T. and Sorensen, B.: 1996, Gmm estimation of a stochastic volatility model: A Monte Carlo study, *Journal of Business and Economic Statistics* **14(3)**, 328–352.

Andrews, D. W. K.: 1991, Heteroskedasticity and autocorrelation consistent covariance matrix estimation, *Econometrica* **59**, 817–858.

Antoine, B., Bonnal, H. and Renault, E.: 2007, On the efficiente use of the informational content of estimating equations: implied probabilities and euclidian empirical likelihood, *Journal Of Econometrics* **138**, 461–487.

Barndorff-Nielsen, O. E., Noclato, E. and Shephard, N. G.: 2002, Some recent developments in stochastic volatility modelling, *Quantitative Finance* **2**, 11–23.

Bickel, P., Klassen, C., Ritov, Y. and Wellner, J.: 1993, *Efficient and Adaptative Estimation for Semiparametric Models*, Johns Hopkins Press.

Bollerslev, T.: 1986, Generalized autoregressive conditional heteroskedasticity, *Journal of Econometrics* **32**, 307–327.

Broto, C. and E., R.: 2004, Estimation methods for stochastic volatility methods: A survey, *Journal of Economic Surveys* **18(5)**, 613–649.

Chib, S., Omori, Y. and Asai, M.: 2009, *Handbook of Financial Time Series*, Springer, chapter Multivariate Stochastic Volatility Models, pp. 365–402.

DasGupta, A.: 2008, *Asymptotic Theory of Statistics and Probability*, Sprimger.

Engle, R. F.: 1982, Autoregressive conditional heteroscedasticity with estimates of the variance of United Kingdom inflation, *Econometrica* **50**, 987–1007.

Francq, C. and Zakoïan, J. M.: 2006, Linear-representation bases estimation of stochastic volatilty models, *Scandinavian Journal of Statistics* **33**, 785–806.

Gallant, R. A. and Tauchen, G.: 1996, Which moments to match, *Econometric Theory* **12(4)**, 657–681.

Geweke, J.: 1994, Bayesian comparison of econometric models. Federal Reserve of Minneapolis Working Paper,.

Ghysels, E., Harvey, A. C. and Renault, E.: 1996, *Statistical Methods in Finance*, North Holland, chapter Stochastic Volatility, pp. 221–238.

Gourieroux, C. A., Monfort, A. and Renault, E.: 1993, Indirect inference, *Journal of Applied Econometrics* **8**, 85–118.

Hall, A.: 2005, *Generalized Method of Moments*, Oxford University Press.

Hampel, F. R., Ronchetti, E. M., Rousseeuw, P. and Stahel, W. A.: 1986, *Robust Statistics: The Approach Based on Influence Functions*, John Wiley & Sons.

Hansen, L. P.: 1982, Large sample properties of Generalized Method of Moments estimators, *Econometrica* **50**, 1029–1054.

Hansen, L. P., Heaton, J. and Yaron, A.: 1996, Finite sample properties od some alternative gmm estimators, *Journal of Business and Economic Statistics* **14**, 262–280.

Harvey, A. C., E., R. and Shephard, N. G.: 1994, Multivariate stochastic variance models, *Review of Economic Studies* **61**, 247–264.

Hotta, L. and Tsay, R.: 1998, Outliers in garch processes. Working Paper, Graduate School of Business, University of Chicago.

Huber, P. J.: 1981, *Robust Statistics*, John Wiley & Sons.

Imbens, G. W., Spady, R. H. and Johnson, P.: 1998, Information theoretic approaches to inference in moment conditions models, *Econometrica* **66**, 333–357.

Jacquier, E., Polson, N. and Rossi, P. E.: 1994, Bayesian analysis of stochastic volatilty models (with discussions)., *Journal of Business and Economic Statistics* **12**, 371–314.

Jungbacker, B. and Koopman, S. J.: 2009, *Handbook of Financial Time Series*, Springer, chapter Parameter Estimation and Practical Aspects of Modeling Stochastic Volatility, pp. 313–344.

Kiefer, J. and Wolfowitz, J.: 1956, Consistency of maximum likelihood estimator in the presence of infinitely many parameters., *Annals of Mathematical Statistics* **27**, 887–906.

Kitamura, Y.: 2006, Empirical likelihood methods in econometrics: Theory and practice. Unpublished Working Paper.

Kitamura, Y. and Stutzer, M.: 1997, An information-theoretic alternative to generalized method of moments estimation, *Econometrica* **65(5)**, 861–874.

Koopman, S. J., Jungbacker, B. and Hol, E.: 2005, Forecasting daily variability of the s&p 100 stock index using historical, realised and implied volatility measurements., *Journal of Empirical Finance* **12**, 445–475.

Liesenfeld, R. and Richard, J.: 2003, Univariate and multivariate stochastic volatility models: Estimation and diagnostics, *Journal of Empirical Finance* **10**, 505–531.

McNeil, A., Frey, R. and Embrechts, P.: 2005, *Quantitative Risk Management*, Princeton University Press.

Meddahi, N.: 2001, An eigenfunction approach for volatility modelling. Working Paper, University of Montreal.

Melino, A. and Turnbull, S. M.: 1990, Pricing foreign currency options with stochastic volatility, *Journal Of Econometrics* **45**, 239–265.

Monfardini, C.: 1998, Estimating stochastic volatility models through indirect inferene, *Econometrics Journal* **1**, 113–128.

Newey, W. K. and West, K. D.: 1987, A simple, positive semi-definite, heteroskedasticity and autocorrelation consistent covariance matrix, *Econometrica* **55**, 703–708.

Newey, W. and McFadden, D.: 1994, *Handbook of Econometrics*, Vol. 4, Elsevier, chapter Large sample estimation and hypothesis testing.

Newey, W. and Smith, R. J.: 2004, High-order properties of gmm and generalized empirical likelihood estimators, *Econometrica* **72**, 219–255.

Owen, A.: 1991, Empirical likelihood for linear models, *The Annals of Statistics* **18(1)**, 1725–1747.

Qin, J. and Lawless, J.: 1994, Empirical likelihood and general estimating equations, *The Annals of Statistics* **20(1)**, 300–325.

Renault, E.: 2009, *Handbook of Financial Time Series*, Springer, chapter Moment-Based Estimation of Stochastic Volatility Models, pp. 269–311.

Ruiz, E.: 1994, Quasi-maximum likelihood estimation of stochastic volatility models, *Journal Of Econometrics* **63**, 284–306.

Schennach, S.: 2007, Point estimation with exponentially tilted empirical likelihood, *Annals of Statistics* **35(2)**, 634–672.

Shepard, N. and Andersen, T.: 2009, *Handbook of Financial Time Series*, Springer, chapter Stochastic Volatility: Origins and Overview, pp. 233–254.

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

Singleton, K. J.: 2006, *Empirical Dynamic Asset Pricing*, Princeton University Press.

Smith, R. J.: 2001, Gel criteria for moment conditions models. Working Paper, University of Bristol.

Takada, T.: 2009, Simulated minimum hellinger distance estimation of stochastic volatility models, *Computational Statistics and Data Analysis* **53**, 2390–2403.

Taylor, S. J.: 1986, *Modelling Financial Time Series*, John Wiley& Sons.

White, H.: 1982, Maximum likelihood estimation of misspecified models, *Econometrica* **50**, 1–25.

Appendix - Tables

	GMM2S	GMMITER	GMMCUE	GEL	ET	ETEL	SGEL	SET	SETEL
mean α	-0.700191	-0.608924	-0.569311	-0.729763	-0.730008	-0.728858	-0.732023	-0.730267	-0.727510
bias α	0.035809	0.127076	0.166689	0.006237	0.005992	0.007142	0.003977	0.005733	0.008490
mse α	0.495912	0.827004	0.383436	0.000723	0.000105	0.000229	0.002954	0.000148	0.000431
mae α	0.494248	0.686760	0.466081	0.016242	0.007843	0.011289	0.014323	0.008779	0.011775
mean β	0.905908	0.918369	0.923634	0.902438	0.901313	0.901860	0.901165	0.901239	0.901681
bias β	0.005908	0.018369	0.023634	0.002438	0.001313	0.001860	0.001165	0.001239	0.001681
mse β	0.008939	0.014804	0.006909	0.000019	0.000005	0.000009	0.000049	0.000007	0.000013
mae β	0.066857	0.092551	0.063312	0.003243	0.001806	0.002431	0.002724	0.001962	0.002344
mean σ	0.236795	0.158262	0.170565	0.386721	0.387840	0.380383	0.378491	0.383913	0.376488
bias σ	-0.126105	-0.204638	-0.192335	0.023821	0.024940	0.017483	0.015591	0.021013	0.013588
mse σ	0.039156	0.067792	0.053258	0.001503	0.001488	0.001911	0.002134	0.001700	0.002037
mae σ	0.168891	0.234061	0.207049	0.031779	0.033321	0.035319	0.037523	0.035307	0.036788

Table 1. Reference SV Model Sample Size 500 - α=-0.736 β=.9 σ=.3629, T=500

	GMM2S	GMMITER	GMMCUE	GEL	ET	ETEL	SGEL	SET	SETEL
mean α	-0.323798	-0.300172	-0.520061	-0.371048	-0.366947	-0.369468	-0.367388	-0.366799	-0.367850
bias α	0.044202	0.067828	-0.152061	-0.003048	0.001053	-0.001468	0.000612	0.001201	0.000150
mse α	0.209281	0.375829	0.034725	0.000468	0.000219	0.000250	0.000331	0.000225	0.000631
mae α	0.311453	0.394674	0.156372	0.012824	0.010469	0.011898	0.012324	0.011163	0.013191
mean β	0.956477	0.959678	0.930090	0.950147	0.950309	0.950104	0.950271	0.950331	0.950167
bias β	0.006477	0.009678	-0.019910	0.000147	0.000309	0.000104	0.000271	0.000331	0.000167
mse β	0.003805	0.006868	0.000635	0.000008	0.000004	0.000005	0.000006	0.000004	0.000010
mae β	0.042185	0.053324	0.020561	0.001767	0.001443	0.001708	0.001526	0.001395	0.001681
mean σ	0.146227	0.098469	0.198858	0.265285	0.272920	0.262856	0.258772	0.269163	0.257539
bias σ	-0.113773	-0.161531	-0.061142	0.005285	0.012920	0.002856	-0.001228	0.009163	-0.002461
mse σ	0.027740	0.040511	0.004846	0.001597	0.001578	0.001870	0.002119	0.001679	0.002234
mae σ	0.142769	0.182002	0.061584	0.031468	0.033283	0.034507	0.037988	0.034831	0.039488

Table 2. Reference SV Model Sample Size 500 - α=-0.368 β=.95 σ=.26

	GMM2S	GMMITER	GMMCUE	GEL	ET	ETEL	SGEL	SET	SETEL
mean α	-0.148610	-0.153037	-0.161124	-0.177110	-0.167514	-0.166558	-0.182551	-0.168109	-0.170785
bias α	-0.001410	-0.005837	-0.013924	-0.029910	-0.020314	-0.019358	-0.035351	-0.020909	-0.023585
mse α	0.122958	0.180502	0.007471	0.002719	0.000792	0.000865	0.004335	0.001014	0.002519
mae α	0.185981	0.222997	0.033918	0.031699	0.020929	0.020838	0.038433	0.021868	0.025349
mean β	0.980010	0.979486	0.979135	0.976228	0.977244	0.977427	0.975315	0.977149	0.976785
bias β	0.000010	-0.000514	-0.000865	-0.003772	-0.002756	-0.002573	-0.004685	-0.002851	-0.003215
mse β	0.002228	0.003236	0.000131	0.000043	0.000017	0.000017	0.000073	0.000023	0.000048
mae β	0.025133	0.030039	0.003657	0.004043	0.002837	0.002788	0.005116	0.002971	0.003441
mean σ	0.078125	0.055286	0.148443	0.169463	0.170654	0.160767	0.167593	0.169098	0.160081
bias σ	-0.087575	-0.110414	-0.017257	0.003763	0.004954	-0.004933	0.001893	0.003398	-0.005619
mse σ	0.016178	0.020346	0.001544	0.001591	0.001388	0.001523	0.001909	0.001439	0.001613
mae σ	0.115562	0.133223	0.017375	0.031635	0.029684	0.030572	0.035919	0.029612	0.032264

Table 3. Reference SV Model Sample Size 500 - α-.1472 β=.98 σ=.1657

Reference SV Model - Sample Size 500.

	GMM2S	GMMITER	GMMCUE	GEL	ET	ETEL	SGEL	SET	SETEL
mean α	-0.845827	-0.840909	-0.672729	-0.736400	-0.728755	-0.727990	-0.730019	-0.729343	-0.726712
bias α	-0.109827	-0.104909	0.063271	-0.000400	0.007245	0.008010	0.005981	0.006657	0.009288
mse α	0.944574	2.522021	1.339914	0.004713	0.000205	0.000555	0.000448	0.000775	0.001030
mae α	0.568116	0.956395	0.602316	0.023396	0.010023	0.015524	0.013468	0.012361	0.016576
mean β	0.887361	0.889161	0.910046	0.901653	0.901630	0.902274	0.901814	0.901582	0.902088
bias β	-0.012639	-0.010839	0.010046	0.001653	0.001630	0.002274	0.001814	0.001582	0.002088
mse β	0.016784	0.043542	0.023987	0.000080	0.000009	0.000016	0.000015	0.000020	0.000028
mae β	0.076241	0.127159	0.081218	0.004287	0.002298	0.003158	0.002857	0.002684	0.003380
mean σ	0.255995	0.139735	0.143907	0.383615	0.385449	0.377069	0.377577	0.383512	0.370742
bias σ	-0.106905	-0.223165	-0.218993	0.020715	0.022549	0.014169	0.014677	0.020612	0.007842
mse σ	0.042192	0.079458	0.067077	0.002284	0.001830	0.002593	0.002906	0.002043	0.003781
mae σ	0.170713	0.256598	0.237429	0.037310	0.036526	0.040098	0.041506	0.038537	0.046534

TABLE 4. Reference SV Model Sample Size 250 - α=-0.736 β=.9 σ=.3629

	GMM2S	GMMITER	GMMCUE	GEL	ET	ETEL	SGEL	SET	SETEL
mean α	-0.453177	-0.510670	-0.582461	-0.369508	-0.365449	-0.367953	-0.367438	-0.365712	-0.366300
bias α	-0.085177	-0.142670	-0.214461	-0.001508	0.002551	0.000047	0.000562	0.002288	0.001700
mse α	0.472332	1.448096	0.080215	0.000396	0.000215	0.000345	0.001415	0.000261	0.001354
mae α	0.392820	0.621217	0.219123	0.013576	0.010837	0.012862	0.016954	0.011440	0.016981
mean β	0.939239	0.932243	0.921192	0.950571	0.950596	0.950449	0.950410	0.950571	0.950487
bias β	-0.010761	-0.017757	-0.028808	0.000571	0.000596	0.000449	0.000410	0.000571	0.000487
mse β	0.008842	0.025796	0.001543	0.000007	0.000005	0.000007	0.000024	0.000006	0.000023
mae β	0.053047	0.082917	0.029686	0.002000	0.001679	0.002015	0.002288	0.001694	0.002338
mean σ	0.171005	0.093941	0.178377	0.262543	0.272532	0.258514	0.255426	0.269208	0.252188
bias σ	-0.088995	-0.166059	-0.081623	0.002543	0.012532	-0.001486	-0.004574	0.009208	-0.007812
mse σ	0.028943	0.047305	0.009177	0.002206	0.001896	0.002725	0.003091	0.002087	0.003352
mae σ	0.142804	0.200365	0.082031	0.036923	0.036060	0.041672	0.045007	0.038874	0.047074

TABLE 5. Reference SV Model, Sample Size 250 - α=-0.368 β=.95 σ=.26

	GMM2S	GMMITER	GMMCUE	GEL	ET	ETEL	SGEL	SET	SETEL
mean α	-0.287550	-0.358624	-0.178688	-0.184597	-0.174181	-0.172286	-0.184011	-0.175501	-0.178481
bias α	-0.140350	-0.211424	-0.031488	-0.037397	-0.026981	-0.025086	-0.036811	-0.028301	-0.031281
mse α	0.480456	1.419835	0.036347	0.003332	0.001820	0.002089	0.004401	0.002261	0.002978
mae α	0.290438	0.426097	0.052332	0.039280	0.027872	0.026591	0.040621	0.029496	0.033409
mean β	0.961360	0.951781	0.976673	0.975317	0.976275	0.976645	0.975205	0.976093	0.975689
bias β	-0.018640	-0.028219	-0.003327	-0.004683	-0.003725	-0.003355	-0.004795	-0.003907	-0.004311
mse β	0.008904	0.025951	0.000614	0.000053	0.000046	0.000051	0.000067	0.000056	0.000068
mae β	0.039143	0.057390	0.006172	0.005015	0.003846	0.003566	0.005370	0.004078	0.004620
mean σ	0.099607	0.058895	0.147883	0.162567	0.165318	0.152367	0.154270	0.165148	0.149183
bias σ	-0.066093	-0.106805	-0.017817	-0.003133	-0.000382	-0.013333	-0.011430	-0.000552	-0.016517
mse σ	0.018376	0.022620	0.003918	0.002296	0.002296	0.002485	0.002566	0.002073	0.002865
mae σ	0.117220	0.138037	0.022491	0.038325	0.037033	0.038657	0.041356	0.035026	0.041933

TABLE 6. Reference SV Model, Sample Size 250 - α-.1472 β=.98 σ=.1657

Reference SV Model - Sample Size 250.

	GMM2S	GMMITER	GMMCUE	GEL	ET	ETEL	SGEL	SET	SETEL
mean α	-0.657239	-0.655518	-0.646651	-0.732105	-0.731168	-0.728442	-0.731064	-0.728545	-0.730313
bias α	0.078761	0.080482	0.089349	0.003895	0.004832	0.007558	0.004936	0.007455	0.005687
mse α	0.324049	0.448948	0.321806	0.000529	0.000103	0.000284	0.000186	0.000103	0.000182
mae α	0.431807	0.524950	0.412616	0.013323	0.007189	0.011075	0.009587	0.007935	0.009229
mean β	0.911393	0.911682	0.912914	0.901757	0.901142	0.901871	0.901144	0.901450	0.901280
bias β	0.011393	0.011682	0.012914	0.001757	0.001142	0.001871	0.001144	0.001450	0.001280
mse β	0.005944	0.008215	0.005888	0.000011	0.000004	0.000009	0.000007	0.000005	0.000007
mae β	0.058491	0.070958	0.056019	0.002499	0.001646	0.002257	0.002140	0.001763	0.002053
mean σ	0.246982	0.220967	0.227270	0.381270	0.386004	0.377610	0.374018	0.380753	0.376954
bias σ	-0.115918	-0.141933	-0.135630	0.018370	0.023104	0.014710	0.011118	0.017853	0.014054
mse σ	0.029136	0.041257	0.034827	0.001272	0.001388	0.001484	0.001415	0.001299	0.001300
mae σ	0.146245	0.173349	0.158733	0.027994	0.031482	0.030527	0.031812	0.031192	0.030416

TABLE 7. Reference SV Model, Sample Size 1000 - α=-0.736 β=.9 σ=.3629

	GMM2S	GMMITER	GMMCUE	GEL	ET	ETEL	SGEL	SET	SETEL
mean α	-0.300991	-0.300362	-0.476716	-0.372371	-0.369634	-0.371808	-0.368965	-0.368831	-0.369742
bias α	0.067009	0.067638	-0.108716	-0.004371	-0.001634	-0.003808	-0.000965	-0.000831	-0.001742
mse α	0.108547	0.154086	0.016805	0.000462	0.000154	0.000212	0.000287	0.000194	0.000300
mae α	0.248294	0.300398	0.110979	0.011809	0.009061	0.010973	0.010886	0.009801	0.010868
mean β	0.959360	0.959477	0.935880	0.949738	0.949870	0.949686	0.949956	0.949983	0.949848
bias β	0.009360	0.009477	-0.014120	-0.000262	-0.000130	-0.000314	-0.000044	-0.000017	-0.000152
mse β	0.002000	0.002828	0.000297	0.000008	0.000003	0.000004	0.000005	0.000003	0.000005
mae β	0.033712	0.040665	0.014485	0.001599	0.001237	0.001470	0.001397	0.001205	0.001375
mean σ	0.156376	0.134259	0.215652	0.264582	0.270840	0.264941	0.260428	0.269523	0.261419
bias σ	-0.103624	-0.125741	-0.044348	0.004582	0.010840	0.004941	0.000428	0.009523	0.001419
mse σ	0.020798	0.028795	0.002460	0.001239	0.001213	0.001342	0.001352	0.001238	0.001391
mae σ	0.122116	0.146295	0.044402	0.027305	0.029037	0.029162	0.030043	0.029670	0.030514

TABLE 8. Reference SV Model, Sample Size 1000 - α=-0.368 β=.95 σ=.26

	GMM2S	GMMITER	GMMCUE	GEL	ET	ETEL	SGEL	SET	SETEL
mean α	-0.114335	-0.119373	-0.150316	-0.173171	-0.164059	-0.163896	-0.174346	-0.164044	-0.165767
bias α	0.032865	0.027827	-0.003116	-0.025971	-0.016859	-0.016696	-0.027146	-0.016844	-0.018567
mse α	0.043901	0.058947	0.004958	0.001341	0.000497	0.000473	0.001637	0.000563	0.001227
mae α	0.146038	0.166462	0.022415	0.027646	0.017471	0.017493	0.029596	0.017999	0.020478
mean β	0.984552	0.983893	0.980499	0.976673	0.977706	0.977774	0.976394	0.977703	0.977488
bias β	0.004552	0.003893	0.000499	-0.003327	-0.002294	-0.002226	-0.003606	-0.002297	-0.002512
mse β	0.000818	0.001091	0.000081	0.000021	0.000010	0.000009	0.000027	0.000012	0.000023
mae β	0.019818	0.022531	0.002246	0.003557	0.002355	0.002325	0.003912	0.002425	0.002743
mean σ	0.078361	0.066519	0.150921	0.174408	0.171587	0.167460	0.173789	0.171004	0.167570
bias σ	-0.087339	-0.099181	-0.014779	0.008708	0.005887	0.001760	0.008089	0.005304	0.001870
mse σ	0.013792	0.016902	0.001312	0.001195	0.000971	0.000976	0.001379	0.000960	0.001094
mae σ	0.104885	0.118200	0.015356	0.026957	0.024514	0.024630	0.029696	0.024281	0.026283

TABLE 9. Reference SV Model, Sample Size 1000 - α-.1472 β=.98 σ=.1657

Reference SV Model - Sample Size 1000.

	GMM2S	GMMITER	GMMCUE	GEL	ET	ETEL	SGEL	SET	SETEL
mean α	-0.387790	-0.329887	-0.392345	-0.706187	-0.583088	-0.705202	-0.734080	-0.688249	-0.737702
bias α	0.348210	0.406113	0.343655	0.029813	0.152912	0.030798	0.001920	0.047751	-0.001702
mse α	0.407410	0.529764	0.448516	0.030564	0.135697	0.059130	0.018108	0.298280	0.003793
mae α	0.500311	0.579126	0.488393	0.075290	0.240673	0.100716	0.035211	0.243153	0.015571
mean β	0.947847	0.955695	0.947246	0.904431	0.920801	0.905838	0.901287	0.907336	0.901759
bias β	0.047847	0.055695	0.047246	0.004431	0.020801	0.005838	0.001287	0.007336	0.001759
mse β	0.007420	0.009644	0.008133	0.000573	0.002498	0.001141	0.000342	0.005458	0.000083
mae β	0.067970	0.078549	0.066343	0.010849	0.033067	0.015050	0.005682	0.033610	0.003415
mean σ	0.177854	0.141572	0.166061	0.345979	0.256893	0.384320	0.376598	0.293342	0.398028
bias σ	-0.185046	-0.221328	-0.196839	-0.016921	-0.106007	0.021420	0.013698	-0.069558	0.035128
mse σ	0.053409	0.071263	0.057475	0.006115	0.023314	0.006100	0.003235	0.020749	0.002738
mae σ	0.204028	0.240183	0.213535	0.061064	0.125973	0.061736	0.045643	0.110763	0.044111

TABLE 10. Reference SV Model, Subset of Instruments - α=-0.736 β=.9 σ=.3629

	GMM2S	GMMITER	GMMCUE	GEL	ET	ETEL	SGEL	SET	SETEL
mean α	-0.176010	-0.167857	-0.424666	-0.371692	-0.361446	-0.367182	-0.372239	-0.362772	-0.367428
bias α	0.191990	0.200143	-0.056666	-0.003692	0.006554	0.000818	-0.004239	0.005228	0.000572
mse α	0.226799	0.303138	0.012526	0.000276	0.000203	0.000169	0.000345	0.000091	0.000117
mae α	0.323536	0.349715	0.070120	0.010155	0.010376	0.008243	0.012456	0.006924	0.006971
mean β	0.976343	0.977480	0.942910	0.951056	0.951271	0.950586	0.951122	0.951025	0.950734
bias β	0.026343	0.027480	-0.007090	0.001056	0.001271	0.000586	0.001122	0.001025	0.000734
mse β	0.004069	0.005476	0.000219	0.000010	0.000006	0.000006	0.000010	0.000003	0.000004
mae β	0.043860	0.047336	0.009105	0.002251	0.001941	0.001689	0.002211	0.001380	0.001501
mean σ	0.100259	0.081734	0.214449	0.289414	0.269170	0.284782	0.285166	0.281602	0.292884
bias σ	-0.159741	-0.178266	-0.045551	0.029414	0.009170	0.024782	0.025166	0.021602	0.032884
mse σ	0.037548	0.044987	0.003587	0.001766	0.001190	0.001708	0.001545	0.001222	0.002284
mae σ	0.176449	0.195545	0.046372	0.032133	0.027369	0.030589	0.029887	0.028479	0.035764

TABLE 11. Reference SV Model, Subset of Instruments - α=-0.368 β=.95 σ=.26

	GMM2S	GMMITER	GMMCUE	GEL	ET	ETEL	SGEL	SET	SETEL
mean α	-0.089063	-0.092245	-0.146779	-0.160451	-0.160487	-0.152710	-0.158337	-0.151514	-0.154633
bias α	0.058137	0.054955	0.000421	-0.013251	-0.013287	-0.005510	-0.011137	-0.004314	-0.007433
mse α	0.074543	0.137264	0.004951	0.000839	0.000555	0.000188	0.000673	0.000117	0.000291
mae α	0.163815	0.178350	0.028959	0.017083	0.015205	0.007540	0.016316	0.006146	0.010725
mean β	0.988033	0.987660	0.980816	0.979301	0.978390	0.979612	0.979438	0.979615	0.979451
bias β	0.008033	0.007660	0.000816	-0.000699	-0.001610	-0.000388	-0.000562	-0.000385	-0.000549
mse β	0.001332	0.002385	0.000081	0.000010	0.000011	0.000003	0.000009	0.000003	0.000004
mae β	0.022137	0.024041	0.003246	0.001962	0.002243	0.001172	0.001775	0.001188	0.001340
mean σ	0.056143	0.047436	0.146157	0.196089	0.176797	0.191231	0.189516	0.186095	0.196360
bias σ	-0.109557	-0.118264	-0.019543	0.030389	0.011097	0.025531	0.023816	0.020395	0.030660
mse σ	0.018120	0.020777	0.002100	0.002081	0.001214	0.001677	0.001812	0.001431	0.002110
mae σ	0.126418	0.135976	0.020115	0.033161	0.026940	0.029457	0.029030	0.028479	0.034580

TABLE 12. Reference SV Model, Subset of Instruments - α-.1472 β=.98 σ=.1657

Reference SV Model - Subset of Instruments.

	GMM2S	GMMITER	GMMCUE	GEL	ET	ETEL	SGEL	SET	SETEL
mean α	-0.898594	-0.840806	-0.807848	-0.704411	-0.710422	-0.711583	-0.727019	-0.752694	-0.725200
bias α	-0.162594	-0.104806	-0.071848	0.031589	0.025578	0.024417	0.008981	-0.016694	0.010800
mse α	0.390480	0.981390	0.843844	0.031866	0.142569	0.013417	0.002774	0.072544	0.000397
mae α	0.421985	0.677693	0.578302	0.077138	0.211659	0.044133	0.016357	0.072182	0.014560
mean β	0.880334	0.887774	0.892161	0.903264	0.902588	0.903792	0.899936	0.896904	0.900473
bias β	-0.019666	-0.012226	-0.007839	0.003264	0.002588	0.003792	-0.000064	-0.003096	0.000473
mse β	0.006902	0.017631	0.015195	0.000610	0.002765	0.000260	0.000052	0.001422	0.000007
mae β	0.056283	0.090791	0.078059	0.010621	0.029255	0.006682	0.001934	0.009835	0.001803
mean σ	0.411426	0.276272	0.277088	0.365962	0.330249	0.392250	0.401249	0.387742	0.398329
bias σ	0.048526	-0.086628	-0.085812	0.003062	-0.032651	0.029350	0.038349	0.024842	0.035429
mse σ	0.035925	0.046839	0.039294	0.005068	0.012450	0.004109	0.002003	0.004555	0.001863
mae σ	0.144036	0.179405	0.162927	0.053270	0.085005	0.051261	0.041188	0.051046	0.039298

TABLE 13. Student t (4) Innovations in Mean - α=-0.736 β=.9 σ=.3629

	GMM2S	GMMITER	GMMCUE	GEL	ET	ETEL	SGEL	SET	SETEL
mean α	-0.418591	-0.409907	-0.436561	-0.366868	-0.364062	-0.366793	-0.365613	-0.364053	-0.365115
bias α	-0.050591	-0.041907	-0.068561	0.001132	0.003938	0.001207	0.002387	0.003947	0.002885
mse α	0.162877	0.504704	0.016724	0.000260	0.000175	0.000259	0.000197	0.000180	0.000215
mae α	0.268665	0.420585	0.078071	0.012478	0.010332	0.012119	0.010687	0.010435	0.010866
mean β	0.944253	0.945494	0.941962	0.950021	0.950277	0.949960	0.950228	0.950265	0.950229
bias β	-0.005747	-0.004506	-0.008038	0.000021	0.000277	-0.000040	0.000228	0.000265	0.000229
mse β	0.002946	0.008605	0.000312	0.000005	0.000002	0.000005	0.000003	0.000003	0.000003
mae β	0.036112	0.056316	0.009965	0.001655	0.001184	0.001629	0.001271	0.001118	0.001276
mean σ	0.258888	0.163107	0.230616	0.308080	0.304090	0.304898	0.305641	0.302799	0.303441
bias σ	-0.001112	-0.096893	-0.029384	0.048080	0.044090	0.044898	0.045641	0.042799	0.043441
mse σ	0.022531	0.036115	0.002240	0.004447	0.003247	0.003896	0.004327	0.003056	0.003854
mae σ	0.120120	0.160839	0.033712	0.053685	0.047893	0.051585	0.053479	0.047292	0.051053

TABLE 14. Student t (4) Innovations in Mean - α=-0.368 β=.95 σ=.26

	GMM2S	GMMITER	GMMCUE	GEL	ET	ETEL	SGEL	SET	SETEL
mean α	-0.182803	-0.207332	-0.153876	-0.178418	-0.172451	-0.171749	-0.179504	-0.174035	-0.173768
bias α	-0.035603	-0.060132	-0.006676	-0.031218	-0.025251	-0.024549	-0.032304	-0.026835	-0.026568
mse α	0.064698	0.362934	0.015031	0.001764	0.001905	0.001836	0.001910	0.002270	0.002064
mae α	0.162150	0.250030	0.023437	0.032214	0.025756	0.025222	0.033189	0.027654	0.027348
mean β	0.975580	0.972639	0.979965	0.975852	0.976312	0.976454	0.975630	0.976079	0.976160
bias β	-0.004420	-0.007361	-0.000035	-0.004148	-0.003688	-0.003546	-0.004370	-0.003921	-0.003840
mse β	0.001189	0.005761	0.000260	0.000027	0.000049	0.000047	0.000031	0.000058	0.000052
mae β	0.021753	0.033207	0.002463	0.004285	0.003778	0.003645	0.004497	0.004054	0.003948
mean σ	0.141434	0.087524	0.153574	0.206849	0.202369	0.199885	0.205119	0.203117	0.200861
bias σ	-0.024266	-0.078176	-0.012126	0.041150	0.036669	0.034185	0.039419	0.037417	0.035161
mse σ	0.013726	0.021573	0.001190	0.003876	0.003752	0.003589	0.003721	0.004012	0.003858
mae σ	0.099130	0.127428	0.013556	0.048306	0.044964	0.044559	0.048999	0.045614	0.045999

TABLE 15. Student t (4) Innovations in Mean - α-.1472 β=.98 σ=.1657

SV Model - Student-t in Mean Equation.

	GMM2S	GMMITER	GMMCUE	GEL	ET	ETEL	SGEL	SET	SETEL
mean α	-0.489227	-0.409877	-0.414510	-0.645279	-0.527524	-0.688817	-0.732088	-0.636541	-0.725662
bias α	0.246773	0.326123	0.321490	0.090721	0.208476	0.047183	0.003912	0.099459	0.010338
mse α	0.294811	0.598153	0.570777	0.062880	0.157706	0.159874	0.034620	0.208947	0.001027
mae α	0.447530	0.597699	0.566140	0.135388	0.277889	0.169003	0.050133	0.260987	0.019382
mean β	0.934580	0.945279	0.944608	0.911168	0.927605	0.905506	0.899820	0.913821	0.901443
bias β	0.034579	0.045279	0.044608	0.011168	0.027605	0.005506	-0.000180	0.013821	0.001443
mse β	0.005378	0.010677	0.010224	0.001163	0.002921	0.003044	0.000682	0.003908	0.000023
mae β	0.060665	0.080809	0.076761	0.018336	0.037783	0.023940	0.006913	0.035885	0.002971
mean σ	0.290529	0.203847	0.211820	0.352261	0.299423	0.379088	0.388014	0.323308	0.398034
bias σ	-0.072371	-0.159053	-0.151080	-0.010639	-0.063477	0.016188	0.025114	-0.039592	0.035134
mse σ	0.040188	0.061685	0.053119	0.009106	0.022248	0.008340	0.003194	0.019586	0.002666
mae σ	0.161437	0.214730	0.199325	0.073127	0.118990	0.069374	0.045031	0.103051	0.043476

TABLE 16. Student t (4) Innovations in Mean, Subset of Instruments - α=-0.736 β=.9 σ=.3629

	GMM2S	GMMITER	GMMCUE	GEL	ET	ETEL	SGEL	SET	SETEL
mean α	-0.212205	-0.188704	-0.381028	-0.370669	-0.361173	-0.367182	-0.373580	-0.361736	-0.366302
bias α	0.155795	0.179296	-0.013028	-0.002669	0.006827	0.000818	-0.005580	0.006264	0.001698
mse α	0.151073	0.250103	0.004146	0.000320	0.000217	0.000351	0.000469	0.000131	0.000170
mae α	0.289612	0.348350	0.041639	0.011552	0.011170	0.011661	0.014248	0.008588	0.008391
mean β	0.971552	0.974813	0.949139	0.950362	0.950847	0.949851	0.950648	0.950749	0.950230
bias β	0.021552	0.024813	-0.000861	0.000362	0.000847	-0.000149	0.000648	0.000749	0.000230
mse β	0.002782	0.004466	0.000074	0.000012	0.000006	0.000011	0.000011	0.000004	0.000006
mae β	0.039291	0.047117	0.005421	0.002340	0.001886	0.002199	0.002342	0.001441	0.001700
mean σ	0.163153	0.114320	0.248484	0.311227	0.304519	0.310048	0.310693	0.303892	0.307685
bias σ	-0.096847	-0.145680	-0.011516	0.051227	0.044518	0.050048	0.050693	0.043892	0.047685
mse σ	0.030464	0.042008	0.001299	0.004046	0.003705	0.004212	0.004070	0.002848	0.003464
mae σ	0.149490	0.184456	0.024816	0.052409	0.048864	0.051927	0.052690	0.045343	0.048842

TABLE 17. Student t (4) Innovations in Mean, Subset of Instruments - α=-0.368 β=.95 σ=.26

	GMM2S	GMMITER	GMMCUE	GEL	ET	ETEL	SGEL	SET	SETEL
mean α	-0.074284	-0.071013	-0.150531	-0.162376	-0.166849	-0.154689	-0.158033	-0.155149	-0.153788
bias α	0.072916	0.076187	-0.003331	-0.015176	-0.019649	-0.007489	-0.010833	-0.007949	-0.006588
mse α	0.028767	0.072554	0.008817	0.000874	0.001352	0.000405	0.000949	0.000434	0.000259
mae α	0.139566	0.156468	0.026179	0.019321	0.021606	0.009982	0.017845	0.009705	0.010202
mean β	0.989995	0.990442	0.980081	0.978495	0.977240	0.978996	0.979339	0.978898	0.979279
bias β	0.009995	0.010442	0.000081	-0.001505	-0.002760	-0.001004	-0.000661	-0.001102	-0.000721
mse β	0.000533	0.001329	0.000149	0.000016	0.000033	0.000011	0.000014	0.000013	0.000005
mae β	0.018938	0.021213	0.003100	0.002503	0.003344	0.001732	0.002077	0.001849	0.001444
mean σ	0.077686	0.057410	0.154182	0.209621	0.203023	0.207680	0.206146	0.205631	0.209157
bias σ	-0.088014	-0.108290	-0.011518	0.043921	0.037323	0.041980	0.040446	0.039931	0.043457
mse σ	0.015682	0.019399	0.001735	0.003441	0.003193	0.003309	0.003071	0.002986	0.003199
mae σ	0.115388	0.130757	0.018132	0.045794	0.042404	0.043876	0.042707	0.042980	0.045337

TABLE 18. Student t (4) Innovations in Mean, Subset of Instruments - α-.1472 β=.98 σ=.1657

SV Model - Student-t in Mean Equation - Subset of Instruments.

	GMM2S	GMMITER	GMMCUE	GEL	ET	ETEL	SGEL	SET	SETEL
mean α	-2.019463	-2.019004	-1.353874	-0.719185	-0.878625	-0.718273	-0.726762	-0.874355	-0.715406
bias α	-1.283463	-1.283004	-0.617874	0.016815	-0.142625	0.017727	0.009238	-0.138355	0.020594
mse α	5.053566	6.598158	3.044399	0.058803	0.680573	0.015097	0.016660	0.644231	0.001293
mae α	1.475510	1.659492	1.051613	0.099185	0.432671	0.048763	0.039363	0.277956	0.024200
mean β	0.720150	0.719250	0.811384	0.898652	0.876085	0.900611	0.897372	0.877314	0.899777
bias β	-0.179850	-0.180750	-0.088616	-0.001348	-0.023915	0.000611	-0.002628	-0.022686	-0.000223
mse β	0.097458	0.128318	0.059642	0.001216	0.013893	0.000300	0.000337	0.012709	0.000020
mae β	0.205075	0.231269	0.146937	0.013264	0.060945	0.006534	0.004285	0.038086	0.002403
mean σ	0.532352	0.431597	0.357978	0.366275	0.346430	0.393303	0.397762	0.381166	0.397085
bias σ	0.169452	0.068697	-0.004922	0.003375	-0.016470	0.030403	0.034862	0.018266	0.034185
mse σ	0.109984	0.085714	0.058204	0.005314	0.023325	0.003714	0.002450	0.014768	0.001909
mae σ	0.251868	0.239819	0.198540	0.051763	0.111622	0.047475	0.043035	0.079160	0.038535

TABLE 19. Student t (4) Innovations in Variance - α=-0.736 β=.9 σ=.3629

	GMM2S	GMMITER	GMMCUE	GEL	ET	ETEL	SGEL	SET	SETEL
mean α	-1.531838	-1.505341	-0.411598	-0.363040	-0.356509	-0.360347	-0.357152	-0.355783	-0.356305
bias α	-1.163838	-1.137341	-0.043598	0.004960	0.011491	0.007653	0.010848	0.012217	0.011695
mse α	4.582102	5.703206	0.006206	0.000298	0.000293	0.000355	0.000448	0.000265	0.000321
mae α	1.265176	1.353192	0.048679	0.013481	0.015033	0.015479	0.015028	0.014312	0.015145
mean β	0.786817	0.789921	0.942956	0.949416	0.949709	0.949327	0.949691	0.949781	0.949705
bias β	-0.163183	-0.160079	-0.007044	-0.000584	-0.000291	-0.000673	-0.000309	-0.000219	-0.000295
mse β	0.089675	0.112105	0.000128	0.000005	0.000002	0.000005	0.000005	0.000002	0.000003
mae β	0.176775	0.189182	0.007696	0.001525	0.001065	0.001646	0.001189	0.000972	0.001124
mean σ	0.438012	0.340150	0.227405	0.294889	0.292910	0.290458	0.282838	0.292852	0.281492
bias σ	0.178012	0.080150	-0.032595	0.034889	0.032910	0.030458	0.022838	0.032852	0.021492
mse σ	0.111763	0.085479	0.001741	0.002467	0.001768	0.002375	0.001898	0.001811	0.001844
mae σ	0.241064	0.225776	0.034656	0.041202	0.037140	0.040675	0.036961	0.038883	0.037063

TABLE 20. Student t (4) Innovations in Variance - α=-0.368 β=.95 σ=.26

	GMM2S	GMMITER	GMMCUE	GEL	ET	ETEL	SGEL	SET	SETEL
mean α	-1.183910	-1.169166	-0.139778	-0.169468	-0.177636	-0.176430	-0.171889	-0.177510	-0.180039
bias α	-1.036710	-1.021966	0.007422	-0.022268	-0.030436	-0.029230	-0.024689	-0.030310	-0.032839
mse α	4.167620	5.284636	0.002062	0.000871	0.003163	0.003335	0.001260	0.002726	0.003574
mae α	1.083632	1.122330	0.019276	0.022964	0.030810	0.030064	0.026382	0.030722	0.033550
mean β	0.833844	0.835416	0.981116	0.976373	0.974736	0.974993	0.975766	0.974741	0.974392
bias β	-0.146156	-0.144584	0.001116	-0.003627	-0.005264	-0.005007	-0.004234	-0.005259	-0.005608
mse β	0.083851	0.106097	0.000032	0.000019	0.000090	0.000091	0.000030	0.000079	0.000099
mae β	0.152473	0.158181	0.002408	0.003689	0.005304	0.005093	0.004429	0.005307	0.005681
mean σ	0.340019	0.254258	0.148694	0.191014	0.199949	0.194585	0.182533	0.197417	0.188237
bias σ	0.174319	0.088558	-0.017006	0.025314	0.034249	0.028885	0.016833	0.031717	0.022537
mse σ	0.097556	0.071992	0.001456	0.001879	0.003394	0.003096	0.001858	0.003056	0.002708
mae σ	0.222331	0.195336	0.017893	0.033934	0.042568	0.040125	0.033289	0.041190	0.037962

TABLE 21. Student t (4) Innovations in Variance - α-.1472 β=.98 σ=.1657

SV Model - Student-t in Volatility Equation.

	GMM2S	GMMITER	GMMCUE	GEL	ET	ETEL	SGEL	SET	SETEL
mean α	-1.624080	-1.558604	-1.060265	-0.734392	-0.676903	-0.956436	-0.853334	-1.077795	-0.725856
bias α	-0.888080	-0.822604	-0.324265	0.001608	0.059097	-0.220436	-0.117334	-0.341795	0.010144
mse α	3.928852	4.590728	2.015175	0.173709	0.230651	0.815326	0.491227	1.434573	0.088954
mae α	1.200330	1.292493	0.849144	0.137095	0.238348	0.374986	0.176277	0.537670	0.045287
mean β	0.773478	0.782450	0.851910	0.896662	0.905272	0.864619	0.879744	0.849333	0.899025
bias β	-0.126522	-0.117550	-0.048090	-0.003338	0.005272	-0.035381	-0.020256	-0.050667	-0.000975
mse β	0.077261	0.090068	0.039782	0.003338	0.004464	0.017650	0.009865	0.027960	0.001882
mae β	0.167782	0.180375	0.118296	0.018190	0.032246	0.054444	0.023862	0.074613	0.005793
mean σ	0.490207	0.430327	0.370190	0.380871	0.359469	0.428566	0.398392	0.407583	0.403044
bias σ	0.127307	0.067427	0.007290	0.017971	-0.003431	0.065666	0.035492	0.044683	0.040144
mse σ	0.097302	0.093233	0.060965	0.007639	0.015327	0.017199	0.006391	0.031044	0.003633
mae σ	0.226197	0.236678	0.191858	0.061614	0.091534	0.096654	0.051057	0.117093	0.048806

TABLE 22. Student t (4) Innovations in Variance, Subset of Instruments - α=-0.736 β=.9 σ=.3629

	GMM2S	GMMITER	GMMCUE	GEL	ET	ETEL	SGEL	SET	SETEL
mean α	-1.197984	-1.098225	-0.361745	-0.370533	-0.361434	-0.368460	-0.371025	-0.358880	-0.365012
bias α	-0.829984	-0.730225	0.006255	-0.002533	0.006566	-0.000460	-0.003025	0.009120	0.002988
mse α	3.604356	3.690784	0.001912	0.000341	0.000276	0.000471	0.000384	0.000188	0.000241
mae α	0.987495	0.971869	0.017620	0.011394	0.013219	0.015002	0.012487	0.011009	0.010551
mean β	0.832283	0.845932	0.949841	0.949360	0.949605	0.948366	0.950019	0.949814	0.949038
bias β	-0.117717	-0.104068	-0.000159	-0.000640	-0.000395	-0.001634	0.000019	-0.000186	-0.000962
mse β	0.071710	0.073722	0.000037	0.000010	0.000007	0.000016	0.000007	0.000002	0.000007
mae β	0.138731	0.136549	0.002583	0.002030	0.001903	0.002706	0.001824	0.001152	0.001797
mean σ	0.387174	0.328301	0.261493	0.316921	0.301799	0.312991	0.310528	0.307241	0.316634
bias σ	0.127174	0.068301	0.001493	0.056921	0.041799	0.052991	0.050528	0.047241	0.056634
mse σ	0.092473	0.081984	0.001088	0.004417	0.002851	0.004073	0.003565	0.002990	0.004241
mae σ	0.208110	0.207164	0.019947	0.057437	0.045345	0.054480	0.051266	0.049407	0.057532

TABLE 23. Student t (4) Innovations in Variance, Subset of Instruments - α=-0.368 β=.95 σ=.26

	GMM2S	GMMITER	GMMCUE	GEL	ET	ETEL	SGEL	SET	SETEL
mean α	-1.013611	-0.966006	-0.160485	-0.162815	-0.170481	-0.159132	-0.158832	-0.158234	-0.152619
bias α	-0.866411	-0.818806	-0.013285	-0.015615	-0.023281	-0.011932	-0.011632	-0.011034	-0.005419
mse α	4.022485	4.187654	0.038329	0.001189	0.001502	0.007726	0.001482	0.000565	0.000261
mae α	0.936151	0.921012	0.031955	0.018731	0.024269	0.014440	0.015531	0.013051	0.008082
mean β	0.857288	0.863844	0.977562	0.977898	0.976214	0.977718	0.978634	0.977919	0.978892
bias β	-0.122712	-0.116156	-0.002438	-0.002102	-0.003786	-0.002282	-0.001366	-0.002081	-0.001108
mse β	0.080732	0.084098	0.000880	0.000030	0.000039	0.000203	0.000033	0.000017	0.000008
mae β	0.132099	0.129967	0.004854	0.002841	0.003945	0.002639	0.002217	0.002398	0.001452
mean σ	0.304510	0.262075	0.168088	0.211233	0.203048	0.207656	0.205108	0.212755	0.214762
bias σ	0.138810	0.096375	0.002388	0.045533	0.037349	0.041956	0.039408	0.047055	0.049062
mse σ	0.086786	0.077900	0.002828	0.003107	0.002707	0.003095	0.002736	0.003547	0.003678
mae σ	0.201054	0.195892	0.017739	0.046314	0.040821	0.043382	0.040756	0.049855	0.050215

TABLE 24. Student t (4) Innovations in Variance, Subset of Instruments - α-.1472 β=.98 σ=.1657

SV Model - Student-t in Volatility Equation - Subset of Instruments.

	GMM2S	GMMITER	GMMCUE	GEL	ET	ETEL	SGEL	SET	SETEL
mean α	-0.896686	-0.921662	-0.796978	-0.728243	-0.835926	-0.729340	-0.734904	-0.757911	-0.732839
bias α	-0.160686	-0.185662	-0.060978	0.007757	-0.099926	0.006660	0.001096	-0.021911	0.003161
mse α	0.758956	1.251152	0.627006	0.029238	0.367820	0.005461	0.001601	0.075302	0.000141
mae α	0.563200	0.764779	0.503486	0.055221	0.350987	0.024025	0.010621	0.085502	0.008789
mean β	0.879116	0.875898	0.892628	0.901656	0.885525	0.903877	0.900704	0.897573	0.901291
bias β	-0.020884	-0.024102	-0.007372	0.001656	-0.014475	0.003877	0.000704	-0.002427	0.001291
mse β	0.013699	0.022533	0.011362	0.000583	0.007103	0.000121	0.000031	0.001436	0.000007
mae β	0.076157	0.103184	0.068596	0.008517	0.048875	0.005174	0.001878	0.012231	0.001886
mean σ	0.287003	0.236533	0.232068	0.355123	0.286266	0.399452	0.397132	0.351652	0.393545
bias σ	-0.075897	-0.126367	-0.130832	-0.007777	-0.076634	0.036552	0.034232	-0.011248	0.030645
mse σ	0.032933	0.049530	0.039221	0.004932	0.019929	0.003515	0.002085	0.008075	0.001812
mae σ	0.150171	0.189751	0.171741	0.055842	0.114077	0.048633	0.041545	0.069468	0.037968

TABLE 25. Level Outlier - α=-0.736 β=.9 σ=.3629

	GMM2S	GMMITER	GMMCUE	GEL	ET	ETEL	SGEL	SET	SETEL
mean α	-0.450450	-0.452413	-0.505817	-0.369780	-0.366575	-0.368322	-0.367022	-0.366164	-0.365422
bias α	-0.082450	-0.084413	-0.137817	-0.001780	0.001425	-0.000322	0.000978	0.001836	0.002578
mse α	0.351303	0.524573	0.029065	0.000364	0.000199	0.000283	0.000401	0.000209	0.000378
mae α	0.375102	0.464943	0.140401	0.012881	0.010511	0.012609	0.012380	0.011065	0.012730
mean β	0.939147	0.938991	0.931884	0.950220	0.950221	0.950119	0.950194	0.950278	0.950372
bias β	-0.010853	-0.011009	-0.018116	0.000220	0.000221	0.000119	0.000194	0.000278	0.000372
mse β	0.006482	0.009584	0.000528	0.000007	0.000003	0.000005	0.000006	0.000003	0.000006
mae β	0.050839	0.062913	0.018567	0.001763	0.001325	0.001726	0.001484	0.001256	0.001481
mean σ	0.185589	0.143397	0.208345	0.266704	0.273863	0.263786	0.259905	0.272609	0.259491
bias σ	-0.074411	-0.116603	-0.051655	0.006704	0.013863	0.003786	-0.000095	0.012609	-0.000509
mse σ	0.026369	0.036223	0.003550	0.001492	0.001554	0.001678	0.001966	0.001545	0.002029
mae σ	0.134897	0.165390	0.051706	0.030328	0.032812	0.033478	0.036517	0.033588	0.036596

TABLE 26. Level Outlier - α=-0.368 β=.95 σ=.26

	GMM2S	GMMITER	GMMCUE	GEL	ET	ETEL	SGEL	SET	SETEL
mean α	-0.227624	-0.232687	-0.156716	-0.177422	-0.167124	-0.166912	-0.178924	-0.168041	-0.169662
bias α	-0.080424	-0.085487	-0.009516	-0.030222	-0.019924	-0.019712	-0.031724	-0.020841	-0.022462
mse α	0.182403	0.256869	0.008321	0.001678	0.000806	0.000916	0.002641	0.000985	0.001197
mae α	0.234099	0.269799	0.031758	0.031825	0.020514	0.020990	0.034853	0.021881	0.024061
mean β	0.969217	0.968663	0.979613	0.976150	0.977225	0.977292	0.975734	0.977080	0.976873
bias β	-0.010783	-0.011337	-0.000387	-0.003850	-0.002775	-0.002708	-0.004266	-0.002920	-0.003127
mse β	0.003380	0.004678	0.000148	0.000025	0.000019	0.000021	0.000044	0.000023	0.000026
mae β	0.031679	0.036386	0.003503	0.004078	0.002851	0.002860	0.004695	0.003064	0.003324
mean σ	0.107893	0.083627	0.145206	0.174273	0.172216	0.163029	0.167475	0.170336	0.162140
bias σ	-0.057807	-0.082073	-0.020494	0.008573	0.006516	-0.002671	0.001775	0.004636	-0.003560
mse σ	0.016458	0.020176	0.002062	0.001507	0.001380	0.001435	0.001771	0.001392	0.001579
mae σ	0.112344	0.127673	0.020619	0.030929	0.029741	0.029958	0.034131	0.029380	0.032292

TABLE 27. Level Outlier - α-.1472 β=.98 σ=.1657

SV Model - Level Outlier.

	GMM2S	GMMITER	GMMCUE	GEL	ET	ETEL	SGEL	SET	SETEL
mean α	-0.606591	-0.573558	-0.538447	-0.687828	-0.625819	-0.706385	-0.774831	-0.809760	-0.733482
bias α	0.129409	0.162442	0.197553	0.048172	0.110181	0.029615	-0.038831	-0.073760	0.002518
mse α	0.584982	0.692991	0.525481	0.080173	0.219997	0.190438	0.116738	0.801588	0.001846
mae α	0.515088	0.569472	0.498117	0.135071	0.306007	0.204829	0.100064	0.469002	0.017395
mean β	0.918228	0.922744	0.927494	0.906621	0.915070	0.905568	0.895372	0.890500	0.902918
bias β	0.018228	0.022744	0.027494	0.006621	0.015070	0.005568	-0.004628	-0.009500	0.002918
mse β	0.010605	0.012582	0.009568	0.001503	0.004040	0.003652	0.002273	0.014845	0.000050
mae β	0.069723	0.077076	0.067698	0.018982	0.041769	0.029291	0.014573	0.064207	0.003939
mean σ	0.248488	0.222641	0.223980	0.323688	0.283371	0.372535	0.363399	0.298043	0.405760
bias σ	-0.114412	-0.140259	-0.138920	-0.039212	-0.079529	0.009635	0.000499	-0.064857	0.042860
mse σ	0.036861	0.047877	0.042372	0.007156	0.016915	0.009111	0.004789	0.025874	0.003402
mae σ	0.159842	0.184737	0.173584	0.065210	0.103195	0.076377	0.049515	0.123514	0.048524

TABLE 28. Level Outlier, Subset of Instruments - α=-0.736 β=.9 σ=.3629

	GMM2S	GMMITER	GMMCUE	GEL	ET	ETEL	SGEL	SET	SETEL
mean α	-0.335845	-0.330789	-0.386053	-0.370569	-0.360749	-0.366858	-0.372957	-0.362454	-0.367322
bias α	0.032155	0.037211	-0.018053	-0.002569	0.007251	0.001142	-0.004957	0.005546	0.000678
mse α	0.393126	0.447313	0.005856	0.000258	0.000216	0.000272	0.000349	0.000100	0.000117
mae α	0.346897	0.376404	0.033920	0.009525	0.010960	0.010322	0.012304	0.007358	0.006922
mean β	0.954821	0.955542	0.948156	0.950982	0.951276	0.950536	0.951031	0.950969	0.950638
bias β	0.004821	0.005542	-0.001844	0.000982	0.001276	0.000536	0.001031	0.000969	0.000638
mse β	0.007061	0.008069	0.000095	0.000009	0.000006	0.000008	0.000010	0.000003	0.000005
mae β	0.046834	0.050766	0.004224	0.002083	0.001983	0.001981	0.002292	0.001400	0.001500
mean σ	0.163024	0.146836	0.230379	0.288998	0.270922	0.285684	0.288388	0.284835	0.293978
bias σ	-0.096976	-0.113164	-0.029621	0.028998	0.010922	0.025684	0.028388	0.024835	0.033978
mse σ	0.028388	0.034878	0.002244	0.001811	0.001298	0.001837	0.001679	0.001516	0.002272
mae σ	0.143643	0.161651	0.031444	0.032089	0.028832	0.032401	0.031920	0.031852	0.036552

TABLE 29. Level Outlier, Subset of Instruments - α=-0.368 β=.95 σ=.26

	GMM2S	GMMITER	GMMCUE	GEL	ET	ETEL	SGEL	SET	SETEL
mean α	-0.188493	-0.208328	-0.132164	-0.159277	-0.163758	-0.151777	-0.158373	-0.151709	-0.154944
bias α	-0.041293	-0.061128	0.015036	-0.012077	-0.016558	-0.004577	-0.011173	-0.004509	-0.007744
mse α	0.277367	0.463592	0.002377	0.000597	0.000662	0.000095	0.000625	0.000122	0.000265
mae α	0.217575	0.248036	0.026094	0.016203	0.017784	0.005898	0.015817	0.006516	0.010448
mean β	0.974756	0.972189	0.982451	0.979374	0.977933	0.979646	0.979411	0.979550	0.979339
bias β	-0.005244	-0.007811	0.002451	-0.000626	-0.002067	-0.000354	-0.000589	-0.000450	-0.000661
mse β	0.004900	0.008202	0.000043	0.000007	0.000013	0.000003	0.000008	0.000003	0.000004
mae β	0.029199	0.033214	0.003226	0.001859	0.002529	0.001038	0.001722	0.001228	0.001289
mean σ	0.100210	0.092477	0.140692	0.196440	0.181443	0.192532	0.189730	0.189435	0.196667
bias σ	-0.065490	-0.073223	-0.025008	0.030740	0.015743	0.026832	0.024030	0.023735	0.030967
mse σ	0.017358	0.020779	0.002875	0.001913	0.001262	0.001657	0.001677	0.001469	0.002163
mae σ	0.115406	0.125359	0.025701	0.033128	0.027601	0.030217	0.027776	0.029240	0.034699

TABLE 30. Level Outlier, Subset of Instruments - α-.1472 β=.98 σ=.1657

SV Model - Level Outlier - Subset of Instruments.

	GMM2S	GMMITER	GMMCUE	GEL	ET	ETEL	SGEL	SET	SETEL
mean α	-0.706232	-0.637902	-0.592404	-0.709767	-0.704501	-0.725353	-0.730109	-0.734256	-0.731349
bias α	0.029768	0.098098	0.143596	0.026233	0.031499	0.010647	0.005891	0.001744	0.004651
mse α	0.465434	0.924587	0.469655	0.015619	0.119894	0.004171	0.001127	0.017193	0.000166
mae α	0.454061	0.648185	0.466139	0.043156	0.177426	0.021347	0.011337	0.036012	0.009249
mean β	0.904874	0.914112	0.920234	0.903850	0.903874	0.903045	0.900937	0.900494	0.900896
bias β	0.004874	0.014112	0.020234	0.003850	0.003874	0.003045	0.000937	0.000494	0.000896
mse β	0.008459	0.016728	0.008446	0.000292	0.002275	0.000089	0.000022	0.000323	0.000005
mae β	0.061305	0.087544	0.063305	0.006366	0.024527	0.004016	0.001827	0.005275	0.001637
mean σ	0.256169	0.177301	0.185598	0.363270	0.300050	0.393189	0.400320	0.377173	0.396990
bias σ	-0.106731	-0.185599	-0.177302	0.000370	-0.062850	0.030289	0.037420	0.014273	0.034090
mse σ	0.033791	0.059072	0.049032	0.005121	0.015349	0.003025	0.001967	0.004401	0.001776
mae σ	0.153312	0.214148	0.195515	0.055298	0.097185	0.045333	0.040743	0.051957	0.038353

TABLE 31. Volatility Outlier - α=-0.736 β=.9 σ=.3629

	GMM2S	GMMITER	GMMCUE	GEL	ET	ETEL	SGEL	SET	SETEL
mean α	-0.351970	-0.352200	-0.508979	-0.370434	-0.365920	-0.368212	-0.367283	-0.365273	-0.367459
bias α	0.016030	0.015800	-0.140979	-0.002434	0.002080	-0.000212	0.000717	0.002727	0.000541
mse α	0.334339	0.610981	0.030704	0.001105	0.000219	0.000269	0.000495	0.000259	0.001637
mae α	0.329172	0.438042	0.144810	0.014213	0.011012	0.012832	0.012775	0.011757	0.013554
mean β	0.952649	0.952689	0.931288	0.949823	0.950035	0.949864	0.949897	0.950127	0.949848
bias β	0.002649	0.002689	-0.018712	-0.000177	0.000035	-0.000136	-0.000103	0.000127	-0.000152
mse β	0.005887	0.010844	0.000560	0.000017	0.000003	0.000005	0.000008	0.000004	0.000027
mae β	0.044410	0.059013	0.019272	0.001791	0.001296	0.001651	0.001511	0.001290	0.001593
mean σ	0.158805	0.108873	0.202559	0.269724	0.274544	0.267376	0.265636	0.273019	0.265015
bias σ	-0.101195	-0.151127	-0.057441	0.009724	0.014544	0.007376	0.005636	0.013019	0.005015
mse σ	0.026434	0.040086	0.004335	0.001910	0.001650	0.001972	0.002104	0.001773	0.002085
mae σ	0.138501	0.180479	0.057666	0.033782	0.034282	0.035755	0.037771	0.035826	0.037569

TABLE 32. Volatility Outlier - α=-0.368 β=.95 σ=.26

	GMM2S	GMMITER	GMMCUE	GEL	ET	ETEL	SGEL	SET	SETEL
mean α	-0.148088	-0.156272	-0.153073	-0.175806	-0.168742	-0.169571	-0.177988	-0.171109	-0.173611
bias α	-0.000888	-0.009072	-0.005873	-0.028606	-0.021542	-0.022371	-0.030788	-0.023909	-0.026411
mse α	0.097686	0.181002	0.001995	0.001490	0.001060	0.001655	0.002637	0.001731	0.002607
mae α	0.179616	0.224732	0.021219	0.029802	0.022380	0.023438	0.034092	0.024905	0.027865
mean β	0.979896	0.978821	0.980121	0.976108	0.976763	0.976686	0.975623	0.976401	0.976064
bias β	-0.000104	-0.001179	0.000121	-0.003892	-0.003237	-0.003314	-0.004377	-0.003599	-0.003936
mse β	0.001817	0.003372	0.000031	0.000025	0.000027	0.000042	0.000047	0.000045	0.000061
mae β	0.024394	0.030485	0.002028	0.004047	0.003334	0.003450	0.004797	0.003721	0.004109
mean σ	0.083003	0.058572	0.150466	0.174770	0.176514	0.168816	0.170293	0.175436	0.167197
bias σ	-0.082697	-0.107128	-0.015234	0.009070	0.010814	0.003116	0.004593	0.009736	0.001497
mse σ	0.015647	0.020630	0.001414	0.001570	0.001758	0.001881	0.001829	0.001875	0.001901
mae σ	0.112508	0.132708	0.015302	0.031762	0.032617	0.032367	0.034844	0.033013	0.033659

TABLE 33. Volatility Outlier - α-.1472 β=.98 σ=.1657

SV Model - Volatility Outlier.

	GMM2S	GMMITER	GMMCUE	GEL	ET	ETEL	SGEL	SET	SETEL
mean α	-0.380933	-0.319525	-0.346457	-0.675964	-0.533546	-0.689775	-0.722365	-0.629596	-0.730389
bias α	0.355067	0.416475	0.389543	0.060036	0.202454	0.046225	0.013635	0.106404	0.005611
mse α	0.400762	0.522664	0.372611	0.053243	0.136697	0.052746	0.017866	0.201125	0.002570
mae α	0.507168	0.583628	0.506768	0.103944	0.263422	0.106389	0.039998	0.254014	0.016877
mean β	0.948491	0.956853	0.953213	0.908027	0.927169	0.907076	0.902300	0.914907	0.902072
bias β	0.048491	0.056853	0.053213	0.008027	0.027169	0.007076	0.002300	0.014907	0.002072
mse β	0.007406	0.009648	0.006851	0.000982	0.002507	0.001024	0.000334	0.003711	0.000057
mae β	0.068937	0.079200	0.068915	0.014429	0.035831	0.015276	0.005944	0.034862	0.003113
mean σ	0.183818	0.142762	0.161442	0.338956	0.248977	0.380012	0.374279	0.283110	0.396593
bias σ	-0.179082	-0.220138	-0.201458	-0.023944	-0.113923	0.017112	0.011379	-0.079790	0.033693
mse σ	0.052152	0.071179	0.060068	0.008752	0.025916	0.006950	0.004083	0.023850	0.002701
mae σ	0.200869	0.239498	0.218347	0.070212	0.132413	0.064921	0.048370	0.118879	0.043555

TABLE 34. Volatility Outlier, Subset of Instruments - α=-0.736 β=.9 σ=.3629

	GMM2S	GMMITER	GMMCUE	GEL	ET	ETEL	SGEL	SET	SETEL
mean α	-0.172707	-0.155808	-0.416848	-0.371238	-0.360159	-0.366675	-0.372505	-0.361854	-0.366928
bias α	0.195293	0.212192	-0.048848	-0.003238	0.007841	0.001325	-0.004505	0.006146	0.001072
mse α	0.205893	0.235857	0.008678	0.000264	0.000227	0.000177	0.000371	0.000112	0.000124
mae α	0.315813	0.339617	0.063105	0.010417	0.011081	0.008793	0.012457	0.007730	0.007175
mean β	0.976594	0.978933	0.943591	0.950694	0.951035	0.950248	0.950705	0.950765	0.950356
bias β	0.026594	0.028933	-0.006409	0.000694	0.001035	0.000248	0.000705	0.000765	0.000356
mse β	0.003823	0.004361	0.000162	0.000008	0.000005	0.000005	0.000009	0.000003	0.000005
mae β	0.042912	0.046096	0.008475	0.002108	0.001795	0.001665	0.002122	0.001287	0.001475
mean σ	0.104251	0.082372	0.217374	0.293798	0.272302	0.287402	0.288896	0.283725	0.295915
bias σ	-0.155749	-0.177628	-0.042626	0.033798	0.012302	0.027402	0.028896	0.023725	0.035915
mse σ	0.036282	0.044006	0.003188	0.002233	0.001357	0.001832	0.001810	0.001541	0.002494
mae σ	0.172382	0.192864	0.044066	0.036650	0.029614	0.032069	0.032802	0.032286	0.038483

TABLE 35. Volatility Outlier, Subset of Instruments - α=-0.368 β=.95 σ=.26

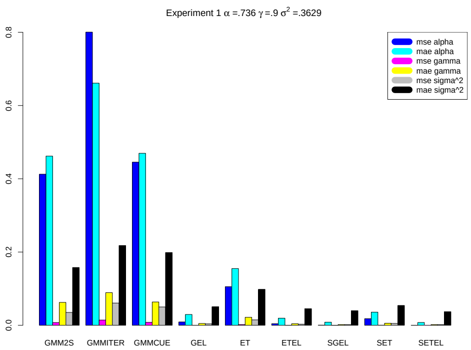
	GMM2S	GMMITER	GMMCUE	GEL	ET	ETEL	SGEL	SET	SETEL
mean α	-0.087077	-0.083241	-0.151294	-0.158665	-0.161511	-0.153171	-0.158281	-0.152353	-0.153574
bias α	0.060123	0.063959	-0.004094	-0.011465	-0.014311	-0.005971	-0.011082	-0.005153	-0.006374
mse α	0.113590	0.127431	0.014150	0.000588	0.000775	0.000351	0.000671	0.000194	0.000245
mae α	0.162835	0.170595	0.028076	0.015739	0.016077	0.007695	0.014863	0.006753	0.009692
mean β	0.988190	0.988723	0.980057	0.979187	0.977970	0.979248	0.979189	0.979258	0.979384
bias β	0.008190	0.008723	0.000057	-0.000813	-0.002030	-0.000752	-0.000811	-0.000742	-0.000616
mse β	0.002128	0.002384	0.000221	0.000010	0.000019	0.000009	0.000010	0.000005	0.000004
mae β	0.022082	0.023135	0.003205	0.001995	0.002513	0.001365	0.001746	0.001348	0.001257
mean σ	0.057910	0.048059	0.148996	0.197172	0.182484	0.192340	0.190732	0.192244	0.199407
bias σ	-0.107790	-0.117641	-0.016704	0.031472	0.016784	0.026640	0.025032	0.026544	0.033707
mse σ	0.018145	0.020719	0.001930	0.002109	0.001506	0.001826	0.001733	0.001906	0.002436
mae σ	0.125675	0.135012	0.018445	0.034274	0.029489	0.030313	0.028883	0.032819	0.037243

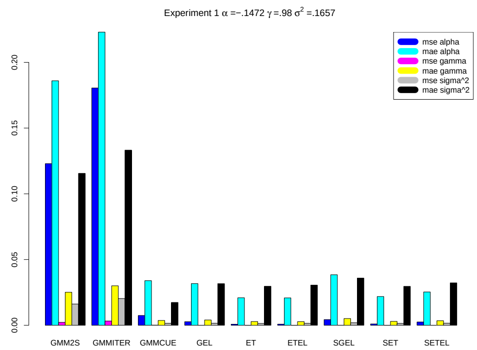
TABLE 36. Volatility Outlier, Subset of Instruments - α-.1472 β=.98 σ=.1657

SV Model - Volatility Outlier - Subset of Instruments.

Figures

Figure 1. MSE and MAE of the estimation of the reference models with sample size 500 and 24 moment conditions.





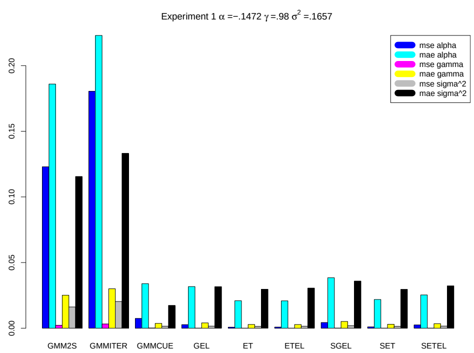


FIGURE 2. Relative Efficiency in the reference models - Effect of sample size - (MSE sample size 250 /MSE sample size 1000). Estimation based on 24 moment conditions

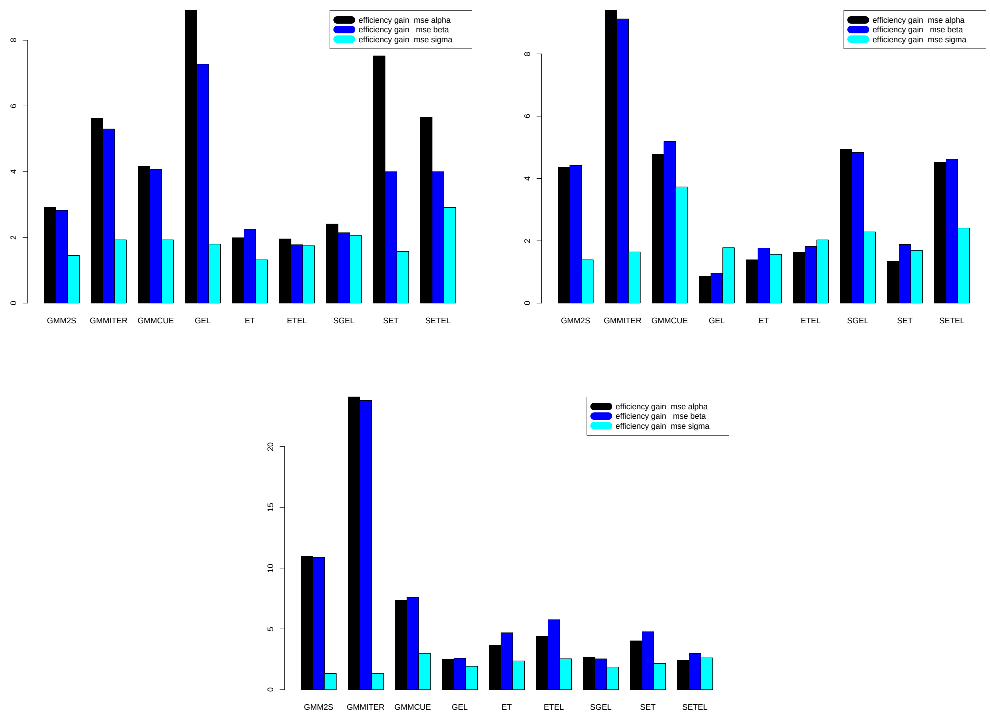


FIGURE 3. MSE and MAE of the estimation of the reference models with sample size 500 and 14 moment conditions.

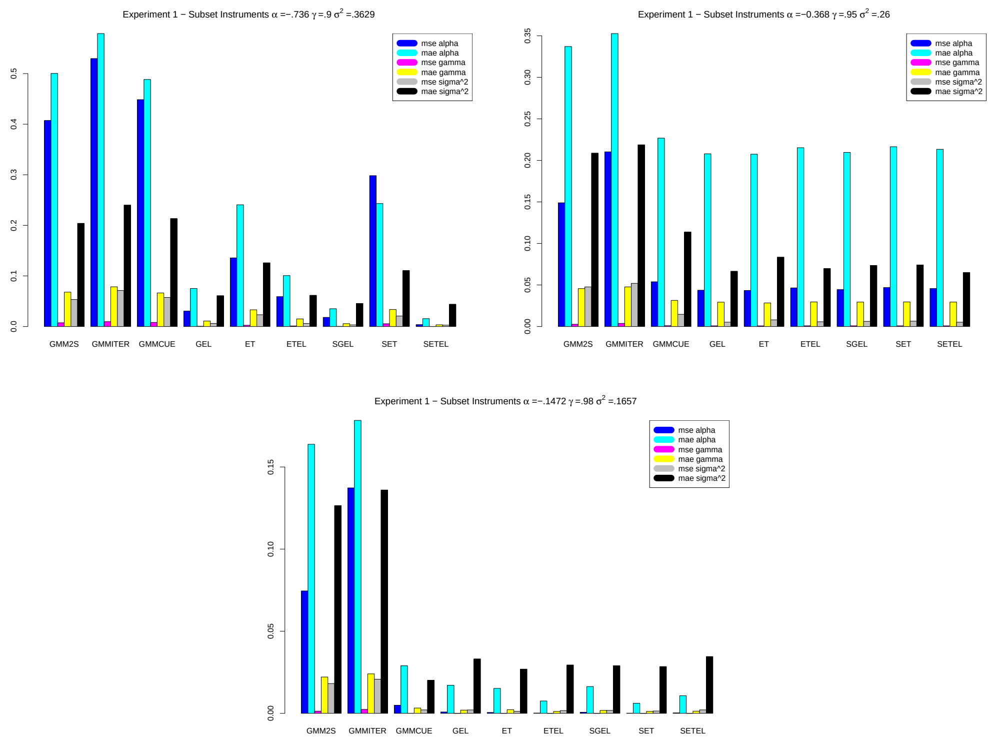


FIGURE 4. Relative Efficiency in the reference models - Effect of number of moment conditions - (MSE 24 moment conditions /MSE 24 moment conditions). Sample size 500.

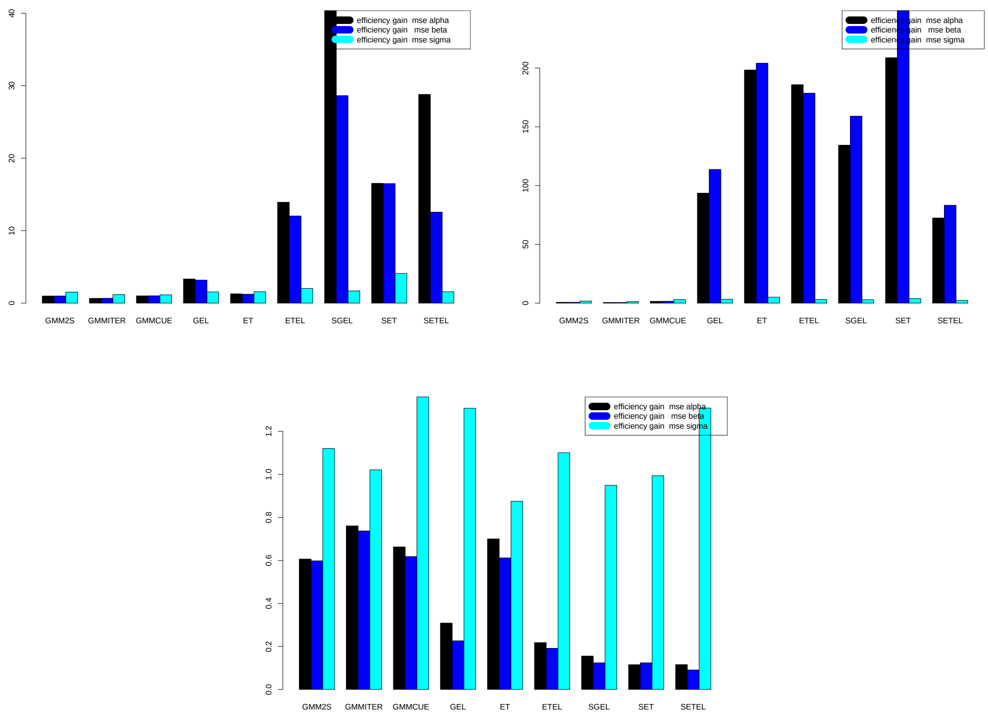


FIGURE 5. MSE and MAE of the estimation of the reference models, modified with Student-t with 4 d.f. innovation in the mean equation. Sample size 500 and 24 moment conditions.

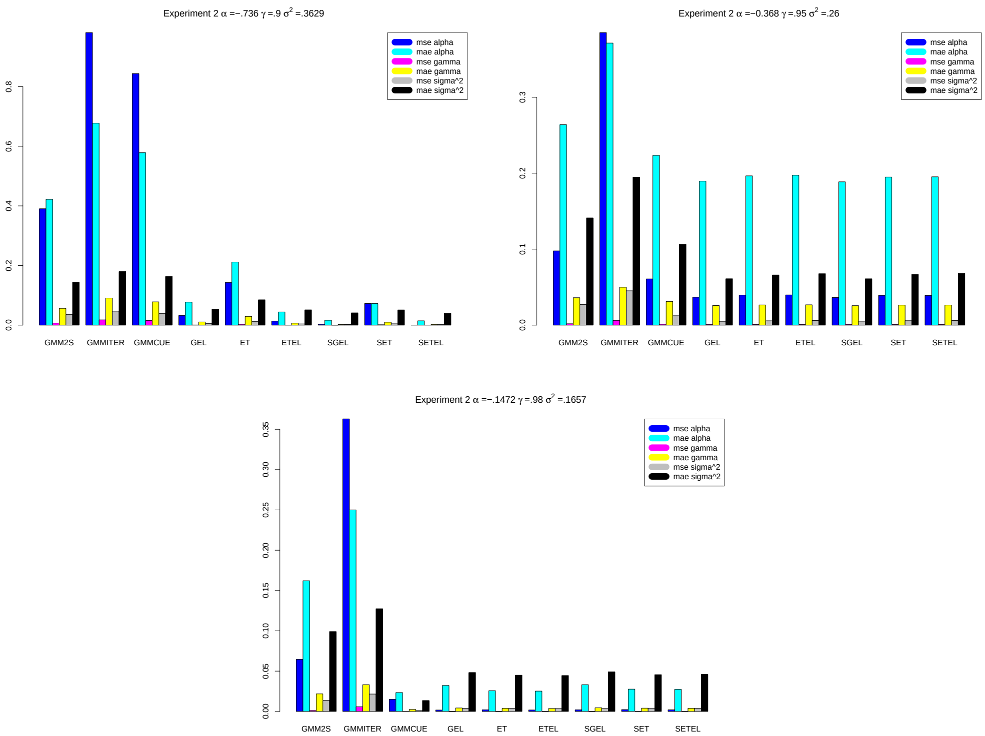


FIGURE 6. MSE and MAE of the estimation of the reference models, modified with Student-t with 4 d.f. innovation in the mean equation. Sample size 500 and 14 moment conditions.

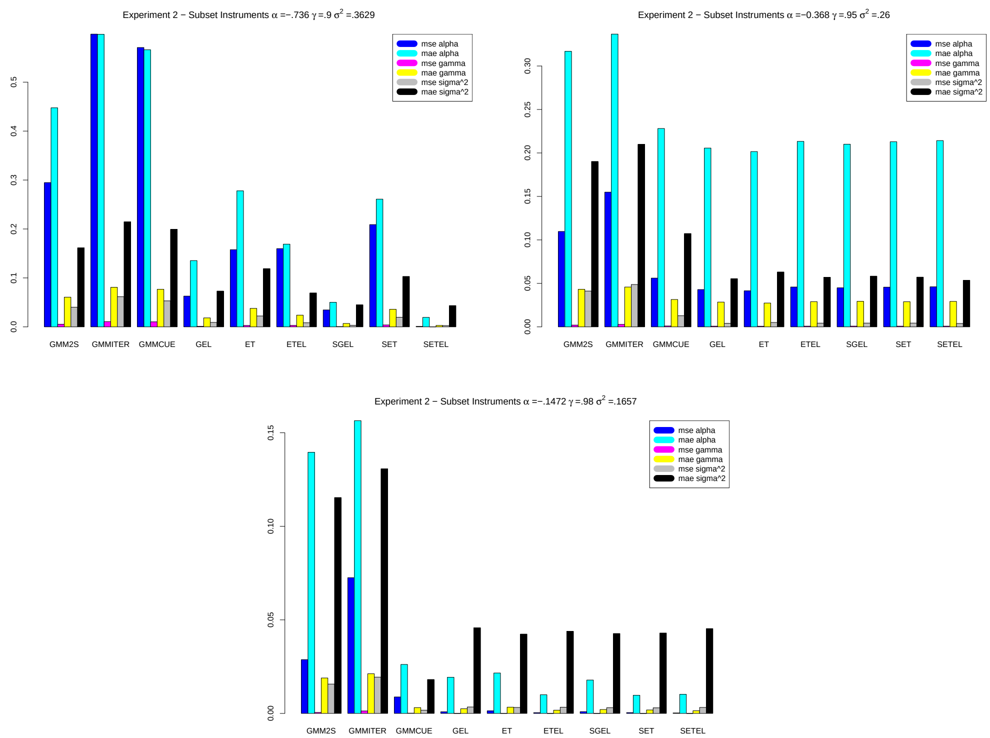


FIGURE 7. Relative Efficiency in the reference models modified with Student-t with 4 d.f innovation in the mean equation - Effect of number of moment conditions - (MSE 14 moment conditions /MSE 24 moment conditions). Sample size 500.

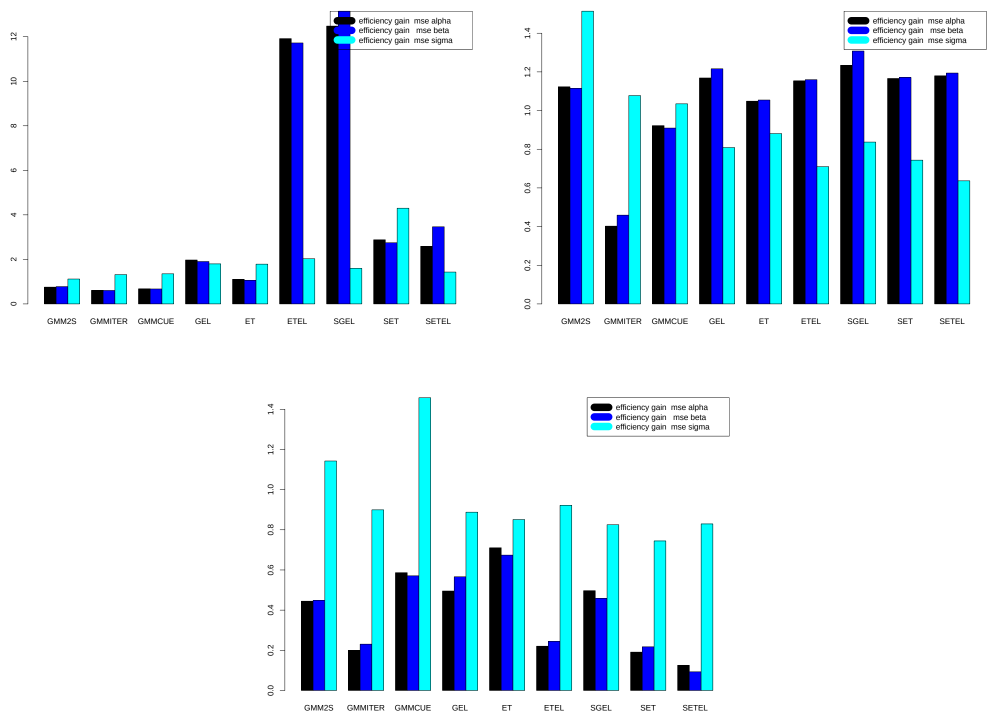
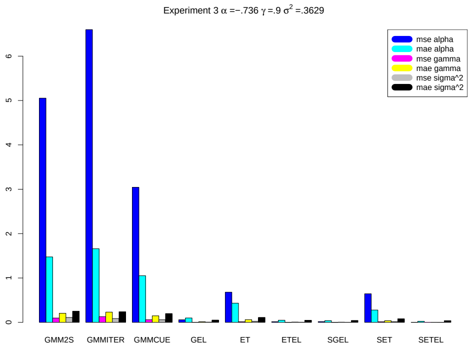
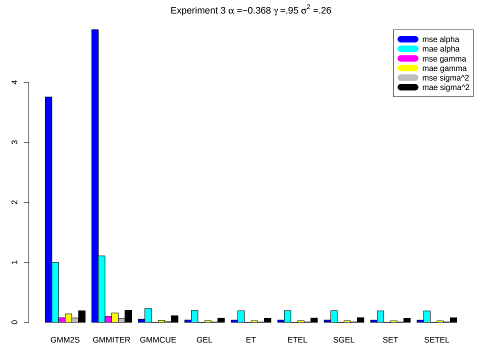


FIGURE 8. MSE and MAE of the estimation of the reference models, modified with Student-t with 4 d.f. innovation in the volatility equation. Sample size 500 and 24 moment conditions.





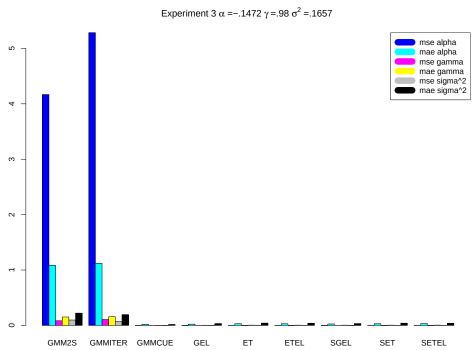


Figure 9. MSE and MAE of the estimation of the reference models, modified with Student-t with 4 d.f. innovation in the volatility equation. Sample size 500 and 14 moment conditions.

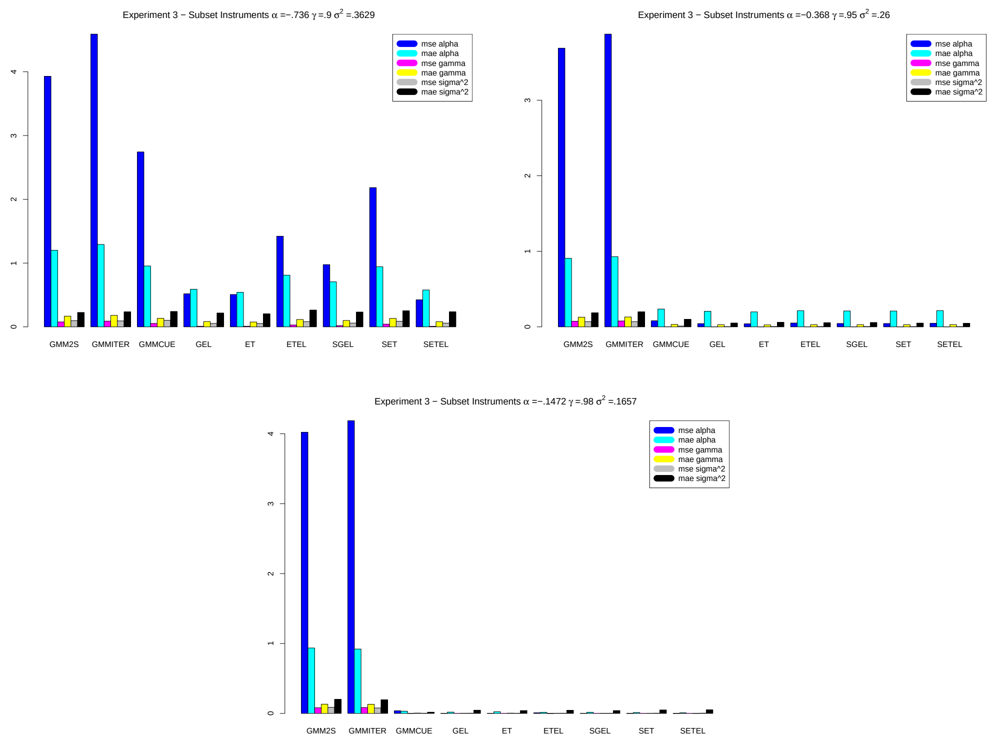


Figure 10. Relative Efficiency in the reference models modified with Student-t with 4 d.f innovation in the volatility equation - Effect of number of moment conditions - (MSE 14 moment conditions /MSE 24 moment conditions). Sample size 500.

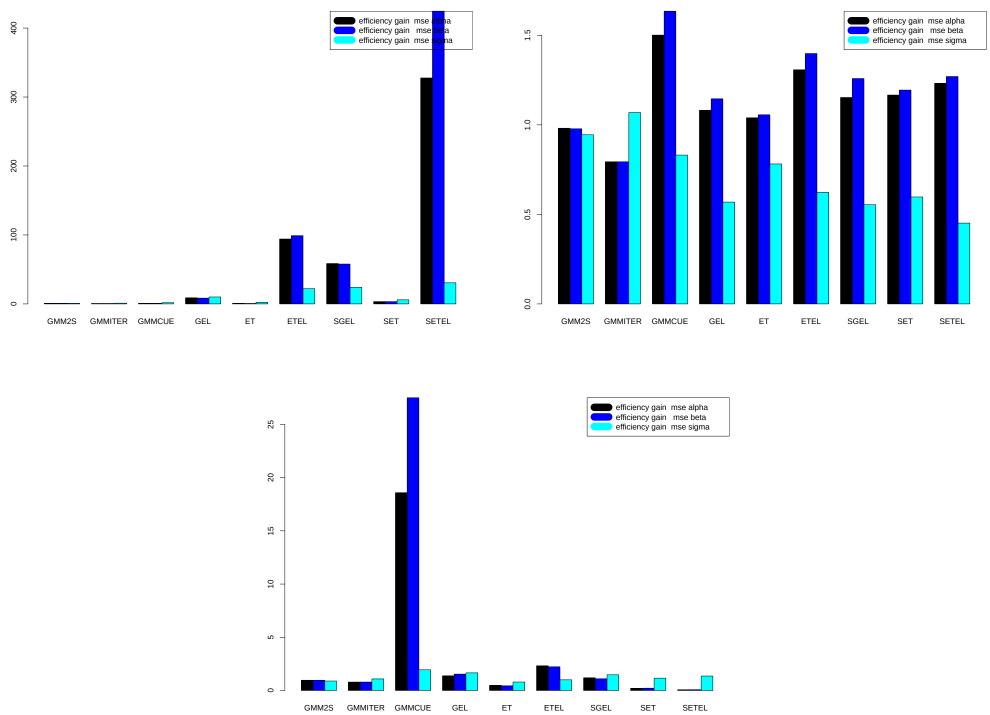


FIGURE 11. MSE and MAE of the estimation of the reference models, modified with Level Outlier. Sample size 500 and 24 moment conditions.

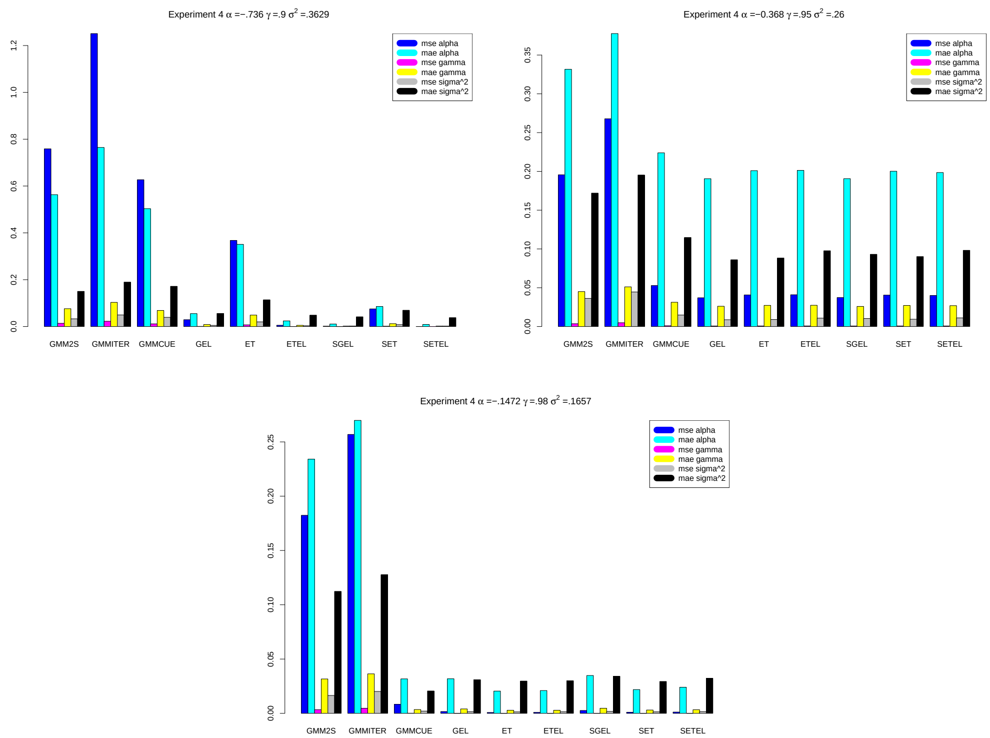


FIGURE 12. MSE and MAE of the estimation of the reference models, modified with Level Outlier. Sample size 500 and 14 moment conditions.

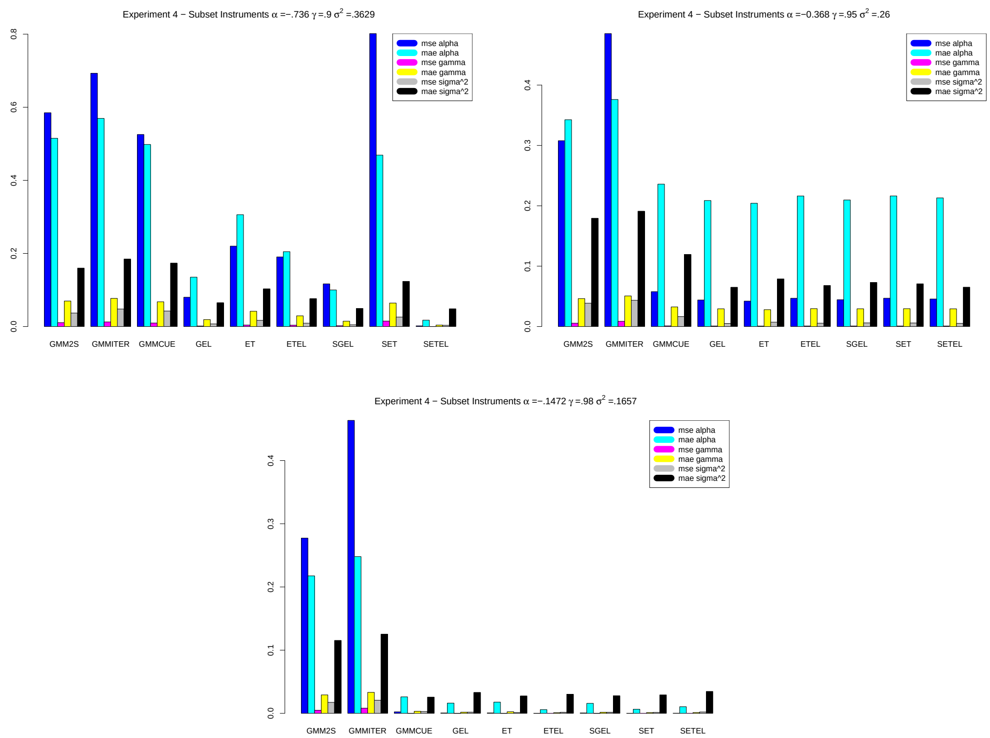


Figure 13. Relative Efficiency in the reference models with level outlier - Effect of number of moment conditions - (MSE 14 moment conditions /MSE 24 moment conditions). Sample size 500.

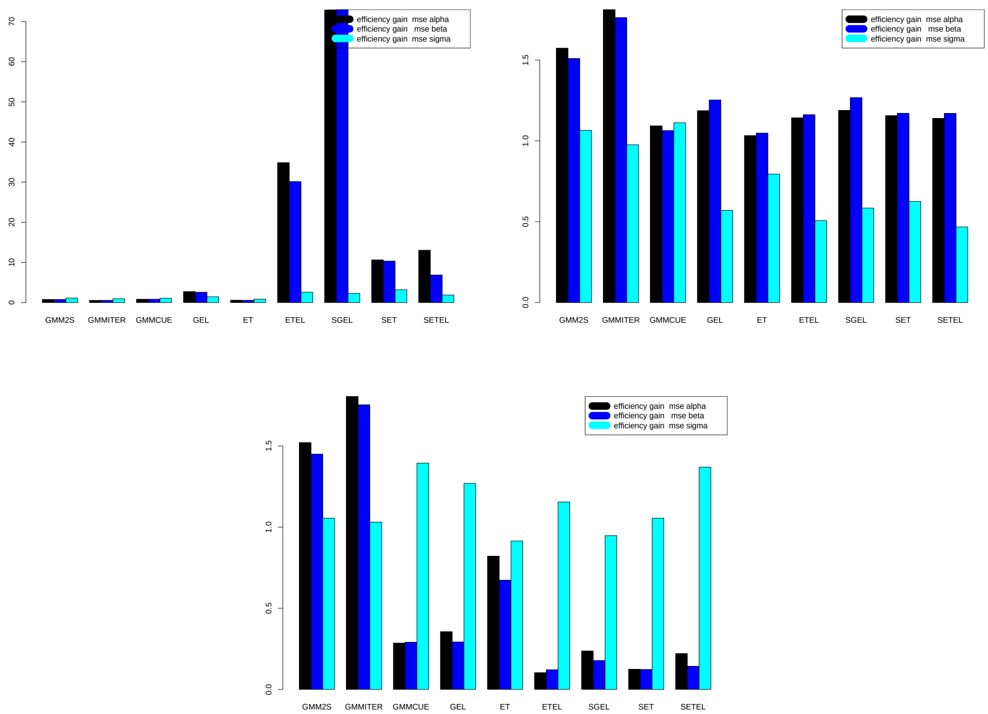


FIGURE 14. MSE and MAE of the estimation of the reference models modified with volatility outlier. Sample size 500 and 24 moment conditions.

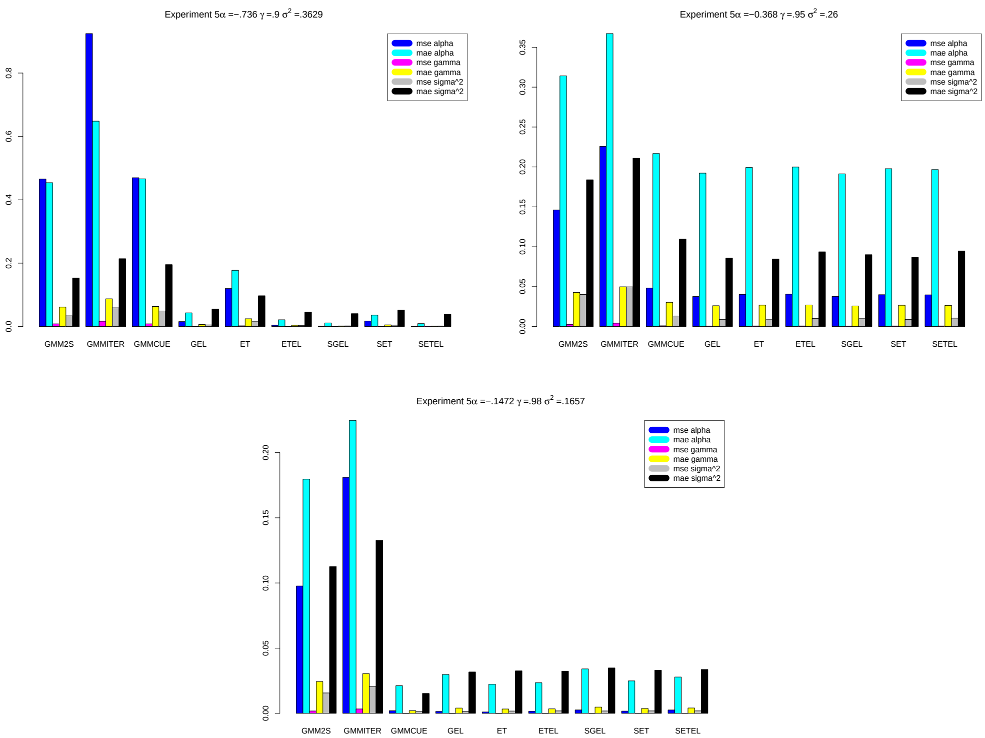


Figure 15. MSE and MAE of the estimation of the reference models modified with volatility outlier. Sample size 500 and 14 moment conditions.

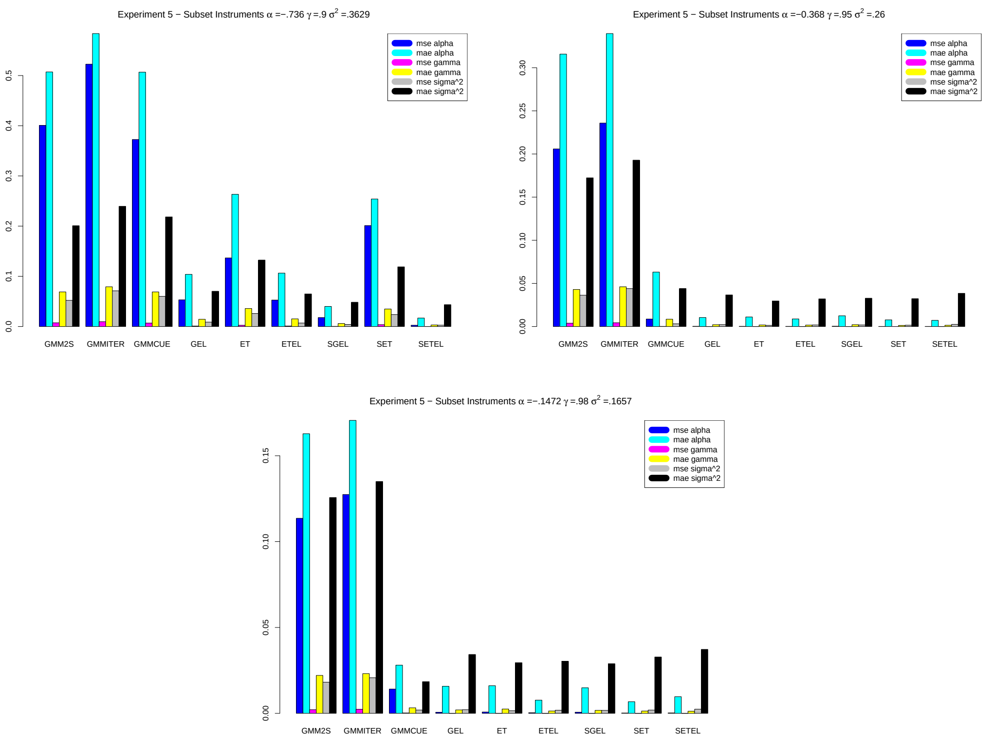


FIGURE 16. Relative Efficiency in the reference models modified with volatility outlier - Effect of number of moment conditions - (MSE 14 moment conditions /MSE 24 moment conditions.) Sample size 500.