Recent Developments in Recursive Estimation for Time Series Models

You Liang[1], A. Thavaneswaran[1], & B. Abraham[2]

[1] Department of Statistics, University of Manitoba, Winnipeg, Canada

[2] Department of Statistics University of Waterloo, Waterloo, Canada

Correspondence: A. Thavaneswaran, Department of Statistics, University of Manitoba, Winnipeg, MB, R3T 2N2, Canada. Tel: 1-204-474-8984. E-mail: athava23@gmail.com

Abstract

Recently there has been a growing interest in joint estimation of the location and scale parameters using combined estimation functions. Combined estimating functions had been studied in Liang et al. (2011) for models with finite variance errors and in Thavaneswaran et al. (2013) for models with infinite variance stable errors. In this paper, first a theorem on recursive estimation based on estimating functions is extended to multi-parameter setup and it is shown that the unified approach can be used to estimate the location parameter recursively for models with finite variance/infinite variance errors. The method is applied for the joint estimation of the location and scale parameters for regression models with ARCH errors and RCA models with GARCH errors.

Keywords: quadratic estimating functions, information, recursive estimation, joint estimation, ARCH, GARCH

1. Introduction

Estimating function theory is well suited to financial data (see Bera et al. (2006)). Recently, Ghahramani and Thavaneswaran (2009, 2011) have studied GARCH model identification and recursive estimation by combining least squares and least absolute deviation estimating functions and the method has been applied to identify several financial time series models. Combined estimating functions had also been studied in Liang et al. (2011) and in Thavaneswaran et al. (2013,2015) for the multi-parameter setup. Thavaneswaran and Ravishanker (2015) studied recursive estimation for circular time series models using estimating functions. In this paper, first the combined estimating function method is applied to obtain joint optimal recursive estimates of the parameters in autoregressive models with t-distributed errors, then regression models with ARCH errors and RCA models with GARCH errors. Combinations of least squares and quadratic estimating function, as well as combinations of least squares and LAD estimating functions, are considered. The following example motivates the use of estimating function theory for recursive estimation of the parameter in certain time series models with stable errors.

Consider an AR(1) process

$$y_t = \phi y_{t-1} + \varepsilon_t,$$

where $\{\varepsilon_t\}$ is an i.i.d. sequence of symmetric stable random variables with characteristic function $c(u) = \exp\left(-|u|^\lambda\right)$ where $0 \leq \lambda \leq 2$. Closed form representations of the density exist only when $\lambda = 1$ (ε_t follows a Cauchy distribution) or when $\lambda = 2$ (ε_t follows a Gaussian distribution). Moreover, the second moments are not finite when $0 \leq \lambda < 2$. Interest centers on estimating the parameter ϕ based on the observations $y_1, \cdots, y_n$. Merkouris (2007) recently has proposed the estimating function approach to estimate the parameter ϕ. Recently in Thavaneswaran et al. (2013) joint estimates of location and scale parameters are derived for a class of autoregressive (AR) models, a class of Random Coefficient Autoregressive (RCA) models with stable errors, as well as for a class of AR models with stable Autoregressive Conditionally Heteroscedastic (ARCH) errors. Moreover in Thavaneswaran et al. (2013) a fast, on-line, recursive parametric estimation for the location parameter based on transformed estimating functions is discussed using simulation studies, and a real financial time series is also discussed in some detail. Recursive (or online) estimation of a parameter where the estimate of the parameter at time $t + 1$ is the estimate of parameter at time t plus an adjustment is advantageous when there is a large stretch of data and observations become available successively over time. Recursive estimation of the parameter based on nonlinear estimating functions had been studied by Thavaneswaran and Heyde (1999) and based on combined

estimating function in Ghahramani and Thavaneswaran (2012).

This paper is organized as follows. In section 2, recursive estimation of location parameters using nonlinear estimating functions is discussed. and in section 3 recursive estimation of location and scale parameters are considered.

2. Recursive Estimation using Nonlinear Estimating Functions

Suppose that $\{\mathbf{y}_t, t = 1, \ldots, n\}$ is a realization of a discrete-time stochastic process and its distribution depends on a vector parameter $\boldsymbol{\theta}$ belonging to an open subset $\boldsymbol{\Theta}$ of the p-dimensional Euclidean space. Let $(\Omega, \mathcal{F}, P_{\boldsymbol{\theta}})$ denote the underlying probability space, and let $\mathcal{F}_t^{\mathbf{y}}$ be the σ-field generated by $\{\mathbf{y}_1, \ldots, \mathbf{y}_t, t \geq 1\}$. Let $\mathbf{h}_t(\boldsymbol{\theta}) = \mathbf{h}_t(\mathbf{y}_1, \ldots, \mathbf{y}_t, \boldsymbol{\theta}), 1 \leq t \leq n$ be specified q-dimensional vectors that are martingale differences. We consider the class $\mathcal{M}$ of zero mean and square integrable p-dimensional martingale estimating functions of the form

$$\mathcal{M} = \left\{ \mathbf{g}_n(\boldsymbol{\theta}) : \mathbf{g}_n(\boldsymbol{\theta}) = \sum_{t=1}^{n} \mathbf{a}_{t-1}(\boldsymbol{\theta}) \mathbf{h}_t(\boldsymbol{\theta}) \right\},$$

where $\mathbf{a}_{t-1}(\boldsymbol{\theta})$ are $p \times q$ matrices depending on $\mathbf{y}_1, \ldots, \mathbf{y}_{t-1}, 1 \leq t \leq n$ and the parameter $\boldsymbol{\theta}$. The estimating functions $\mathbf{g}_n(\boldsymbol{\theta})$ are further assumed to be almost surely differentiable with respect to the components of $\boldsymbol{\theta}$ and such that $\mathrm{E}\left[\frac{\partial \mathbf{g}_n(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}} \Big| \mathcal{F}_{n-1}^{\mathbf{y}} \right]$ and $\mathrm{E}[\mathbf{g}_n(\boldsymbol{\theta}) \mathbf{g}_n(\boldsymbol{\theta})' | \mathcal{F}_{n-1}^{\mathbf{y}}]$ are nonsingular for all $\boldsymbol{\theta} \in \boldsymbol{\Theta}$ and for each $n \geq 1$. The expectations are always taken with respect to $P_{\boldsymbol{\theta}}$. Estimators of $\boldsymbol{\theta}$ can be obtained by solving the estimating equation $\mathbf{g}_n(\boldsymbol{\theta}) = \mathbf{0}$. Furthermore, the $p \times p$ matrix $\mathrm{E}[\mathbf{g}_n(\boldsymbol{\theta}) \mathbf{g}_n(\boldsymbol{\theta})' | \mathcal{F}_{n-1}^{\mathbf{y}}]$ is assumed to be positive definite for all $\boldsymbol{\theta} \in \boldsymbol{\Theta}$. Then in the class of all zero mean and square integrable martingale estimating functions $\mathcal{M}$, the optimal estimating function $\mathbf{g}_n^*(\boldsymbol{\theta})$ which maximizes, in the partial order of nonnegative definite matrices, the information matrix

$$\mathbf{I}_{\mathbf{g}_n}(\boldsymbol{\theta}) = \left(\sum_{t=1}^{n} \mathbf{a}_{t-1}(\boldsymbol{\theta}) \mathrm{E}\left[\frac{\partial \mathbf{h}_t(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}} \Big| \mathcal{F}_{t-1}^{\mathbf{y}} \right] \right)' \left(\sum_{t=1}^{n} \mathrm{E}[(\mathbf{a}_{t-1}(\boldsymbol{\theta}) \mathbf{h}_t(\boldsymbol{\theta}))(\mathbf{a}_{t-1}(\boldsymbol{\theta}) \mathbf{h}_t(\boldsymbol{\theta}))' | \mathcal{F}_{t-1}^{\mathbf{y}}] \right)^{-1}$$
$$\times \left(\sum_{t=1}^{n} \mathbf{a}_{t-1}(\boldsymbol{\theta}) \mathrm{E}\left[\frac{\partial \mathbf{h}_t(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}} \Big| \mathcal{F}_{t-1}^{\mathbf{y}} \right] \right)$$

is given by

$$\mathbf{g}_n^*(\boldsymbol{\theta}) = \sum_{t=1}^{n} \mathbf{a}_{t-1}^*(\boldsymbol{\theta}) \mathbf{h}_t(\boldsymbol{\theta}) = \sum_{t=1}^{n} \left(\mathrm{E}\left[\frac{\partial \mathbf{h}_t(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}} \Big| \mathcal{F}_{t-1}^{\mathbf{y}} \right] \right)' (\mathrm{E}[\mathbf{h}_t(\boldsymbol{\theta}) \mathbf{h}_t(\boldsymbol{\theta})' | \mathcal{F}_{t-1}^{\mathbf{y}}])^{-1} \mathbf{h}_t,$$

and the corresponding optimal information reduces to

$$\mathbf{I}_{\mathbf{g}_n^*}(\boldsymbol{\theta}) = \sum_{t=1}^{n} \left(\mathrm{E}\left[\frac{\partial \mathbf{h}_t(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}} \Big| \mathcal{F}_{t-1}^{\mathbf{y}} \right] \right)' \left(\mathrm{E}[\mathbf{h}_t(\boldsymbol{\theta}) \mathbf{h}_t(\boldsymbol{\theta})' | \mathcal{F}_{t-1}^{\mathbf{y}}] \right)^{-1} \left(\mathrm{E}\left[\frac{\partial \mathbf{h}_t(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}} \Big| \mathcal{F}_{t-1}^{\mathbf{y}} \right] \right).$$

For the estimating function of the form $G(\boldsymbol{\theta}, \mathcal{F}_{t-1}^{y}) = \sum_{t=1}^{n} \mathbf{b}_{t-1}(\boldsymbol{\theta}, \mathcal{F}_{t-1}^{y}) g(m_t(\boldsymbol{\theta}, \mathcal{F}_{t-1}^{y}))$ based on the martingale differences $m_t(\boldsymbol{\theta}, \mathcal{F}_{t-1}^{y})) = y_t - \mu_t(\boldsymbol{\theta}, \mathcal{F}_{t-1}^{y})$, the following Theorem provides the recursive form of the optimal estimator based on optimal estimating function in the multi-parameter case. The proof is similar to the theorem in Thavaneswaran and Heyde (1999) for the scalar parameter case.

Theorem 1. Let $\mu_t(\boldsymbol{\theta}, \mathcal{F}_{t-1}^{y}) = E[y_t | \mathcal{F}_{t-1}^{y}]$ be differentiable respect to $\boldsymbol{\theta}$. Then the recursive estimator for $\boldsymbol{\theta}$ based on the optimal estimating function $G(\boldsymbol{\theta}, \mathcal{F}_{t-1}^{y})$ is given by

$$\hat{\boldsymbol{\theta}}_t = \hat{\boldsymbol{\theta}}_{t-1} + \mathbf{K}_t \mathbf{b}_{t-1}^*(\hat{\boldsymbol{\theta}}_{t-1}, \mathcal{F}_{t-1}^{y}) g(m_t(\hat{\boldsymbol{\theta}}_{t-1}, \mathcal{F}_{t-1}^{y})),$$

$$\mathbf{K}_t = \mathbf{K}_{t-1} \left(\mathbf{I}_p - \left[\mathbf{b}_{t-1}^*(\hat{\boldsymbol{\theta}}_{t-1}, \mathcal{F}_{t-1}^{y}) \frac{\partial g(m_t(\hat{\boldsymbol{\theta}}_{t-1}, \mathcal{F}_{t-1}^{y}))}{\partial \boldsymbol{\theta}} \right.\right.$$
$$\left.\left. + \frac{\partial \mathbf{b}_{t-1}^*(\hat{\boldsymbol{\theta}}_{t-1}, \mathcal{F}_{t-1}^{y})}{\partial \boldsymbol{\theta}} g(m_t(\hat{\boldsymbol{\theta}}_{t-1}, \mathcal{F}_{t-1}^{y})) \right] \mathbf{K}_{t-1} \right)^{-1},$$

where $\mathbf{I}_p$ is the identity matrix and $\mathbf{b}_{t-1}^*$ is a function of $g, \boldsymbol{\theta}$ and the observations. If $g(x) = x$, then

$$\mathbf{b}_{t-1}^* = \frac{\partial \mu_t(\hat{\boldsymbol{\theta}}_{t-1}, \mathcal{F}_{t-1}^{y}) / \partial \boldsymbol{\theta}}{Var(g | \mathcal{F}_{t-1}^{y})},$$

while for any function g (e.g. if g is the score function, then)

$$\mathbf{b}_{t-1}^* = \frac{[\partial \mu_t(\hat{\boldsymbol{\theta}}_{t-1}, \mathcal{F}_{t-1}^y)/\partial \boldsymbol{\theta}][\partial g(m_t(\boldsymbol{\theta}, \mathcal{F}_{t-1}^y))/\partial m_t]}{Var(g|\mathcal{F}_{t-1}^y)}.$$

For the scalar parameter case where b_{t-1}^* dose not depend on θ, we have the following corollary as a special case.

Corollary 2. *In the class $\mathcal{G}$ of all unbiased estimating function $g(m_t(\theta, \mathcal{F}_{t-1}^y)))$ based on the martingale difference $m_t(\theta, \mathcal{F}_{t-1}^y)) = y_t - \mu_t(\theta, \mathcal{F}_{t-1}^y)$. The recursive estimator for θ based on the optimal estimating function is given by*

$$\hat{\theta}_t = \hat{\theta}_{t-1} + \frac{K_{t-1}b_{t-1}^*}{1 + [\partial \mu_t(\hat{\theta}_{t-1}, \mathcal{F}_{t-1}^y)/\partial \theta]K_{t-1}b_{t-1}^*} g(m_t(\theta, \mathcal{F}_{t-1}^y))), \qquad (1)$$

$$K_t = \frac{K_{t-1}}{1 + [\partial \mu_t(\hat{\theta}_{t-1}, \mathcal{F}_{t-1}^y)/\partial \theta]K_{t-1}b_{t-1}^*}, \qquad (2)$$

where b_{t-1}^* is a function of the observations. If $g(x) = x$, then

$$b_{t-1}^* = \frac{\partial \mu_t(\hat{\theta}_{t-1}, \mathcal{F}_{t-1}^y)/\partial \theta}{Var(g|\mathcal{F}_{t-1}^y)},$$

while for any function g (e.g. if g is the score function, then)

$$b_{t-1}^* = \frac{[\partial \mu_t(\hat{\theta}_{t-1}, \mathcal{F}_{t-1}^y)/\partial \theta][\partial g(m_t(\theta, \mathcal{F}_{t-1}^y))/\partial m_t]}{Var(g|\mathcal{F}_{t-1}^y)}.$$

Corollary 3 (Thavaneswaran and Abraham (1988)). *For the nonlinear time series models of the form*

$$y_t = \theta f(t - 1, y) + \varepsilon_t,$$

where ε_t is an uncorrelated sequence with mean zero and variance σ_ε^2, the recursive estimate based on the optimal linear estimating function $\sum_{t=1}^n a_{t-1}(y_t - \theta f(t - 1, y))$ is given by

$$\hat{\theta}_t = \hat{\theta}_{t-1} + \frac{K_{t-1}a_{t-1}^*}{1 + f(t - 1, y)K_{t-1}a_{t-1}^*},$$

where

$$K_t = \frac{K_{t-1}}{1 + f(t - 1, y)K_{t-1}a_{t-1}^*},$$

and $a_{t-1}^* = -f(t - 1, y)/\sigma_\varepsilon^2$.

We describe recursive estimation of $\boldsymbol{\phi}$ for known scale c. Based on $g_n^*(\theta)$ in Thavaneswaran et al.(2013), the optimal estimate is obtained by solving the estimating equation

$$k(u, \lambda) \sum_{t=p+1}^n \mathbf{y}_{t-1} \sin\left[u\left(\frac{y_t - \boldsymbol{\phi}'\mathbf{y}_{t-1}}{c}\right)\right] = 0,$$

where $k(u, \lambda) = \frac{2ue^{-|u|^\lambda}}{c(1 - e^{-|2u|^\lambda})}$. By using vector form of the recursive algorithm (1) and (2) and letting $g = \sin\left[u\left(\frac{y_t - \boldsymbol{\phi}'\mathbf{y}_{t-1}}{c}\right)\right]$ and $\mathbf{b}_{t-1}^* = k(u, \lambda)\mathbf{y}_{t-1}$, it is easy to show that the recursive equations for $\boldsymbol{\phi}$ become

$$\mathbf{K}_t = \left\{\mathbf{I}_p + \frac{uk(u, \lambda)\mathbf{y}_{t-1}\mathbf{y}'_{t-1}}{c} \cos\left[u\left(\frac{y_t - \hat{\boldsymbol{\phi}}'_{t-1}\mathbf{y}_{t-1}}{c}\right)\right]\mathbf{K}_{t-1}\right\}^{-1} \mathbf{K}_{t-1}, \qquad (3)$$

where $\mathbf{I}_p$ is the $p \times p$ identity matrix, and

$$\hat{\boldsymbol{\phi}}_t = \hat{\boldsymbol{\phi}}_{t-1} + \mathbf{K}_t k(u, \lambda)\mathbf{y}_{t-1} \sin\left[u\left(\frac{y_t - \hat{\boldsymbol{\phi}}'_{t-1}\mathbf{y}_{t-1}}{c}\right)\right]. \qquad (4)$$

Recently recursive estimation for *location parameters* has been studied and applied to real data for infinite variance stable processes (Thavaneswaran et al. (2013)). Now we show that the results can be obtained as a direct corollary to Theorem 1. Consider the RCA(p) process was defined as $y_t = (\boldsymbol{\phi}' + \mathbf{b}_t')\mathbf{y}_{t-1} + \varepsilon_t$, where $\boldsymbol{\phi} = (\phi_1, \cdots, \phi_p)'$, and $\mathbf{b}_t = (b_{t,1}, \cdots, b_{t,p})'$. Assume that $\{\varepsilon_t\}$ is a sequence of i.i.d λ-stable random variables with location parameter zero and known scale parameter c. Assume that the components of $\{\mathbf{b}_t\}$ are mutually independent, and are independent of $\{\varepsilon_t\}$. Assume that $\{b_{t,j}\}$ is a sequence of i.i.d λ-stable random variables with location parameter zero and known scale parameter $\beta_j = c_{b,j}$. Let $\boldsymbol{\beta} = (\beta_1, \cdots, \beta_p)' = (c_{b,1}, \cdots, c_{b,p})'$. Then, the distribution of y_t, conditional on the past, follows a stable distribution with location parameter $\boldsymbol{\phi}'\mathbf{y}_{t-1}$ and scale parameter $c_t(\boldsymbol{\phi}, \boldsymbol{\beta})$, where variation of the the RCA process $|c_t|^\lambda = |\boldsymbol{\beta}'\mathbf{y}_{t-1}|^\lambda + |c|^\lambda$.

Based on martingale difference $\sin\left[u_1\left(\frac{y_t - \boldsymbol{\phi}'\mathbf{y}_{t-1}}{c_t}\right)\right]$, the optimal estimate of $\boldsymbol{\phi}$ is obtained by solving the estimating equation

$$k(u, \lambda) \sum_{t=p+1}^{n} \frac{\mathbf{y}_{t-1}}{c_t} \sin\left[u\left(\frac{y_t - \boldsymbol{\phi}'\mathbf{y}_{t-1}}{c_t}\right)\right] = 0.$$

where $k(u, \lambda) = \frac{2u \exp(-|u|^\lambda)}{1 - \exp(-|2u|^\lambda)}$. By using the recursive algorithm (2.5) and (2.6) and letting $g = \sin\left[u\left(\frac{y_t - \boldsymbol{\phi}'\mathbf{y}_{t-1}}{c_t}\right)\right]$ and $\mathbf{b}_{t-1}^* = k(u, \lambda)\mathbf{y}_{t-1}$, Now it is easily shown that the recursive equations become

$$\mathbf{K}_t = \left\{\mathbf{I}_p + \frac{uk(u, \lambda)\mathbf{y}_{t-1}\mathbf{y}_{t-1}'}{c_t^2} \cos\left[u\left(\frac{y_t - \hat{\boldsymbol{\phi}}_{t-1}'\mathbf{y}_{t-1}}{c_t}\right)\right]\mathbf{K}_{t-1}\right\}^{-1} \mathbf{K}_{t-1} \tag{5}$$

and

$$\hat{\boldsymbol{\phi}}_t = \hat{\boldsymbol{\phi}}_{t-1} + \mathbf{K}_t \frac{k(u, \lambda)\mathbf{y}_{t-1}}{c_t} \sin\left[u\left(\frac{y_t - \hat{\boldsymbol{\phi}}_{t-1}'\mathbf{y}_{t-1}}{c_t}\right)\right]. \tag{6}$$

The algorithm described in (5) and (6) gives the new estimate at time t as the old estimate at time $t-1$ plus an adjustment. Given initial values $\boldsymbol{\phi}_0$ and $\mathbf{J}_0$, we can compute the estimate recursively. The recursive estimate $\hat{\boldsymbol{\phi}}_t$ in (6) is usually referred to as an 'on-line' estimate and it can be mathematically computed in a feasible way, especially when data arise sequentially.

As a special case of the RCA (p) model with stable errors where $\mathbf{b}_t = 0$, the AR (p) model with stable errors is defined as $y_t = \boldsymbol{\phi}'\mathbf{y}_{t-1} + \varepsilon_t$, where $\boldsymbol{\phi}' = (\phi_1, \ldots, \phi_p)$, $\mathbf{y}_{t-1}' = (y_{t-1}, \ldots, y_{t-p})$, and $\varepsilon_t's$ are i.i.d. random variables following a λ-stable distribution with location parameter zero and scale parameter c. That is $\mathbf{b}_t = 0$ in the RCA model, $c_t = c$ and in (3) and (4), the optimal estimate of $\boldsymbol{\phi}$ is simply obtained by solving recursive equations

$$\mathbf{K}_t = \left\{\mathbf{I}_p + \frac{uk(u, \lambda)\mathbf{y}_{t-1}\mathbf{y}_{t-1}'}{c^2} \cos\left[u\left(\frac{y_t - \hat{\boldsymbol{\phi}}_{t-1}'\mathbf{y}_{t-1}}{c}\right)\right]\mathbf{K}_{t-1}\right\}^{-1} \mathbf{K}_{t-1}, \tag{7}$$

where $\mathbf{I}_p$ is the $p \times p$ identity matrix, and

$$\hat{\boldsymbol{\phi}}_t = \hat{\boldsymbol{\phi}}_{t-1} + \mathbf{K}_t \frac{k(u, \lambda)\mathbf{y}_{t-1}}{c} \sin\left[u\left(\frac{y_t - \hat{\boldsymbol{\phi}}_{t-1}'\mathbf{y}_{t-1}}{c}\right)\right]. \tag{8}$$

3. Joint Recursive Estimation of the Location and Scale Parameters

For the single parameter case, recursive estimation has been studied in Thavaneswaran and Abraham (1988), Thavaneswaran and Heyde (1999), and Ghahramani and Thavaneswaran (2012) by using estimating functions. In this section we study the recursive estimation of the location and scale parameters for nonlinear time series models with finite variance and infinite variance error.

Now we consider a real-valued discrete-time stochastic process $\{y_t, t = 1, 2, \ldots\}$ with conditional moments

$$\mu_t(\boldsymbol{\theta}) = \mathrm{E}\left[y_t | \mathscr{F}_{t-1}^y\right], \tag{9}$$

$$\sigma_t^2(\boldsymbol{\theta}) = \mathrm{Var}\left(y_t | \mathscr{F}_{t-1}^y\right). \tag{10}$$

In order to estimate the parameter $\boldsymbol{\theta}$ based on the observations $y_1, \ldots, y_n$, we consider two classes of martingale differences, viz., $\{m_t(\boldsymbol{\theta}) = y_t - \mu_t(\boldsymbol{\theta}), t = 1, \ldots, n\}$ and the generalized martingale differences $\{M_t(\boldsymbol{\theta}) = q(m_t(\boldsymbol{\theta})) - \mathrm{E}[q(m_t(\boldsymbol{\theta}))|\mathscr{F}_{t-1}^y], t = 1, \ldots, n\}$ such that the quadratic estimating functions and the LS and LAD combinations becomes as special cases. The quadratic variations of $m_t(\boldsymbol{\theta})$, $M_t(\boldsymbol{\theta})$ and, the quadratic covariation of $m_t(\boldsymbol{\theta})$ and $M_t(\boldsymbol{\theta})$ are respectively

$$\langle m \rangle_t = \mathrm{E}[m_t^2(\boldsymbol{\theta})|\mathscr{F}_{t-1}^y] = \sigma_t^2(\boldsymbol{\theta}),$$

$$\langle M \rangle_t = E[q^2(m_t(\boldsymbol{\theta})|\mathscr{F}^y_{t-1}] - \left(E[q(m_t(\boldsymbol{\theta}))|\mathscr{F}^y_{t-1}]\right)^2,$$

and

$$\langle m, M \rangle_t = E[m_t(\boldsymbol{\theta})q(m_t(\boldsymbol{\theta}))|\mathscr{F}^y_{t-1}].$$

Here, q is any differentiable function with respect to $\boldsymbol{\theta}$ chosen in a way such that $\langle M \rangle_t$ and $\langle m, M \rangle_t$ exist. The optimal estimating functions based on the martingale differences $m_t(\boldsymbol{\theta})$ and $M_t(\boldsymbol{\theta})$ are respectively

$$\mathbf{g}^*_m(\boldsymbol{\theta}) = -\sum_{t=1}^n \frac{\partial \mu_t(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}} \frac{m_t}{\langle m \rangle_t}$$

and

$$\mathbf{g}^*_M(\boldsymbol{\theta}) = \sum_{t=1}^n \left(E\left[\frac{\partial q(m_t(\boldsymbol{\theta}))}{\partial \boldsymbol{\theta}} \middle| \mathscr{F}^y_{t-1} \right] - \frac{\partial E\left[q(m_t(\boldsymbol{\theta}))| \mathscr{F}^y_{t-1} \right]}{\partial \boldsymbol{\theta}} \right) \frac{M_t}{\langle M \rangle_t}.$$

Then the information associated with $\mathbf{g}^*_m(\boldsymbol{\theta})$ and $\mathbf{g}^*_M(\boldsymbol{\theta})$ are respectively

$$\mathbf{I}_{\mathbf{g}^*_m}(\boldsymbol{\theta}) = \sum_{t=1}^n \frac{\partial \mu_t(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}} \frac{\partial \mu_t(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}'} \frac{1}{\langle m \rangle_t}$$

and

$$\mathbf{I}_{\mathbf{g}^*_M}(\boldsymbol{\theta}) = \sum_{t=1}^n \left(E\left[\frac{\partial q(m_t(\boldsymbol{\theta}))}{\partial \boldsymbol{\theta}} \frac{\partial q(m_t(\boldsymbol{\theta}))}{\partial \boldsymbol{\theta}'} \middle| \mathscr{F}^y_{t-1} \right] \right.$$
$$\left. - \frac{\partial E\left[q(m_t(\boldsymbol{\theta}))| \mathscr{F}^y_{i-1} \right]}{\partial \boldsymbol{\theta}} \frac{\partial E\left[q(m_t(\boldsymbol{\theta}))| \mathscr{F}^y_{t-1} \right]}{\partial \boldsymbol{\theta}'} \right) \frac{1}{\langle M \rangle_t}.$$

Let g_1, g_2 be fixed unbiased estimating functions having finite and positive variances, and such that the expectations of $\partial g_1/\partial \theta$ and $\partial g_2/\partial \theta$ are finite with $E[\partial g_1/\partial \theta] \neq 0$. The following lemma which gives the form of the combined estimating function as a linear combination of orthogonal estimating functions as well as nonorthogonal estimating functions can be used to obtain the recursive estimates of the location and scale parameters.

Lemma 4. *In the class of all unbiased estimating functions* $\{g = \alpha_1 g_1 + \alpha_2 g_2\}$, *the one which maximizes the information* $(E[\partial g/\partial \theta])^2/Var(g)$ *is given:*

(a) by a linear combination of non-orthogonal estimating functions as

$$g^* = \alpha_1^* g_1 + \alpha_2^* g_2$$

where

$$\alpha_1^* = \frac{E\left[\frac{\partial g_1}{\partial \theta} \right] Var(g_2) - E\left[\frac{\partial g_2}{\partial \theta} \right] Cov(g_1, g_2)}{Var(g_1)Var(g_2) - Cov^2(g_1, g_2)}. \tag{11}$$

and

$$\alpha_2^* = \frac{E\left[\frac{\partial g_2}{\partial \theta} \right] Var(g_1) - E\left[\frac{\partial g_1}{\partial \theta} \right] Cov(g_1, g_2)}{Var(g_1)Var(g_2) - Cov^2(g_1, g_2)}; \tag{12}$$

(b) by a linear combination of orthogonal estimating functions g_1 and $\psi = g_2 - \frac{Cov(g_1, g_2)}{Var(g_1)} g_1$ as

$$g^* = \frac{E\left[\frac{\partial g_1}{\partial \theta} \right]}{Var(g_1)} g_1 + \frac{E\left[\frac{\partial g_2}{\partial \theta} \right] - \frac{Cov(g_1, g_2)}{Var(g_1)} E\left[\frac{\partial g_1}{\partial \theta} \right]}{Var(g_2) + \frac{Cov^2(g_1, g_2)}{Var(g_1)} - 2\frac{Cov^2(g_1, g_2)}{Var(g_1)}} \psi = \alpha_1^* g_1 + \alpha_2^* g_2,$$

where α_1^* and α_2^* are given by (9) and (10).

For the discrete-time stochastic process $\{y_t\}$, the following theorem first extends the results in Liang et al. (2011) for *quadratic* estimating functions to the combined estimating functions based on the martingale differences $m_t(\boldsymbol{\theta})$ and the *generalized* martingale differences $M_t(\boldsymbol{\theta})$, and then provides form of the estimates based on the generalized combined estimating functions.

Theorem 5. *For the general model in (7) and (8), in the class of all combined estimating functions of the form*

$$\mathscr{G}_C = \{\mathbf{g}_C(\boldsymbol{\theta}) : \mathbf{g}_C(\boldsymbol{\theta}) = \sum_{t=1}^{n} (\mathbf{a}_{t-1}(\boldsymbol{\theta}) m_t(\boldsymbol{\theta}) + \mathbf{b}_{t-1}(\boldsymbol{\theta}) M_t(\boldsymbol{\theta}))\},$$

(a) the optimal estimating function is given by

$$\mathbf{g}_C^*(\boldsymbol{\theta}) = \sum_{t=1}^{n} \left(\mathbf{a}_{t-1}^*(\boldsymbol{\theta}) m_t(\boldsymbol{\theta}) + \mathbf{b}_{t-1}^*(\boldsymbol{\theta}) M_t(\boldsymbol{\theta}) \right),$$

where

$$
\mathbf{a}_{t-1}^*(\boldsymbol{\theta}) = \left(1 - \frac{\langle m, M \rangle_t^2}{\langle m \rangle_t \langle M \rangle_t} \right)^{-1} \left(-\frac{\partial \mu_t(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}} \frac{1}{\langle m \rangle_t} \right. \\
\left. - \left(E\left[\frac{\partial q(m_t(\boldsymbol{\theta}))}{\partial \boldsymbol{\theta}} \middle| \mathscr{F}_{t-1}^y \right] - \frac{\partial E\left[q(m_t(\boldsymbol{\theta})) | \mathscr{F}_{t-1}^y \right]}{\partial \boldsymbol{\theta}} \right) \frac{\langle m, M \rangle_t}{\langle m \rangle_t \langle M \rangle_t} \right) \tag{13}
$$

and

$$
\mathbf{b}_{t-1}^*(\boldsymbol{\theta}) = \left(1 - \frac{\langle m, M \rangle_t^2}{\langle m \rangle_t \langle M \rangle_t} \right)^{-1} \left(\frac{\partial \mu_t(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}} \frac{\langle m, M \rangle_t}{\langle m \rangle_t \langle M \rangle_t} \right. \\
\left. + \left(E\left[\frac{\partial q(m_t(\boldsymbol{\theta}))}{\partial \boldsymbol{\theta}} \middle| \mathscr{F}_{t-1}^y \right] - \frac{\partial E\left[q(m_t(\boldsymbol{\theta})) | \mathscr{F}_{t-1}^y \right]}{\partial \boldsymbol{\theta}} \right) \frac{1}{\langle M \rangle_t} \right); \tag{14}
$$

(b) the optimal estimating function is equal to $\mathbf{g}_m^*(\boldsymbol{\theta})$ if

$$
\frac{\partial \mu_t}{\partial \boldsymbol{\theta}} \frac{\langle m, M \rangle_t}{\langle m \rangle_t} = -\left(E\left[\frac{\partial q(m_t(\boldsymbol{\theta}))}{\partial \boldsymbol{\theta}} \middle| \mathscr{F}_{i-1}^y \right] - \frac{\partial E\left[q(m_t(\boldsymbol{\theta})) | \mathscr{F}_{i-1}^y \right]}{\partial \boldsymbol{\theta}} \right), \text{ for each } i;
$$

(c) the recursive estimate for $\boldsymbol{\theta}$ is given by

$$\hat{\boldsymbol{\theta}}_t = \hat{\boldsymbol{\theta}}_{t-1} + \mathbf{K}_t \left(\mathbf{a}_{t-1}^*(\hat{\boldsymbol{\theta}}_{t-1}) m_t(\hat{\boldsymbol{\theta}}_{t-1}) + \mathbf{b}_{t-1}^*(\hat{\boldsymbol{\theta}}_{t-1}) M_t(\hat{\boldsymbol{\theta}}_{t-1}) \right), \tag{15}$$

$$\mathbf{K}_i = \mathbf{K}_{i-1} \left(\mathbf{I}_p - \left(\mathbf{a}_{t-1}^*(\hat{\boldsymbol{\theta}}_{t-1}) \frac{\partial m_t(\hat{\boldsymbol{\theta}}_{t-1})}{\partial \boldsymbol{\theta}'} + \frac{\partial \mathbf{a}_{t-1}^*(\hat{\boldsymbol{\theta}}_{t-1})}{\partial \boldsymbol{\theta}} m_t(\hat{\boldsymbol{\theta}}_{t-1}) \right. \right.$$

$$\left. \left. + \mathbf{b}_{t-1}^*(\hat{\boldsymbol{\theta}}_{t-1}) \frac{\partial M_t(\hat{\boldsymbol{\theta}}_{t-1})}{\partial \boldsymbol{\theta}'} + \frac{\partial \mathbf{b}_{t-1}^*(\hat{\boldsymbol{\theta}}_{t-1})}{\partial \boldsymbol{\theta}} M_t(\hat{\boldsymbol{\theta}}_{t-1}) \right) \mathbf{K}_{t-1} \right)^{-1}, \tag{16}$$

where $\mathbf{I}_p$ is $p \times p$ identity matrix, and $\mathbf{a}_{t-1}^*$ and $\mathbf{b}_{t-1}^*$ can be calculated by substituting $\hat{\boldsymbol{\theta}}_{t-1}$ in equation (13) and (14) respectively.

(d) for the scalar parameter case, the recursive estimate of θ is given by

$$\hat{\theta}_t = \hat{\theta}_{t-1} + K_t(a_{t-1}^*(\hat{\theta}_{t-1}) m_i(\hat{\theta}_{t-1}) + b_{t-1}^*(\hat{\theta}_{t-1}) M_t(\hat{\theta}_{t-1})), \tag{17}$$

$$K_t = K_{t-1} \left(1 - \left(a_{t-1}^*(\hat{\theta}_{t-1}) \frac{\partial m_t(\hat{\theta}_{t-1})}{\partial \theta} + \frac{\partial a_{t-1}^*(\hat{\theta}_{t-1})}{\partial \theta} m_t(\hat{\theta}_{t-1}) \right. \right.$$

$$\left. \left. + b_{t-1}^*(\hat{\theta}_{t-1}) \frac{\partial M_t(\hat{\theta}_{i-1})}{\partial \theta} + \frac{\partial b_{t-1}^*(\hat{\theta}_{t-1})}{\partial \theta} M_i(\hat{\theta}_{t-1}) \right) K_{t-1} \right)^{-1}. \tag{18}$$

Proof. The proof of part (a) and part (b) are somewhat similar to the proof given in Liang et al. (2011). Part (c) and Part (d) extend the results in Thavaneswaran and Heyde (1999) for the generalized combined estimating function $\mathbf{g}_C^*(\boldsymbol{\theta})$ with the vector parameter. Detailed proof of the theorem is given in the Appendix A. □

Note 1. The optimal information matrix based on the first i observations is given by $-E\left[\frac{\partial \mathbf{g}_C^*(\theta)}{\partial \theta}\Big| \mathscr{F}_{t-1}^y\right]$ and hence, $K_t^{-1} = -\sum_{s=1}^t \frac{\partial \mathbf{g}_C^*(\hat{\theta}_{t-1})}{\partial \theta}$ can be interpreted as the observed information matrix associated with the optimal combined estimating function $\mathbf{g}_C^*(\theta)$.

3.1 Autoregressive Models with Student's t Errors

Consider an autoregressive model of the form

$$y_t = \theta y_{t-1} + \sigma(\theta)\varepsilon_t$$

where $\sigma(\theta)$ is any differentiable function of θ, and $\{\varepsilon_t\}$ are uncorrelated random variables having the student t distribution with density function

$$g(x) = \frac{\Gamma\left(\frac{v+1}{2}\right)}{\sqrt{\pi v}\Gamma\left(\frac{v}{2}\right)}\left(1 + \frac{x^2}{v}\right)^{-\frac{v+1}{2}}, \quad -\infty < x < \infty.$$

Then the conditional mean and variance of y_t are $\mu_t(\theta) = \theta y_{t-1}$ and $\sigma_t^2(\theta) = \frac{\sigma^2(\theta)v}{v-2}$ for $v > 2$. Maximum likelihood estimation for this model becomes more complicated and the LAD estimating function method had been discussed Thavaneswaran and Heyde (1999) and combined estimating function method has been discussed in Ghahramani and Thavaneswaran (2009).

We consider two classes of estimating functions generated by the martingale differences $\{m_t = y_t - \mu_t, t = 1, \ldots, n\}$ and $\{M_t = -\text{sgn}(m_t), t = 1, \ldots, n\}$. For $v > 2$, it can be easily shown that the LS estimating function based on m_t is given by

$$g_m^*(\theta) = -\frac{v-2}{\sigma^2(\theta)v}\sum_{t=1}^n y_{t-1}(y_t - \theta y_{t-1}),$$

with information

$$I_{g_m^*}(\theta) = \frac{v-2}{\sigma^2(\theta)v}\sum_{t=1}^n y_{t-1}^2.$$

On the other hand, for any value of v, the LAD estimating function based on M_t is given by

$$g_M^*(\theta) = -\frac{\Gamma\left(\frac{v+1}{2}\right)}{\sigma(\theta)\sqrt{\pi v}\Gamma\left(\frac{v}{2}\right)}\sum_{t=1}^n y_{t-1}\,\text{sgn}(y_t - \theta y_{t-1}),$$

with information

$$I_{g_M^*}(\theta) = \frac{4\left(\Gamma\left(\frac{v+1}{2}\right)\right)^2}{\sigma^2(\theta)\pi v\left(\Gamma\left(\frac{v}{2}\right)\right)^2}\sum_{t=1}^n y_{t-1}^2.$$

For $v \leq 2$, the LS estimating function is not defined, whereas the LAD estimating function provides an estimate of θ. For $v > 2$, neither the LS nor the LAD estimating function is optimal. In the class of all combined estimating functions of the form $\{g_C(\theta) : g_C(\theta) = \sum_{t=1}^n (a_{t-1}m_t + b_{t-1}M_t)\}$, the optimal estimating function is given by

$$g_C^*(\theta) = \sum_{t=1}^n \left(a_{t-1}^* m_t + b_{t-1}^* M_t\right)$$

where

$$a_{t-1}^* = -\frac{v-2}{\sigma^2(\theta)v(1-\rho^2)}\left(1 - \frac{2\Gamma\left(\frac{v+1}{2}\right)\rho}{\sqrt{\pi(v-2)}\Gamma\left(\frac{v}{2}\right)}\right)y_{t-1} = \frac{c_1 y_{t-1}}{\sigma^2(\theta)}$$

$$b_{t-1}^* = -\frac{v-2}{\sigma(\theta)v(1-\rho^2)}\left(\sqrt{\frac{v}{v-2}}\rho - \frac{2\Gamma\left(\frac{v+1}{2}\right)}{\sqrt{\pi v}\Gamma\left(\frac{v}{2}\right)}\frac{v}{v-2}\right)y_{t-1} = \frac{c_2 y_{t-1}}{\sigma(\theta)}.$$

where c_1 and c_2 are the constants and $\rho = E\left[|m_t|\big|\mathscr{F}_{t-1}^y\right]/\sigma_t(\theta)$ for $v > 2$.

Furthermore, the conditional information $I_{g_C^*}(\theta)$ is given by

$$I_{g_C^*}(\theta) = \frac{1}{1-\rho^2}\left(I_{g_m^*}(\theta) + I_{g_M^*}(\theta) - \frac{4\Gamma\left(\frac{v+1}{2}\right)\sqrt{v-2}\rho}{\sigma^2(\theta)\sqrt{\pi v}\Gamma\left(\frac{v}{2}\right)}\sum_{t=1}^n y_{t-1}^2\right).$$

In order to obtain the recursive equations for θ, we first derive the derivatives of $a_{t-1}^*, b_{t-1}^*, m_t, M_t$ with respect to θ. It is easily to calculate that

$$\frac{\partial a_{t-1}^*}{\partial \theta} = -\frac{2c_1 y_{t-1}}{\sigma^3(\theta)} \frac{\partial \sigma(\theta)}{\partial \theta}, \quad \frac{\partial b_{t-1}^*}{\partial \theta} = -\frac{c_1 y_{t-1}}{\sigma^2(\theta)} \frac{\partial \sigma(\theta)}{\partial \theta}, \quad \frac{\partial m_t}{\partial \theta} = -y_{t-1}, \quad \frac{\partial m_t}{\partial \theta} = 2y_{t-1}\delta(y_t - \theta y_{t-1})$$

where δ is the dirac delta function. Then the recursive estimate of θ is given by

$$\hat{\theta}_t = \hat{\theta}_{t-1} + K_t \left(\frac{c_1 y_{t-1}(y_t - \hat{\theta}_{t-1} y_{t-1})}{\sigma^2(\hat{\theta}_{t-1})} - \frac{c_2 y_{t-1} \operatorname{sgn}(y_t - \hat{\theta}_{t-1} y_{t-1})}{\sigma(\hat{\theta}_{t-1})} \right),$$

$$K_t = K_{t-1} \left(1 + \left(\frac{c_1 y_{t-1}^2}{\sigma^2(\hat{\theta}_{t-1})} + \frac{2c_1 y_{t-1}(y_t - \hat{\theta}_{t-1}y_{t-1})}{\sigma^3(\hat{\theta}_{t-1})} \frac{\partial \sigma(\hat{\theta}_{t-1})}{\partial \theta} \right. \right.$$
$$\left. \left. - \frac{2c_2 y_{t-1}^2 \delta(y_t - \hat{\theta}_{t-1} y_{t-1})}{\sigma(\hat{\theta}_{t-1})} - \frac{c_2 y_{t-1} \operatorname{sgn}(y_t - \hat{\theta}_{t-1} y_{t-1})}{\sigma^2(\hat{\theta}_{t-1})} \frac{\partial \sigma(\hat{\theta}_{t-1})}{\partial \theta} \right) K_{t-1} \right)^{-1}.$$

3.2 Recursive Estimation for Location and Scale Parameters of a Cauchy Distribution

The Cauchy distribution has the probability density function

$$f(x; \theta, b) = \frac{1}{\pi} \left[\frac{b}{(x - \theta)^2 + b^2} \right]$$

where θ is the location parameter, specifying the location of the peak of the distribution, and b is the scale parameter which specifies the half-width at half-maximum (HWHM). Now we are interested in how to estimate the parameters θ and b jointly. Because the parameters of the Cauchy distribution do not correspond to a mean and variance and attempt to estimate the parameters of the Cauchy distribution by using a sample mean and a sample variance will not succeed. On the other hand, maximum likelihood can also be used to estimate the parameters θ and b. However, this tends to be complicated by the fact that this requires finding the roots of a high degree polynomial, and there can be multiple roots that represent local maxima. Also, while the maximum likelihood estimator is asymptotically efficient, it is relatively inefficient for small samples. In order to solve the estimation problem, we use he estimating function method and obtain the recursive estimates. The advantages of the method include the simplicity of constructing the estimating equations and the explicit calculation of the estimates.

If $X_1, \cdots, X_n$ are independent and identically distributed random variables from the Cauchy distribution, then based on the martingale differences

$$\mathbf{h}_t(\theta, b) = \left(\sin[u_1(x_t - \theta)], \cos[u_2(x_t - \theta)] - e^{-b|u_2|} \right)'$$

the estimating function is given by

$$g^*(\theta, b) = \sum_{t=1}^n \left(\begin{array}{c} \frac{-2u_1 e^{-b|u_1|}}{1 - e^{-2b|u_1|}} \sin[u_1(x_t - \theta)] \\ \frac{2|u_2| e^{-b|u_2|}}{1 - e^{-2b|u_2|}} \left(\cos[u_2(x_t - \theta)] - e^{-b|u_2|} \right) \end{array} \right).$$

For simplicity, let $\beta = e^{-b|u_2|}$, then the estimating function becomes

$$g^*(\theta, b) = \sum_{t=1}^n \left(\begin{array}{c} \frac{-2u_1 \beta^{\left| \frac{u_1}{u_2} \right|}}{1 - \beta^{2\left| \frac{u_1}{u_2} \right|}} \sin[u_1(x_t - \theta)] \\ \frac{2|u_2|\beta}{1 - \beta^2} \left(\cos[u_2(x_t - \theta)] - \beta \right) \end{array} \right).$$

Therefore the recursive equations for θ and β are

$$\left(\begin{array}{c} \hat{\theta}_t \\ \hat{\beta}_t \end{array} \right) = \left(\begin{array}{c} \hat{\theta}_{t-1} \\ \hat{\beta}_{t-1} \end{array} \right) + \mathbf{K}_t \left(\begin{array}{c} \frac{-2u_1 \hat{\beta}_{t-1}^{\left| \frac{u_1}{u_2} \right|}}{1 - \hat{\beta}_{t-1}^{2\left| \frac{u_1}{u_2} \right|}} \sin[u_1(x_t - \hat{\theta}_{t-1})] \\ \frac{2|u_2|\hat{\beta}_{t-1}}{1 - \hat{\beta}_{t-1}^2} \left(\cos[u_2(x_t - \hat{\theta}_{t-1})] - \hat{\beta}_{t-1} \right) \end{array} \right)$$

and

$$\mathbf{K}_t^{-1} = \mathbf{K}_{t-1}^{-1} - \left(\begin{array}{cc} \frac{2u_1^2 \hat{\beta}_{t-1}^{\left| \frac{u_1}{u_2} \right|}}{1 - \hat{\beta}_{t-1}^{2\left| \frac{u_1}{u_2} \right|}} \cos[u_1(x_t - \hat{\theta}_{t-1})] & \frac{-2u_1 \left| \frac{u_1}{u_2} \right| \hat{\beta}_{t-1}^{\left| \frac{u_1}{u_2} \right| - 1} \left(1 + \hat{\beta}_{t-1}^{2\left| \frac{u_1}{u_2} \right|} \right)}{\left(1 - \hat{\beta}_{t-1}^{2\left| \frac{u_1}{u_2} \right|} \right)^2} \sin[u_1(x_t - \hat{\theta}_{t-1})] \\ \frac{2u_2 |u_2| \hat{\beta}_{t-1}}{1 - \hat{\beta}_{t-1}^2} \sin[u_2(x_t - \hat{\theta}_{t-1})] & \frac{2|u_2|(1 + \hat{\beta}_{t-1}^2) \cos[u_2(x_t - \hat{\theta}_{t-1})] - 4|u_2|\hat{\beta}_{t-1}}{(1 - \hat{\beta}_{t-1}^2)^2} \end{array} \right).$$

3.3 AR(p) Models with Stable Errors

For the AR(p) process given by $y_t = \boldsymbol{\phi}' \mathbf{y}_{t-1} + \varepsilon_t$, where $\boldsymbol{\phi}' = (\phi_1, \ldots, \phi_p)$, $\mathbf{y}'_{t-1} = (y_{t-1}, \ldots, y_{t-p})$, and $\varepsilon'_t s$ are i.i.d. random variables following a λ-stable distribution with location parameter zero and scale parameter c. In order to estimate the vector parameter $\boldsymbol{\theta} = (\boldsymbol{\phi}', c)'$, we consider a class of martingale differences

$$\mathbf{h}_t(\boldsymbol{\theta}) = \left\{ \sin\left[u_1 \left(\frac{y_t - \boldsymbol{\phi}' \mathbf{y}_{t-1}}{c} \right) \right], \cos\left[u_2 \left(\frac{y_t - \boldsymbol{\phi}' \mathbf{y}_{t-1}}{c} \right) \right] - \exp(-|u_2|^\lambda), t = p+1, \ldots, n \right\}.$$

Then the optimal estimating function based on $\mathbf{h}_t(\boldsymbol{\theta})$ is

$$\mathbf{g}_n^*(\boldsymbol{\theta}) = \begin{pmatrix} \frac{2u_1 e^{-|u_1|^\lambda}}{c(1 - e^{-|2u_1|^\lambda})} \sum_{t=p+1}^n \mathbf{y}_{t-1} \sin\left[u_1 \left(\frac{y_t - \boldsymbol{\phi}' \mathbf{y}_{t-1}}{c} \right) \right] \\ \\ -\frac{2\lambda |u_2|^\lambda e^{-|u_2|^\lambda}}{c \left(1 + e^{-|2u_2|^\lambda} - 2e^{-2|u_2|^\lambda} \right)} \sum_{t=p+1}^n \left\{ \cos\left[u_2 \left(\frac{y_t - \boldsymbol{\phi}' \mathbf{y}_{t-1}}{c} \right) \right] - \exp(-|u_2|^\lambda) \right\} \end{pmatrix} \tag{19}$$

and the corresponding information matrix is

$$\mathbf{I}_{\mathbf{g}_n^*} = \begin{pmatrix} \frac{2u_1^2 \exp(-2|u_1|^\lambda)}{c^2 [1 - \exp(-|2u_1|^\lambda)]} \sum_{t=p+1}^n \mathbf{y}_{t-1} \mathbf{y}'_{t-1} & 0 \\ \\ 0 & \frac{2n\lambda^2 u_2^{2\lambda} e^{-2|u_2|^\lambda}}{c^2 \left(1 + e^{-|2u_2|^\lambda} - 2e^{-2|u_2|^\lambda} \right)} \end{pmatrix}.$$

Let

$$k_1(u_1, \lambda) = \frac{2u_1 e^{-|u_1|^\lambda}}{1 - e^{-|2u_1|^\lambda}}, k_2(u_2, \lambda) = \frac{2\lambda |u_2|^\lambda e^{-|u_2|^\lambda}}{1 + e^{-|2u_2|^\lambda} - 2e^{-2|u_2|^\lambda}}.$$

Therefore the recursive equations for θ and β are

$$\begin{pmatrix} \hat{\boldsymbol{\phi}}_t \\ \hat{c}_t \end{pmatrix} = \begin{pmatrix} \hat{\boldsymbol{\phi}}_{t-1} \\ \hat{c}_{t-1} \end{pmatrix} + \mathbf{K}_t \begin{pmatrix} \frac{k_1(u_1, \lambda) \mathbf{y}_{t-1}}{\hat{c}_{t-1}} \sin\left[u_1 \left(\frac{y_t - \boldsymbol{\phi}'_{t-1} \mathbf{y}_{t-1}}{\hat{c}_{t-1}} \right) \right] \\ \\ -\frac{k_2(u_2, \lambda)}{\hat{c}_{t-1}} \left\{ \cos\left[u_2 \left(\frac{y_t - \boldsymbol{\phi}'_{t-1} \mathbf{y}_{t-1}}{\hat{c}_{t-1}} \right) \right] - \exp(-|u_2|^\lambda) \right\} \end{pmatrix}$$

and

$$\mathbf{K}_t^{-1} = \mathbf{K}_{t-1}^{-1} - \begin{pmatrix} D_{11} & D_{12} \\ D_{21} & D_{22} \end{pmatrix},$$

where

$$D_{11} = -\frac{u_1 k_1(u_1, \lambda) \mathbf{y}_{t-1} \mathbf{y}'_{t-1}}{\hat{c}_{t-1}^2} \cos\left[u_1 \left(\frac{y_t - \boldsymbol{\phi}'_{t-1} \mathbf{y}_{t-1}}{\hat{c}_{t-1}} \right) \right]$$

$$D_{12} = -\frac{k_1(u_1, \lambda) \mathbf{y}_{t-1}}{\hat{c}_{t-1}^2} \sin\left[u_1 \left(\frac{y_t - \boldsymbol{\phi}'_{t-1} \mathbf{y}_{t-1}}{\hat{c}_{t-1}} \right) \right]$$

$$\quad - \frac{u_1 k_1(u_1, \lambda) \mathbf{y}_{t-1} (y_t - \boldsymbol{\phi}'_{t-1} \mathbf{y}_{t-1})}{\hat{c}_{t-1}^3} \cos\left[u_1 \left(\frac{y_t - \boldsymbol{\phi}'_{t-1} \mathbf{y}_{t-1}}{\hat{c}_{t-1}} \right) \right]$$

$$D_{21} = -\frac{u_2 k_2(u_2, \lambda) \mathbf{y}'_{t-1}}{\hat{c}_{t-1}^2} \sin\left[u_2 \left(\frac{y_t - \boldsymbol{\phi}'_{t-1} \mathbf{y}_{t-1}}{\hat{c}_{t-1}} \right) \right]$$

$$D_{22} = \frac{k_2(u_2, \lambda)}{\hat{c}_{t-1}^2} \left\{ \cos\left[u_2 \left(\frac{y_t - \boldsymbol{\phi}'_{t-1} \mathbf{y}_{t-1}}{\hat{c}_{t-1}} \right) \right] - \exp(-|u_2|^\lambda) \right\}$$

$$\quad - \frac{u_2 k_2(u_2, \lambda) (y_t - \boldsymbol{\phi}'_{t-1} \mathbf{y}_{t-1})}{\hat{c}_{t-1}^3} \sin\left[u_2 \left(\frac{y_t - \boldsymbol{\phi}'_{t-1} \mathbf{y}_{t-1}}{\hat{c}_{t-1}} \right) \right].$$

.

3.4 Regression Model with ARCH Errors

Consider a regression model with ARCH (s) errors ε_t of the form

$$y_t = \boldsymbol{\beta}' \mathbf{x_t} + \varepsilon_t \tag{20}$$

such that $\mathrm{E}\left[\varepsilon_t | \mathscr{F}^y_{t-1}\right] = 0$, and $\mathrm{Var}\left(\varepsilon_t | \mathscr{F}^y_{t-1}\right) = h_t = \alpha_0 + \alpha_1 \varepsilon^2_{t-1} + \ldots + \alpha_s \varepsilon^2_{t-s}$. In this model, the conditional mean is $\mu_t = \mathbf{x_t}' \boldsymbol{\beta}$, the conditional variance is $\sigma_t^2 = h_t$, and the conditional skewness and excess kurtosis are assumed to be constants γ and κ, respectively. It follows from Theorem 1 that the optimal component quadratic estimating function for the parameter vector $\boldsymbol{\theta} = (\beta_1, \ldots, \beta_r, \alpha_0, \ldots, \alpha_s)' = (\boldsymbol{\beta}', \boldsymbol{\alpha}')'$ based on $m_t = y_t - \boldsymbol{\beta}' \mathbf{x_t}$ and $M_t = m_t^2 - h_t$ is

$$
\begin{aligned}
\mathbf{g}_C^*(\boldsymbol{\beta}) = \ & \frac{1}{\kappa + 2 - \gamma^2} \sum_{t=1}^n \frac{1}{h_t^2} \left(\left(-h_t(\kappa + 2)\mathbf{x}_t - 2h_t^{1/2}\gamma \sum_{j=1}^s \alpha_j \mathbf{x}_{t-j} \varepsilon_{t-j} \right) m_t \right. \\
& \left. + \left(h_t^{1/2} \gamma \mathbf{x}_t + 2 \sum_{j=1}^s \alpha_j \mathbf{x}_{t-j} \varepsilon_{t-j} \right) M_t \right)
\end{aligned}
$$

and

$$
\begin{aligned}
\mathbf{g}_C^*(\boldsymbol{\alpha}) = \ & \frac{1}{\kappa + 2 - \gamma^2} \sum_{t=1}^n \frac{1}{h_t^2} \left(h_t^{1/2} \gamma (1, \varepsilon^2_{t-1}, \ldots, \varepsilon^2_{t-s})' m_t \right. \\
& \left. - (1, \varepsilon^2_{t-1}, \ldots, \varepsilon^2_{t-s})' M_t \right).
\end{aligned}
$$

Moreover, the information matrix for $\boldsymbol{\theta} = (\boldsymbol{\beta}', \boldsymbol{\alpha}')'$ is given by

$$
\mathbf{I} = \left(1 - \frac{\gamma^2}{\kappa + 2} \right)^{-1} \begin{pmatrix} \mathbf{I}_{\beta\beta} & \mathbf{I}_{\beta\alpha} \\ \mathbf{I}_{\alpha\beta} & \mathbf{I}_{\alpha\alpha} \end{pmatrix}, \tag{21}
$$

where

$$
\mathbf{I}_{\beta\beta} = \sum_{t=1}^n \left(\frac{\mathbf{x}_t \mathbf{x}_t'}{h_t} + \frac{4 \left(\sum_{j=1}^s \alpha_j \mathbf{x}_{t-j} \varepsilon_{t-j} \right) \left(\sum_{j=1}^s \alpha_j \mathbf{x}_{t-j}' \varepsilon_{t-j} \right)}{h_t^2 (\kappa + 2)} + \frac{2\gamma \mathbf{x}_t \left(\sum_{j=1}^s \alpha_j \mathbf{x}_{t-j}' \varepsilon_{t-j} \right)}{h_t^{3/2}(\kappa + 2)} \right),
$$

$$
\mathbf{I}_{\beta\alpha} = - \sum_{t=1}^n \frac{\left(h_t^{1/2} \gamma \mathbf{x}_t + 2 \sum_{j=1}^s \alpha_j \mathbf{x}_{t-j} \varepsilon_{t-j} \right) (1, \varepsilon^2_{t-1}, \ldots, \varepsilon^2_{t-s})}{h_t^2 (\kappa + 2)},
$$

$$
\mathbf{I}_{\alpha\beta} = \mathbf{I}'_{\beta\alpha} = - \sum_{t=1}^n \frac{(1, \varepsilon^2_{t-1}, \ldots, \varepsilon^2_{t-s})' \left(h_t^{1/2} \gamma \mathbf{x}_t' + 2 \sum_{j=1}^s \alpha_j \mathbf{x}_{t-j}' \varepsilon_{t-j} \right)}{h_t^2 (\kappa + 2)},
$$

and

$$
\mathbf{I}_{\alpha\alpha} = \sum_{t=1}^n \frac{(1, \varepsilon^2_{t-1}, \ldots, \varepsilon^2_{t-s})' (1, \varepsilon^2_{t-1}, \ldots, \varepsilon^2_{t-s})}{h_t^2 (\kappa + 2)}.
$$

It is of interest to note that when $\{\varepsilon_t\}$ are conditionally Gaussian such that $\gamma = 0, \kappa = 0$,

$$
E \left[\frac{\left(\sum_{j=1}^s \alpha_j \mathbf{x}_{t-j} \varepsilon_{t-j} \right) (1, \varepsilon^2_{t-1}, \ldots, \varepsilon^2_{t-s})}{h_t^2 (\kappa + 2)} \right] = \mathbf{0}.
$$

The optimal quadratic estimating functions for $\boldsymbol{\beta}$ and $\boldsymbol{\alpha}$ based on the estimating functions $m_t = y_t - \mathbf{x}_t \beta$ and $M_t = m_t^2 - h_t$, are respectively given by

$$
\mathbf{g}_C^*(\boldsymbol{\beta}) = \sum_{t=1}^n \frac{1}{h_t^2} \left(-h_t \mathbf{x}_t m_t + \left(\sum_{j=1}^s \alpha_j \mathbf{x}_{t-j} \varepsilon_{t-j} \right) M_t \right)
$$

and

$$
\mathbf{g}_C^*(\boldsymbol{\alpha}) = - \sum_{t=1}^n \frac{1}{h_t^2} (1, \varepsilon^2_{t-1}, \ldots, \varepsilon^2_{t-s})' M_t.
$$

Moreover, the information matrix for $\boldsymbol{\theta} = (\boldsymbol{\beta}', \boldsymbol{\alpha}')'$ in (3.8) has $\mathbf{I}_{\beta\alpha} = \mathbf{I}_{\alpha\beta} = \mathbf{0}$,

$$
\mathbf{I}_{\beta\beta} = \sum_{t=1}^n \frac{h_t \mathbf{x}_t \mathbf{x}_t' + 2 \left(\sum_{j=1}^s \alpha_j \mathbf{x}_{t-j} \varepsilon_{t-j} \right) \left(\sum_{j=1}^s \alpha_j \mathbf{x}_{t-j}' \varepsilon_{t-j} \right)}{h_t^2},
$$

and

$$\mathbf{I}_{\alpha\alpha} = \sum_{t=1}^{n} \frac{(1, \varepsilon_{t-1}^2, \ldots, \varepsilon_{t-s}^2)'(1, \varepsilon_{t-1}^2, \ldots, \varepsilon_{t-s}^2)}{2h_t^2}.$$

Now we consider the recursive estimation for β and α. For $\mathbf{g}_C^*(\boldsymbol{\beta})$,

$$\mathbf{a}_{t-1}^* = \frac{1}{(\kappa + 2 - \gamma^2)h_t^2}\left(-h_t(\kappa+2)\mathbf{x_t} - 2h_t^{1/2}\gamma \sum_{j=1}^{s}\alpha_j \mathbf{x}_{t-j}\varepsilon_{t-j}\right)$$

and

$$\mathbf{b}_{t-1}^* = \frac{1}{(\kappa + 2 - \gamma^2)h_t^2}\left(h_t^{1/2}\gamma\mathbf{x}_t + 2\sum_{j=1}^{s}\alpha_j \mathbf{x}_{t-j}\varepsilon_{t-j}\right).$$

Hence,

$$\frac{\partial \mathbf{a}_{t-1}^*}{\partial \boldsymbol{\beta}} = \frac{1}{(\kappa + 2 - \gamma^2)h_t^2}\left(-2(\kappa+2)\mathbf{x}_t\left(\sum_{j=1}^{s}\alpha_j \mathbf{x}_{t-j}'\varepsilon_{t-j}\right)\right.$$
$$\left. -6h_t^{-1/2}\gamma\left(\sum_{j=1}^{s}\alpha_j \mathbf{x}_{t-j}\varepsilon_{t-j}\right)\left(\sum_{j=1}^{s}\alpha_j \mathbf{x}_{t-j}'\varepsilon_{t-j}\right) + 2\gamma h_t^{1/2}\left(\sum_{j=1}^{s}\alpha_j \mathbf{x}_{t-j}\mathbf{x}_{t-j}'\right)\right),$$

and

$$\frac{\partial \mathbf{b}_{t-1}^*}{\partial \boldsymbol{\beta}} = \frac{1}{(\kappa + 2 - \gamma^2)h_t^2}\left(3\gamma h_t^{-1/2}\mathbf{x}_t\left(\sum_{j=1}^{s}\alpha_j \mathbf{x}_{t-j}'\varepsilon_{t-j}\right)\right.$$
$$\left. +8h_t^{-1}\left(\sum_{j=1}^{s}\alpha_j \mathbf{x}_{t-j}\varepsilon_{t-j}\right)\left(\sum_{j=1}^{s}\alpha_j \mathbf{x}_{t-j}'\varepsilon_{t-j}\right) - 2\left(\sum_{j=1}^{s}\alpha_j \mathbf{x}_{t-j}\mathbf{x}_{t-j}'\right).\right)$$

Also

$$\frac{m_t}{\partial \boldsymbol{\beta}} = -\mathbf{x}_t, \frac{M_t}{\partial \boldsymbol{\beta}} = 2(y_t - \boldsymbol{\beta}'\mathbf{x}_t)\mathbf{x}_t + 2\sum_{j=1}^{s}\alpha_j \mathbf{x}_{t-j}'\varepsilon_{t-j},$$

therefore, by using the recursive algorithm (2.5) and (2.6), the recursive equations for β are given by

$$\hat{\boldsymbol{\beta}}_t = \hat{\boldsymbol{\beta}}_{t-1} + \mathbf{K}_t\left(\mathbf{a}_{t-1}^*(\hat{\alpha}_{t-1}, \hat{\boldsymbol{\beta}}_{t-1})m_t(\hat{\boldsymbol{\beta}}_{t-1}) + \mathbf{b}_{t-1}^*(\hat{\alpha}_{t-1}, \hat{\boldsymbol{\beta}}_{t-1})M_t(\hat{\alpha}_{t-1}, \hat{\boldsymbol{\beta}}_{t-1})\right),$$

$$\mathbf{K}_t = \mathbf{K}_{t-1}\left(\mathbf{I}_r - \left(\mathbf{a}_{t-1}^*(\hat{\alpha}_{t-1}, \hat{\boldsymbol{\beta}}_{t-1})\frac{\partial m_t(\hat{\boldsymbol{\beta}}_{t-1})}{\partial \boldsymbol{\beta}'} + \frac{\partial \mathbf{a}_{t-1}^*(\hat{\alpha}_{t-1}, \hat{\boldsymbol{\beta}}_{t-1})}{\partial \boldsymbol{\beta}}m_t(\hat{\boldsymbol{\beta}}_{t-1})\right.\right.$$
$$\left.\left. + \ \mathbf{b}_{t-1}^*(\hat{\alpha}_{t-1}, \hat{\boldsymbol{\beta}}_{t-1})\frac{\partial M_t(\hat{\alpha}_{t-1}, \hat{\boldsymbol{\beta}}_{t-1})}{\partial \boldsymbol{\beta}'} + \frac{\partial \mathbf{b}_{t-1}^*(\hat{\alpha}_{t-1}, \hat{\boldsymbol{\beta}}_{t-1}))}{\partial \boldsymbol{\beta}}M_t(\hat{\alpha}_{t-1}, \hat{\boldsymbol{\beta}}_{t-1})\right)\mathbf{K}_{t-1}\right)^{-1},$$

One the other hand, for $\mathbf{g}_C^*(\boldsymbol{\alpha})$,

$$\mathbf{c}_{t-1}^* = -\frac{1}{h_t^2}(1, \varepsilon_{t-1}^2, \ldots, \varepsilon_{t-s}^2)'$$

and

$$\frac{\partial \mathbf{c}_{t-1}^*}{\partial \boldsymbol{\alpha}} = \frac{2}{h_t^3}(1, \varepsilon_{t-1}^2, \ldots, \varepsilon_{t-s}^2)'(1, \varepsilon_{t-1}^2, \ldots, \varepsilon_{t-s}^2).$$

Then, by using with the result Also

$$\frac{\partial m_t}{\partial \boldsymbol{\alpha}} = 0, \frac{\partial M_t}{\partial \boldsymbol{\alpha}} = (1, \varepsilon_{t-1}^2, \ldots, \varepsilon_{t-s}^2)',$$

the recursive equations for β are given by

$$\hat{\alpha}_t = \hat{\alpha}_{t-1} + \mathbf{K}_t\left(\mathbf{c}_{t-1}^*(\hat{\alpha}_{t-1})M_t(\hat{\boldsymbol{\beta}}_{t-1}, \hat{\alpha}_{t-1})\right),$$

$$\mathbf{K}_t = \mathbf{K}_{t-1}\left(\mathbf{I}_s - \left(\mathbf{c}_{t-1}^*(\hat{\alpha}_{t-1})\frac{\partial M_t(\hat{\boldsymbol{\beta}}_{t-1}, \hat{\alpha}_{t-1})}{\partial \boldsymbol{\alpha}'} + \frac{\partial \mathbf{c}_{t-1}^*(\hat{\alpha}_{t-1})}{\partial \boldsymbol{\alpha}}M_t(\hat{\boldsymbol{\beta}}_{t-1}, \hat{\alpha}_{t-1})\right)\mathbf{K}_{t-1}\right)^{-1},$$

Example. We consider the simple linear regression model with the ARCH (1) error defined as $y_t = \beta y_{t-1} + \varepsilon_t$ where $E\left[\varepsilon_t | \mathcal{F}_{t-1}^y\right] = 0$, and $\text{Var}\left(\varepsilon_t | \mathcal{F}_{t-1}^y\right) = h_t = \alpha_0 + \alpha_1 \varepsilon_{t-1}^2$. The parameter of interest are $(\beta, \alpha') = (\beta, \alpha_0, \alpha_1)$. If we assume $\gamma = 0$ and $\kappa = 0$, then

$$a_{t-1}^* = -\frac{x_t}{h_t}, b_{t-1}^* = \frac{\alpha_1 \varepsilon_{t-1} x_{t-1}}{h_t^2}, \mathbf{c}_{t-1}^* = -\frac{1}{h_t^2}(1, \varepsilon_{t-1}^2),$$

and

$$\frac{\partial a_{t-1}^*}{\partial \beta} = \frac{2\alpha_1 x_t \varepsilon_{t-1} x_{t-1}}{h_t^2}, \frac{\partial b_{t-1}^*}{\partial \beta} = \frac{4\alpha_1^2 \varepsilon_{t-1}^2 x_{t-1}^2}{h_t^3} - \frac{\alpha_1 x_{t-1}^2}{h_t^2}, \frac{\partial \mathbf{c}_{t-1}^*}{\partial \alpha} = \frac{2}{h_t^3}(1, \varepsilon_{t-1}^2)'(1, \varepsilon_{t-1}^2).$$

Also, the derivatives of m_t and M_t with respect to β and α are given by

$$\frac{\partial m_t}{\partial \beta} = -x_t, \frac{\partial M_t}{\partial \beta} = 2(y_t - \beta x_t)x_t + 2\alpha_1 \varepsilon_{t-1} x_{t-1}, \frac{\partial m_t}{\alpha} = \mathbf{0}, \frac{\partial m_t}{\alpha} = (1, \varepsilon_{t-1}^2)'.$$

Then the recursive equations for β and α are derived as

$$\hat{\beta}_t = \hat{\beta}_{t-1} + K_t \left(-\frac{x_t(y_t - \hat{\beta}_t x_t)}{h_t(\hat{\alpha}_{t-1})} + \frac{\hat{\alpha}_{1,t-1} \varepsilon_{t-1} x_{t-1}((y_t - \hat{\beta}_t x_t)^2 - h_t(\hat{\alpha}_{t-1}))}{h_t(\hat{\alpha}_{t-1})} \right),$$

$$K_t = K_{t-1} \left(1 - \left(\frac{x_t^2}{h_t(\hat{\alpha}_{t-1})} + \frac{2\hat{\alpha}_{1,t-1} x_t \varepsilon_{t-1} x_{t-1}(y_t - \hat{\beta}_t x_t)}{h_t(\hat{\alpha}_{t-1})} \right. \right.$$

$$+ \frac{\hat{\alpha}_{1,t-1} \varepsilon_{t-1} x_{t-1}(2(y_t - \hat{\beta}_t x_t)x_t + 2\hat{\alpha}_{1,t-1} \varepsilon_{t-1} x_{t-1})}{h_t^2(\hat{\alpha}_{t-1})}$$

$$\left. \left. + \frac{(4\hat{\alpha}_{1,t-1}^2 \varepsilon_{t-1}^2 x_{t-1}^2 - \hat{\alpha}_{1,t-1} x_{t-1}^2 h_t(\hat{\alpha}_{t-1}))((y_t - \hat{\beta}_t x_t)^2 - h_t(\hat{\alpha}_{t-1}))}{h_t^3(\hat{\alpha}_{t-1})} \right) K_{t-1} \right)^{-1}$$

and

$$\hat{\alpha}_t = \hat{\alpha}_{t-1} - \mathbf{K}_t \frac{(y_t - \hat{\beta}_t x_t)^2 - h_t(\hat{\alpha}_{t-1})}{h_t^2(\hat{\alpha}_{t-1})}(1, \varepsilon_{t-1}^2)',$$

$$\mathbf{K}_t = \mathbf{K}_{t-1} \left(\mathbf{I}_2 - \left(-\frac{1}{h_t^2(\hat{\alpha}_{t-1})}(1, \varepsilon_{t-1}^2)'(1, \varepsilon_{t-1}^2) \right. \right.$$

$$\left. \left. + \frac{2((y_t - \hat{\beta}_t x_t)^2 - h_t(\hat{\alpha}_{t-1}))}{h_t^2(\hat{\alpha}_{t-1})}(1, \varepsilon_{t-1}^2)'(1, \varepsilon_{t-1}^2) \right) \mathbf{K}_{t-1} \right)^{-1}.$$

3.5 RCA Models with GARCH Errors

For the RCA model with GARCH errors of the form

$$y_t = (\theta + b_t)y_{t-1} + \sqrt{h_t}\varepsilon_t$$

$$h_t = \omega + \sum_{j=1}^p \alpha_j \varepsilon_{t-j}^2 + \sum_{j=1}^q \beta_j h_{t-j}, \tag{22}$$

where $\{b_t\}$ and $\{\varepsilon_t\}$ are uncorrelated zero mean processes with unknown variance $\delta = \sigma_b^2$ and variance $\sigma_\varepsilon^2(\theta)$ with unknown parameter θ, respectively. Further, we denote the skewness and excess kurtosis of $\{b_t\}$ by γ_b, κ_b which are known, and of $\{\varepsilon_t\}$ by $\gamma_\varepsilon(\theta), \kappa_\varepsilon(\theta)$, respectively. In this model, the conditional mean is $\mu_t = y_{t-1}\theta$ and the conditional variance is $\sigma_t^2 = y_{t-1}^2 \delta + h_t \sigma_\varepsilon^2(\theta)$. The parameter θ appears simultaneously in the mean and variance. Let $m_t = y_t - \mu_t$ and $s_t = m_t^2 - \sigma_t^2$ such that $\langle m \rangle_t = y_{t-1}^2 \delta + h_t \sigma_\varepsilon^2$, $\langle M \rangle_t = y_{t-1}^4 \delta^2(\kappa_b + 2) + h_t^2 \sigma_\varepsilon^4(\kappa_\varepsilon + 2) + 4y_{t-1}^2 \delta h_t \sigma_\varepsilon^2$, $\langle m, M \rangle_t = y_{t-1}^3 \delta^{3/2} \gamma_b + h_t^{3/2} \sigma_\varepsilon^3 \gamma_\varepsilon$, the conditional skewness is $\gamma_t = \langle m, M \rangle_t / \sigma_t^3$, and the conditional excess kurtosis is $\kappa_t = \langle M \rangle_t / \sigma_t^4 - 2$. Then It follows from Theorem 1 that the optimal component quadratic estimating function for the parameter vector $\theta = (\theta, \delta, \omega, \alpha_1, \cdots, \alpha_p, \beta_1, \cdots, \beta_q)'$ is given by

$$\mathbf{g}_C^*(\theta) = \sum_{t=1}^n \left(\mathbf{a}_{t-1}^* m_t + \mathbf{b}_{t-1}^* M_t \right), \tag{23}$$

where

$$
\mathbf{a}^*_{t-1} = \left(1 - \frac{\langle m, M\rangle_t^2}{\langle m\rangle_t \langle M\rangle_t}\right)^{-1} \left(-\frac{y_{t-1}}{\langle m\rangle_t} + \left(h_t \frac{\partial \sigma_\varepsilon^2}{\partial \theta} + \sigma_\varepsilon^2 \frac{\partial h_t}{\partial \theta}\right) \frac{\langle m, M\rangle_t}{\langle m\rangle_t \langle M\rangle_t},\right.
$$
$$
y_{t-1}^2 \frac{\langle m, M\rangle_t}{\langle m\rangle_t \langle M\rangle_t}, \sigma_\varepsilon^2 \frac{\partial h_t}{\partial \omega} \frac{\langle m, M\rangle_t}{\langle m\rangle_t \langle M\rangle_t}, \sigma_\varepsilon^2 \frac{\partial h_t}{\partial \alpha_1} \frac{\langle m, M\rangle_t}{\langle m\rangle_t \langle M\rangle_t}, \cdots, \sigma_\varepsilon^2 \frac{\partial h_t}{\partial \alpha_p} \frac{\langle m, M\rangle_t}{\langle m\rangle_t \langle M\rangle_t},
$$
$$
\left.\sigma_\varepsilon^2 \frac{\partial h_t}{\partial \beta_1} \frac{\langle m, M\rangle_t}{\langle m\rangle_t \langle M\rangle_t}, \cdots, \sigma_\varepsilon^2 \frac{\partial h_t}{\partial \beta_q} \frac{\langle m, M\rangle_t}{\langle m\rangle_t \langle M\rangle_t}\right)'
$$

and

$$
\mathbf{b}^*_{t-1} = \left(1 - \frac{\langle m, M\rangle_t^2}{\langle m\rangle_t \langle M\rangle_t}\right)^{-1} \left(\frac{y_{t-1}\langle m, M\rangle_t}{\langle m\rangle_t \langle M\rangle_t} - \left(h_t \frac{\partial \sigma_\varepsilon^2}{\partial \theta} + \sigma_\varepsilon^2 \frac{\partial h_t}{\partial \theta}\right) \frac{1}{\langle M\rangle_t},\right.
$$
$$
- y_{t-1}^2 \frac{1}{\langle M\rangle_t}, -\sigma_\varepsilon^2 \frac{\partial h_t}{\partial \omega} \frac{1}{\langle M\rangle_t}, -\sigma_\varepsilon^2 \frac{\partial h_t}{\partial \alpha_1} \frac{1}{\langle M\rangle_t}, \cdots, -\sigma_\varepsilon^2 \frac{\partial h_t}{\partial \alpha_p} \frac{1}{\langle M\rangle_t},
$$
$$
\left. -\sigma_\varepsilon^2 \frac{\partial h_t}{\partial \beta_1} \frac{1}{\langle M\rangle_t}, \cdots, -\sigma_\varepsilon^2 \frac{\partial h_t}{\partial \beta_q} \frac{1}{\langle M\rangle_t}\right)'.
$$

The calculation of $\frac{\partial h_t}{\partial \boldsymbol{\theta}} = (\frac{\partial h_t}{\partial \theta}, 0, \frac{\partial h_t}{\partial \omega}, \frac{\partial h_t}{\partial \alpha_1}, \cdots, \frac{\partial h_t}{\partial \alpha_p}, \frac{\partial h_t}{\partial \beta_1}, \cdots, \frac{\partial h_t}{\partial \beta})'_q$ is not straight forward and we the recursive form is necessary to take into account. Computation of $\frac{\partial h_t}{\partial \boldsymbol{\theta}}$ yields:

$$
\frac{\partial h_t}{\partial \theta} = -2\sum_{j=1}^{p} \alpha_j y_{t-j-1}\varepsilon_{t-j} h_{t-j}^{-1/2} - \sum_{j=1}^{p} \alpha_j \varepsilon_{t-j}^2 h_{t-j}^{-1}\frac{\partial h_{t-j}}{\partial \theta} + \sum_{j=1}^{q} \beta_j \frac{\partial h_{t-j}}{\partial \theta}
$$

$$
\frac{\partial h_t}{\partial \omega} = 1 - \sum_{j=1}^{p} \alpha_j \varepsilon_{t-j}^2 h_{t-j}^{-1}\frac{\partial h_{t-j}}{\partial \omega} + \sum_{j=1}^{q} \beta_j \frac{\partial h_{t-j}}{\partial \omega}
$$

$$
\frac{\partial h_t}{\partial \alpha_j} = \varepsilon_{t-j}^2 - \sum_{j=1}^{p} \alpha_j \varepsilon_{t-j}^2 h_{t-j}^{-1}\frac{\partial h_{t-j}}{\partial \alpha_j} + \sum_{j=1}^{q} \beta_j \frac{\partial h_{t-j}}{\partial \alpha_j}
$$

$$
\frac{\partial h_t}{\partial \beta_j} = h_{t-j} - \sum_{j=1}^{p} \alpha_j \varepsilon_{t-j}^2 h_{t-j}^{-1}\frac{\partial h_{t-j}}{\partial \beta_j} + \sum_{j=1}^{q} \beta_j \frac{\partial h_{t-j}}{\partial \beta_j}.
$$

Example. We consider the RCA model GARCH (1, 1) error defined as

$$
y_t = (\theta + b_t)y_{t-1} + \sqrt{h_t}\varepsilon_t
$$
$$
h_t = \omega + \alpha_1 \varepsilon_{t-1}^2 + \beta_1 h_{t-1}, \tag{24}
$$

where $\{b_t\}$ and $\{\varepsilon_t\}$ are Gaussian processes with $\gamma_b = 0, \kappa_b = 0, \gamma_\varepsilon(\theta) = 0$ and $\kappa_\varepsilon(\theta) = 0$. In order to estimate the parameter vector $\boldsymbol{\theta} = (\theta, \delta, \omega, \alpha_1, \beta_1)$, we have

$$
\mathbf{a}^*_{t-1} = \left(-\frac{y_{t-1}}{y_{t-1}^2 \delta + h_t \sigma_\varepsilon^2}, 0, 0, 0, 0\right)'
$$

and

$$
\mathbf{b}^*_{t-1} = (b^*_{t-1,1}, b^*_{t-1,2}, b^*_{t-1,3}, b^*_{t-1,4}, b^*_{t-1,5})
$$
$$
= -\frac{1}{2(y_{t-1}^2 \delta + h_t \sigma_\varepsilon^2)^2}\left(h_t \frac{\partial \sigma_\varepsilon^2}{\partial \theta} + \sigma_\varepsilon^2 \frac{\partial h_t}{\partial \theta}, y_{t-1}^2, \sigma_\varepsilon^2 \frac{\partial h_t}{\partial \omega}, \sigma_\varepsilon^2 \frac{\partial h_t}{\partial \alpha_1}, \sigma_\varepsilon^2 \frac{\partial h_t}{\partial \beta_1}\right)'.
$$

Then takeing the derivative of $\mathbf{a}^*_{t-1}$ and $\mathbf{b}^*_{t-1}$ with respect to $\boldsymbol{\theta}$ we obtain

$$
\frac{\partial \mathbf{a}^*_{t-1}}{\partial \boldsymbol{\theta}} = \frac{y_{t-1}}{(y_{t-1}^2 \delta + h_t \sigma_\varepsilon^2)^2}\left(\begin{array}{ccccc} \left(h_t \frac{\partial \sigma_\varepsilon^2}{\partial \theta} + \sigma_\varepsilon^2 \frac{\partial h_t}{\partial \theta}\right) & y_{t-1}^2 & \sigma_\varepsilon^2 \frac{\partial h_t}{\partial \omega} & \sigma_\varepsilon^2 \frac{\partial h_t}{\partial \alpha_1} & \sigma_\varepsilon^2 \frac{\partial h_t}{\partial \beta_1} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} \end{array}\right),
$$

and

$$
\frac{\partial \mathbf{b}^*_{t-1}}{\partial \boldsymbol{\theta}} = \frac{1}{(y_{t-1}^2 \delta + h_t \sigma_\varepsilon^2)^3}(B_{ij})_{1\le i\le 5, 1\le j\le 5},
$$

where

$$B_{11} = \frac{1}{2}\left(-(y_{t-1}^2\delta + h_t\sigma_\varepsilon^2)\left(h_t\frac{\partial^2\sigma_\varepsilon^2}{\partial\theta^2} + 2\frac{\partial h_t}{\partial\theta}\frac{\partial\sigma_\varepsilon^2}{\partial\theta} + \sigma_\varepsilon^2\frac{\partial^2 h_t}{\partial\theta^2}\right) + 2\left(h_t\frac{\partial\sigma_\varepsilon^2}{\partial\theta} + \sigma_\varepsilon^2\frac{\partial h_t}{\partial\theta}\right)^2\right)$$

$$B_{12} = y_{t-1}^2\left(h_t\frac{\partial\sigma_\varepsilon^2}{\partial\theta} + \sigma_\varepsilon^2\frac{\partial h_t}{\partial\theta}\right) = B_{21}$$

$$B_{13} = \frac{1}{2}\left(-(y_{t-1}^2\delta + h_t\sigma_\varepsilon^2)\left(\frac{\partial h_t}{\partial\omega}\frac{\partial\sigma_\varepsilon^2}{\partial\theta} + \sigma_\varepsilon^2\frac{\partial^2 h_t}{\partial\theta\partial\omega}\right) + 2\left(h_t\frac{\partial\sigma_\varepsilon^2}{\partial\theta} + \sigma_\varepsilon^2\frac{\partial h_t}{\partial\theta}\right)\sigma_\varepsilon^2\frac{\partial h_t}{\partial\omega}\right) = B_{31}$$

$$B_{14} = \frac{1}{2}\left(-(y_{t-1}^2\delta + h_t\sigma_\varepsilon^2)\left(\frac{\partial h_t}{\partial\alpha_1}\frac{\partial\sigma_\varepsilon^2}{\partial\theta} + \sigma_\varepsilon^2\frac{\partial^2 h_t}{\partial\theta\partial\alpha_1}\right) + 2\left(h_t\frac{\partial\sigma_\varepsilon^2}{\partial\theta} + \sigma_\varepsilon^2\frac{\partial h_t}{\partial\theta}\right)\sigma_\varepsilon^2\frac{\partial h_t}{\partial\alpha_1}\right) = B_{41}$$

$$B_{15} = \frac{1}{2}\left(-(y_{t-1}^2\delta + h_t\sigma_\varepsilon^2)\left(\frac{\partial h_t}{\partial\beta_1}\frac{\partial\sigma_\varepsilon^2}{\partial\theta} + \sigma_\varepsilon^2\frac{\partial^2 h_t}{\partial\theta\partial\beta_1}\right) + 2\left(h_t\frac{\partial\sigma_\varepsilon^2}{\partial\theta} + \sigma_\varepsilon^2\frac{\partial h_t}{\partial\theta}\right)\sigma_\varepsilon^2\frac{\partial h_t}{\partial\beta_1}\right) = B_{51}$$

$$B_{22} = y_{t-1}^4$$

$$B_{23} = y_{t-1}^2\sigma_\varepsilon^2\frac{\partial h_t}{\partial\omega} = B_{32}$$

$$B_{24} = y_{t-1}^2\sigma_\varepsilon^2\frac{\partial h_t}{\partial\alpha_1} = B_{42}$$

$$B_{25} = y_{t-1}^2\sigma_\varepsilon^2\frac{\partial h_t}{\partial\beta_1} = B_{52}$$

$$B_{33} = \frac{1}{2}\left(-(y_{t-1}^2\delta + h_t\sigma_\varepsilon^2)\sigma_\varepsilon^2\frac{\partial^2 h_t}{\partial\omega^2} + 2\sigma_\varepsilon^4\left(\frac{\partial h_t}{\partial\omega}\right)^2\right)$$

$$B_{34} = \frac{1}{2}\left(-(y_{t-1}^2\delta + h_t\sigma_\varepsilon^2)\sigma_\varepsilon^2\frac{\partial^2 h_t}{\partial\omega\partial\alpha_1} + 2\sigma_\varepsilon^4\frac{\partial h_t}{\partial\omega}\frac{\partial h_t}{\partial\alpha_1}\right) = B_{43}$$

$$B_{35} = \frac{1}{2}\left(-(y_{t-1}^2\delta + h_t\sigma_\varepsilon^2)\sigma_\varepsilon^2\frac{\partial^2 h_t}{\partial\omega\partial\beta_1} + 2\sigma_\varepsilon^4\frac{\partial h_t}{\partial\omega}\frac{\partial h_t}{\partial\beta_1}\right) = B_{53}$$

$$B_{44} = \frac{1}{2}\left(-(y_{t-1}^2\delta + h_t\sigma_\varepsilon^2)\sigma_\varepsilon^2\frac{\partial^2 h_t}{\partial\alpha_1^2} + 2\sigma_\varepsilon^4\left(\frac{\partial h_t}{\partial\alpha_1}\right)^2\right)$$

$$B_{45} = \frac{1}{2}\left(-(y_{t-1}^2\delta + h_t\sigma_\varepsilon^2)\sigma_\varepsilon^2\frac{\partial^2 h_t}{\partial\alpha_1\partial\beta_1} + 2\sigma_\varepsilon^4\frac{\partial h_t}{\partial\alpha_1}\frac{\partial h_t}{\partial\beta_1}\right) = B_{54}$$

$$B_{55} = \frac{1}{2}\left(-(y_{t-1}^2\delta + h_t\sigma_\varepsilon^2)\sigma_\varepsilon^2\frac{\partial^2 h_t}{\partial\partial\beta_1^2} + 2\sigma_\varepsilon^4\left(\frac{\partial h_t}{\partial\beta_1}\right)^2\right)$$

Also, since

$$\frac{\partial m_t}{\partial\boldsymbol{\theta}} = (-y_t, 0, 0, 0, 0)'$$

and

$$\frac{\partial M_t}{\partial\boldsymbol{\theta}} = \left(-2(y_t - (\theta + b_t)y_{t-1})y_{t-1} - \left(h_t\frac{\partial\sigma_\varepsilon^2}{\partial\theta} + \sigma_\varepsilon^2\frac{\partial h_t}{\partial\theta}\right), -y_{t-1}^2, -\sigma_\varepsilon^2\frac{\partial h_t}{\partial\omega}, -\sigma_\varepsilon^2\frac{\partial h_t}{\partial\alpha_1}, -\sigma_\varepsilon^2\frac{\partial h_t}{\partial\beta_1}\right)'$$

the recursive estimate of $\boldsymbol{\theta}$ are given by

$$\hat{\boldsymbol{\theta}}_t = \hat{\boldsymbol{\theta}}_{t-1} + \mathbf{K}_t\left(\mathbf{a}_{t-1}^*(\hat{\boldsymbol{\theta}}_{t-1})m_t(\hat{\boldsymbol{\theta}}_{t-1}) + \mathbf{b}_{t-1}^*(\hat{\boldsymbol{\theta}}_{t-1})M_t(\hat{\boldsymbol{\theta}}_{t-1})\right),$$

$$\mathbf{K}_t = \mathbf{K}_{t-1}\left(\mathbf{I}_5 - \left(\mathbf{a}_{t-1}^*(\hat{\boldsymbol{\theta}}_{t-1})\frac{\partial m_t(\hat{\boldsymbol{\theta}}_{t-1})}{\partial\boldsymbol{\theta}'} + \frac{\partial\mathbf{a}_{t-1}^*(\hat{\boldsymbol{\theta}}_{t-1})}{\partial\boldsymbol{\theta}}m_t(\hat{\boldsymbol{\theta}}_{t-1})\right.\right.$$

$$\left.\left. + \mathbf{b}_{t-1}^*(\hat{\boldsymbol{\theta}}_{t-1})\frac{\partial M_t(\hat{\boldsymbol{\theta}}_{t-1})}{\partial\boldsymbol{\theta}'} + \frac{\partial\mathbf{b}_{t-1}^*(\hat{\boldsymbol{\theta}}_{t-1})}{\partial\boldsymbol{\theta}}M_t(\hat{\boldsymbol{\theta}}_{t-1})\right)\mathbf{K}_{t-1}\right)^{-1}.$$

4. Appendix

Proof of Lemma 4. Method 1 (General estimating functions). For the class of unbiased estimating function $\mathcal{G} = \{g : g = \alpha_1 g_1 + \alpha_2 g_2\}$, we want to derive the coefficients α_1 and α_2 such that $\min_{\alpha_1, \alpha_2} E[S - \alpha_1 g_1 - \alpha_2 g_2]^2$ is obtained. Take the derivative with respect to α_1 and α_2 we have

$$E[(S - \alpha_1 g_1 - \alpha_2 g_2) g_1] = 0, E[(S - \alpha_1 g_1 - \alpha_2 g_2) g_2] = 0$$

which are equivalent to

$$Cov(S, g_1) - \alpha_1 Var(g_1) - \alpha_2 Cov(g_1, g_2) = 0 \tag{25}$$

$$Cov(S, g_2) - \alpha_1 Cov(g_1, g_2) - \alpha_2 Var(g_2) = 0. \tag{26}$$

From equation (3) we have

$$\alpha_1 = \frac{Cov(S, g_2) - \alpha_2 Var(g_2)}{Cov(g_1, g_2)}.$$

Substituting α_1 to equation (2) we have

$$\alpha_2 = \frac{Cov(S, g_1) - \frac{Cov(S, g_2) Var(g_1)}{Cov(g_1, g_2)}}{\frac{Var(g_1) Var(g_2)}{Cov(g_1, g_2)} - Cov(g_1, g_2)}$$

$$= \frac{Cov(S, g_1) Cov(g_1, g_2) - Cov(S, g_2) Var(g_1)}{Var(g_1) Var(g_2) - Cov^2(g_1, g_2)}$$

$$= \frac{E\left[\frac{\partial g_2}{\partial \theta}\right] Var(g_1) - E\left[\frac{\partial g_1}{\partial \theta}\right] Cov(g_1, g_2)}{Var(g_1) Var(g_2) - Cov^2(g_1, g_2)}.$$

Similarly,

$$\alpha_1 = \frac{E\left[\frac{\partial g_1}{\partial \theta}\right] Var(g_2) - E\left[\frac{\partial g_2}{\partial \theta}\right] Cov(g_1, g_2)}{Var(g_1) Var(g_2) - Cov^2(g_1, g_2)}.$$

Alternatively, the optimal combined estimating function based on the orthogonal estimating functions g_1 and $\psi_2 = g_2 - \frac{Cov(g_1, g_2)}{Var(g_1)} g_1$ is given by

$$g_C = \frac{E\left[\frac{\partial g_1}{\partial \theta}\right]}{Var(g_1)} g_1 + \frac{E\left[\frac{\partial g_2}{\partial \theta}\right] - \frac{Cov(g_1, g_2)}{Var(g_1)} E\left[\frac{\partial g_1}{\partial \theta}\right]}{Var(g_2) + \frac{Cov^2(g_1, g_2)}{Var(g_1)} - 2\frac{Cov^2(g_1, g_2)}{Var(g_1)}} \left(g_2 - \frac{Cov(g_1, g_2)}{Var(g_1)} g_1\right).$$

Expanding the above equation, the coefficients for g_1 and g_2 are given by

$$\alpha_2 = \frac{E\left[\frac{\partial g_2}{\partial \theta}\right] Var(g_1) - E\left[\frac{\partial g_1}{\partial \theta}\right] Cov(g_1, g_2)}{Var(g_1) Var(g_2) - Cov^2(g_1, g_2)},$$

and

$$\alpha_1 = \frac{E\left[\frac{\partial g_1}{\partial \theta}\right]}{Var(g_1)} - \frac{E\left[\frac{\partial g_2}{\partial \theta}\right] Var(g_1) - E\left[\frac{\partial g_1}{\partial \theta}\right] Cov(g_1, g_2)}{Var(g_1) Var(g_2) - Cov^2(g_1, g_2)} \times \frac{Cov(g_1, g_2)}{Var(g_1)}$$

$$= \frac{E\left[\frac{\partial g_1}{\partial \theta}\right] Var(g_2) - E\left[\frac{\partial g_2}{\partial \theta}\right] Cov(g_1, g_2)}{Var(g_1) Var(g_2) - Cov^2(g_1, g_2)}.$$

$\square$

Proof of Theorem 1. We choose two orthogonal martingale differences $m_i(\theta) = x_i - \mu_i(\theta)$ and $\psi_i(\theta) = M_i(\theta) - \frac{\langle m, M \rangle_i}{\langle m \rangle_i} m_i(\theta)$, where the conditional variance of $\psi_i(\theta)$ is given by $\langle \psi \rangle_i = \langle M \rangle_i - \frac{\langle m, M \rangle_i^2}{\langle m \rangle_i}$. That is, $m_i(\theta)$ and $\psi_i(\theta)$ are uncorrelated with conditional variance $\langle m \rangle_i$ and $\langle \psi \rangle_i$, respectively. Moreover, the optimal martingale estimating function and associated

information based on the martingale differences $\psi_i(\boldsymbol{\theta})$ are

$$
\begin{aligned}
\mathbf{g}_\Psi^*(\boldsymbol{\theta}) &= \sum_{t=1}^{n} \left(\frac{\partial \mu_i}{\partial \boldsymbol{\theta}} \frac{\langle m, M \rangle_i}{\langle m \rangle_i} + \mathrm{E}\left[\frac{\partial M_i}{\partial \boldsymbol{\theta}} \middle| \mathscr{F}_{i-1}^x \right] \right) \frac{\psi_i}{\langle \psi \rangle_i} \\
&= \sum_{t=1}^{n} \left(1 - \frac{\langle m, M \rangle_i^2}{\langle m \rangle_i \langle M \rangle_i} \right)^{-1} \left(\left(-\frac{\partial \mu_i}{\partial \boldsymbol{\theta}} \frac{\langle m, M \rangle_i^2}{\langle m \rangle_i^2 \langle M \rangle_i} - \mathrm{E}\left[\frac{\partial M_i}{\partial \boldsymbol{\theta}} \middle| \mathscr{F}_{t-1}^y \right] \frac{\langle m, M \rangle_i}{\langle m \rangle_i \langle M \rangle_i} \right) m_i \right. \\
&\quad \left. + \left(\frac{\partial \mu_i}{\partial \boldsymbol{\theta}} \frac{\langle m, M \rangle_i}{\langle m \rangle_i \langle M \rangle_i} + \mathrm{E}\left[\frac{\partial M_i}{\partial \boldsymbol{\theta}} \middle| \mathscr{F}_{i-1}^x \right] \frac{1}{\langle M \rangle_i} \right) M_i \right)
\end{aligned}
$$

and

$$
\begin{aligned}
\mathbf{I}_{\mathbf{g}_\Psi^*}(\boldsymbol{\theta}) &= \sum_{t=1}^{n} \left(\frac{\partial \mu_i}{\partial \boldsymbol{\theta}'} \frac{\langle m, M \rangle_i}{\langle m \rangle_i} + \mathrm{E}\left[\frac{\partial M_i}{\partial \boldsymbol{\theta}} \middle| \mathscr{F}_{i-1}^x \right] \right) \left(\frac{\partial \mu_i}{\partial \boldsymbol{\theta}'} \frac{\langle m, M \rangle_i}{\langle m \rangle_i} + \mathrm{E}\left[\frac{\partial M_i}{\partial \boldsymbol{\theta}'} \middle| \mathscr{F}_{i-1}^x \right] \right) \frac{1}{\langle \psi \rangle_i} \\
&= \sum_{t=1}^{n} \left(1 - \frac{\langle m, M \rangle_i^2}{\langle m \rangle_i \langle M \rangle_i} \right)^{-1} \left(\frac{\partial \mu_i}{\partial \boldsymbol{\theta}} \frac{\partial \mu_i}{\partial \boldsymbol{\theta}'} \frac{\langle m, M \rangle_i^2}{\langle m \rangle_i^2 \langle M \rangle_i} + \mathrm{E}\left[\frac{\partial M_i}{\partial \boldsymbol{\theta}} \frac{\partial M_i}{\partial \boldsymbol{\theta}'} \middle| \mathscr{F}_{i-1}^x \right] \frac{1}{\langle M \rangle_i} \right. \\
&\quad \left. + \left(\frac{\partial \mu_i}{\partial \boldsymbol{\theta}} \frac{\partial M_i}{\partial \boldsymbol{\theta}'} + \frac{\partial M_i}{\partial \boldsymbol{\theta}} \frac{\partial \mu_i}{\partial \boldsymbol{\theta}'} \right) \frac{\langle m, M \rangle_i}{\langle m \rangle_i \langle M \rangle_i} \right).
\end{aligned}
$$

Then the combined estimating function based on m_i and ψ_i becomes

$$
\begin{aligned}
\mathbf{g}_C^*(\boldsymbol{\theta}) &= \sum_{t=1}^{n} \left(1 - \frac{\langle m, M \rangle_i^2}{\langle m \rangle_i \langle M \rangle_i} \right)^{-1} \left(\left(-\frac{\partial \mu_i}{\partial \boldsymbol{\theta}} \frac{1}{\langle m \rangle_i} - \mathrm{E}\left[\frac{\partial M_i}{\partial \boldsymbol{\theta}} \middle| \mathscr{F}_{i-1}^x \right] \frac{\langle m, M \rangle_i}{\langle m \rangle_i \langle M \rangle_i} \right) m_i \right. \\
&\quad \left. + \left(\frac{\partial \mu_i}{\partial \boldsymbol{\theta}} \frac{\langle m, M \rangle_i}{\langle m \rangle_i \langle M \rangle_i} + \mathrm{E}\left[\frac{\partial M_i}{\partial \boldsymbol{\theta}} \middle| \mathscr{F}_{i-1}^x \right] \frac{1}{\langle M \rangle_i} \right) M_i \right),
\end{aligned}
$$

and satisfies the sufficient condition for optimality

$$
\mathrm{E}\left[\frac{\partial \mathbf{g}_C(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}} \middle| \mathscr{F}_{i-1}^x \right] = \mathrm{Cov}(\mathbf{g}_C(\boldsymbol{\theta}), \mathbf{g}_C^*(\boldsymbol{\theta}) \mid \mathscr{F}_{i-1}^x) \, K, \ \forall \mathbf{g}_C(\boldsymbol{\theta}) \in \mathscr{G}_C,
$$

where K is a constant matrix. Hence, $\mathbf{g}_C^*(\boldsymbol{\theta})$ is optimal in the class $\mathscr{G}_C$, and part (a) follows. Since m_i and ψ_i are orthogonal, the information $\mathbf{I}_{\mathbf{g}_C^*}(\boldsymbol{\theta}) = \mathbf{I}_{\mathbf{g}_m^*}(\boldsymbol{\theta}) + \mathbf{I}_{\mathbf{g}_\Psi^*}(\boldsymbol{\theta})$ and part (b) follows. Hence, neither $\mathbf{g}_m^*(\boldsymbol{\theta})$ nor $\mathbf{g}_M^*(\boldsymbol{\theta})$ is fully informative, that is, $\mathbf{I}_{\mathbf{g}_C^*}(\boldsymbol{\theta}) \geq \mathbf{I}_{\mathbf{g}_m^*}(\boldsymbol{\theta})$ and $\mathbf{I}_{\mathbf{g}_C^*}(\boldsymbol{\theta}) \geq \mathbf{I}_{\mathbf{g}_M^*}(\boldsymbol{\theta})$. Part (c) follows immediately from the fact that if $\frac{\partial \mu_i}{\partial \boldsymbol{\theta}} \frac{\langle m, M \rangle_i}{\langle m \rangle_i} = -\mathrm{E}\left[\frac{\partial M_i}{\partial \boldsymbol{\theta}} \middle| \mathscr{F}_{i-1}^x \right]$ for each i, then $\mathbf{g}_\Psi^*(\boldsymbol{\theta}) = \mathbf{0}$ and $\mathbf{I}_{\mathbf{g}_\Psi^*}(\boldsymbol{\theta}) = \mathbf{0}$.

Proof of part (d) is as follows. The optimal combined estimating function based on $m_i((\boldsymbol{\theta}))$ and $M_i((\boldsymbol{\theta}))$ is given by

$$
\mathbf{g}^*(\boldsymbol{\theta}) = \sum_{i=1}^{n} \mathbf{a}_{i-1}^*(\boldsymbol{\theta}) m_i(\boldsymbol{\theta}) + \mathbf{b}_{i-1}^*(\boldsymbol{\theta}) M_i(\boldsymbol{\theta}).
$$

Then using Taylor's expansion for $\mathbf{g}_n^*(\boldsymbol{\theta})$ we have

$$
\begin{aligned}
&\sum_{i=1}^{n} \mathbf{a}_{i-1}^*(\hat{\boldsymbol{\theta}}) m_i(\hat{\boldsymbol{\theta}}) + \mathbf{b}_{i-1}^*(\hat{\boldsymbol{\theta}}) M_i(\hat{\boldsymbol{\theta}}) + \left(\sum_{i=1}^{n} \mathbf{a}_{i-1}^*(\hat{\boldsymbol{\theta}}) \frac{\partial m_i(\hat{\boldsymbol{\theta}})}{\partial \boldsymbol{\theta}} + \frac{\partial \mathbf{a}_{i-1}^*(\hat{\boldsymbol{\theta}})}{\partial \boldsymbol{\theta}} M_i(\hat{\boldsymbol{\theta}}) \right. \\
&\quad \left. + \mathbf{b}_{i-1}^*(\hat{\boldsymbol{\theta}}) \frac{\partial m_i(\hat{\boldsymbol{\theta}})}{\partial \boldsymbol{\theta}} + \frac{\partial \mathbf{b}_{i-1}^*(\hat{\boldsymbol{\theta}})}{\partial \boldsymbol{\theta}} M_i(\hat{\boldsymbol{\theta}}) \right) (\boldsymbol{\theta} - \hat{\boldsymbol{\theta}}) = 0.
\end{aligned}
$$

Substituting the recursive estimate for θ at each step, the estimate based on the first $i - 1$ observations is given by

$$
\begin{aligned}
\hat{\theta}_{i-1} = & \left\{ \sum_{s=1}^{i-1} \left(\mathbf{a}_{s-1}^*(\hat{\theta}_{s-1}) \frac{\partial m_s(\hat{\theta}_{s-1})}{\partial \theta} + \frac{\partial \mathbf{a}_{s-1}^*(\hat{\theta}_{s-1})}{\partial \theta} m_i(\hat{\theta}_{s-1}) + \mathbf{b}_{s-1}^*(\hat{\theta}_{s-1}) \frac{\partial M_s(\hat{\theta}_{s-1})}{\partial \theta} \right. \right. \\
& \left. \left. + \frac{\partial \mathbf{b}_{s-1}^*(\hat{\theta}_{s-1})}{\partial \theta} M_i(\hat{\theta}_{s-1}) \right) \right\}^{-1} \left(- \sum_{s=1}^{i-1} \left(\mathbf{a}_{s-1}^*(\hat{\theta}_{s-1}) m_s(\hat{\theta}_{s-1}) + \mathbf{b}_{s-1}^*(\hat{\theta}_{s-1}) M_s(\hat{\theta}_{s-1}) \right) \right. \\
& + \sum_{s=1}^{i-1} \left(\mathbf{a}_{s-1}^*(\hat{\theta}_{s-1}) \frac{\partial m_s(\hat{\theta}_{s-1})}{\partial \theta} + \frac{\partial \mathbf{a}_{s-1}^*(\hat{\theta}_{s-1})}{\partial \theta} m_i(\hat{\theta}_{s-1}) + \mathbf{b}_{s-1}^*(\hat{\theta}_{s-1}) \frac{\partial M_s(\hat{\theta}_{s-1})}{\partial \theta} \right. \\
& \left. \left. + \frac{\partial \mathbf{b}_{s-1}^*(\hat{\theta}_{s-1})}{\partial \theta} M_i(\hat{\theta}_{s-1}) \right) \hat{\theta}_{s-1} \right).
\end{aligned}
$$

When the ith observation becomes available, the estimate becomes

$$
\begin{aligned}
\hat{\theta}_{i} = & \left\{ \sum_{s=1}^{i} \left(\mathbf{a}_{s-1}^*(\hat{\theta}_{s-1}) \frac{\partial m_s(\hat{\theta}_{s-1})}{\partial \theta} + \frac{\partial \mathbf{a}_{s-1}^*(\hat{\theta}_{s-1})}{\partial \theta} m_i(\hat{\theta}_{s-1}) + \mathbf{b}_{s-1}^*(\hat{\theta}_{s-1}) \frac{\partial M_s(\hat{\theta}_{s-1})}{\partial \theta} \right. \right. \\
& \left. \left. + \frac{\partial \mathbf{b}_{s-1}^*(\hat{\theta}_{s-1})}{\partial \theta} M_i(\hat{\theta}_{s-1}) \right) \right\}^{-1} \left(- \sum_{s=1}^{i} \left(\mathbf{a}_{s-1}^*(\hat{\theta}_{s-1}) m_s(\hat{\theta}_{s-1}) + \mathbf{b}_{s-1}^*(\hat{\theta}_{s-1}) M_s(\hat{\theta}_{s-1}) \right) \right. \\
& + \sum_{s=1}^{i} \left(\mathbf{a}_{s-1}^*(\hat{\theta}_{s-1}) \frac{\partial m_s(\hat{\theta}_{s-1})}{\partial \theta} + \frac{\partial \mathbf{a}_{s-1}^*(\hat{\theta}_{s-1})}{\partial \theta} m_i(\hat{\theta}_{s-1}) + \mathbf{b}_{s-1}^*(\hat{\theta}_{s-1}) \frac{\partial M_s(\hat{\theta}_{s-1})}{\partial \theta} \right. \\
& \left. \left. + \frac{\partial \mathbf{b}_{s-1}^*(\hat{\theta}_{s-1})}{\partial \theta} M_i(\hat{\theta}_{s-1}) \right) \hat{\theta}_{s-1} \right).
\end{aligned}
$$

Let

$$
\begin{aligned}
\mathbf{K}_i^{-1} = & - \sum_{s=1}^{i} \left(\mathbf{a}_{s-1}^*(\hat{\theta}_{s-1}) \frac{\partial m_s(\hat{\theta}_{s-1})}{\partial \theta} + \frac{\partial \mathbf{a}_{s-1}^*(\hat{\theta}_{s-1})}{\partial \theta} m_i(\hat{\theta}_{s-1}) + \mathbf{b}_{s-1}^*(\hat{\theta}_{s-1}) \frac{\partial M_s(\hat{\theta}_{s-1})}{\partial \theta} \right. \\
& \left. + \frac{\partial \mathbf{b}_{s-1}^*(\hat{\theta}_{s-1})}{\partial \theta} M_i(\hat{\theta}_{s-1}) \right),
\end{aligned}
$$

then

$$
\begin{aligned}
\mathbf{K}_i^{-1} = & \mathbf{K}_{i-1}^{-1} - \left(\mathbf{a}_{s-1}^*(\hat{\theta}_{s-1}) \frac{\partial m_s(\hat{\theta}_{s-1})}{\partial \theta} + \frac{\partial \mathbf{a}_{s-1}^*(\hat{\theta}_{s-1})}{\partial \theta} m_i(\hat{\theta}_{s-1}) + \mathbf{b}_{s-1}^*(\hat{\theta}_{s-1}) \frac{\partial M_s(\hat{\theta}_{s-1})}{\partial \theta} \right. \\
& \left. + \frac{\partial \mathbf{b}_{s-1}^*(\hat{\theta}_{s-1})}{\partial \theta} M_i(\hat{\theta}_{s-1}) \right),
\end{aligned}
$$

and

$$
\begin{aligned}
\hat{\theta}_i - \hat{\theta}_{i-1} &= \mathbf{K}_i \left(\mathbf{K}_i^{-1} \hat{\theta}_i - \mathbf{K}_i^{-1} \hat{\theta}_{i-1} \right) \\
&= \mathbf{K}_i \left(\mathbf{a}_{i-1}^*(\hat{\theta}_{i-1}) m_i(\hat{\theta}_{i-1}) + \mathbf{b}_{i-1}^*(\hat{\theta}_{i-1}) M_i(\hat{\theta}_{i-1}) \right).
\end{aligned}
$$

Hence it is easy to show that the recursive equations for θ take the form (15) - (16). □

For the RCA model with GARCH errors , the information matrix of the optimal quadratic estimating function for θ is given by

$$
\mathbf{I}_{\mathbf{g}_C^*}(\theta) = \begin{pmatrix}
I_{\theta\theta} & I_{\theta\delta} & I_{\theta\omega} & \mathbf{I}_{\theta\alpha'} & \mathbf{I}_{\theta\beta'} \\
I_{\delta\theta} & I_{\delta\delta} & I_{\delta\omega} & \mathbf{I}_{\delta\alpha'} & \mathbf{I}_{\delta\beta'} \\
I_{\omega\theta} & I_{\omega\delta} & I_{\omega\omega} & \mathbf{I}_{\omega\alpha'} & \mathbf{I}_{\omega\beta'} \\
\mathbf{I}_{\alpha\theta} & \mathbf{I}_{\alpha\delta} & \mathbf{I}_{\alpha\omega} & \mathbf{I}_{\alpha\alpha'} & \mathbf{I}_{\alpha\beta'} \\
\mathbf{I}_{\beta\theta} & \mathbf{I}_{\beta\delta} & \mathbf{I}_{\beta\omega} & \mathbf{I}_{\beta\alpha'} & \mathbf{I}_{\beta\beta'}
\end{pmatrix},
$$

where

$$I_{\theta\theta} = \sum_{t=1}^{n} \left(1 - \frac{\langle m, M\rangle_t^2}{\langle m\rangle_t \langle M\rangle_t}\right)^{-1} \left(\frac{y_{t-1}^2}{\langle m\rangle_t} + \left(h_t \frac{\partial \sigma_\varepsilon^2}{\partial \theta} + \sigma_\varepsilon^2 \frac{\partial h_t}{\partial \theta}\right)^2 \frac{1}{\langle M\rangle_t}\right.$$
$$\left. -2y_{t-1}\left(h_t \frac{\partial \sigma_\varepsilon^2}{\partial \theta} + \sigma_\varepsilon^2 \frac{\partial h_t}{\partial \theta}\right) \frac{\langle m, M\rangle_t}{\langle m\rangle_t \langle M\rangle_t}\right),$$

$$I_{\theta\delta} = I_{\delta\theta} = \sum_{t=1}^{n} \left(1 - \frac{\langle m, M\rangle_t^2}{\langle m\rangle_t \langle M\rangle_t}\right)^{-1} \left(\left(h_t \frac{\partial \sigma_\varepsilon^2}{\partial \theta} + \sigma_\varepsilon^2 \frac{\partial h_t}{\partial \theta}\right) \frac{1}{\langle M\rangle_t} - \frac{y_{t-1}\langle m, M\rangle_t}{\langle m\rangle_t \langle M\rangle_t}\right) y_{t-1}^2,$$

$$I_{\delta\delta} = \sum_{t=1}^{n} \left(1 - \frac{\langle m, M\rangle_t^2}{\langle m\rangle_t \langle M\rangle_t}\right)^{-1} \frac{y_{t-1}^4}{\langle M\rangle_t},$$

$$I_{\theta\omega} = I_{\omega\theta} = \sum_{t=1}^{n} \left(1 - \frac{\langle m, M\rangle_t^2}{\langle m\rangle_t \langle M\rangle_t}\right)^{-1} \left(\left(h_t \frac{\partial \sigma_\varepsilon^2}{\partial \theta} + \sigma_\varepsilon^2 \frac{\partial h_t}{\partial \theta}\right) \frac{1}{\langle M\rangle_t} - \frac{y_{t-1}\langle m, M\rangle_t}{\langle m\rangle_t \langle M\rangle_t}\right) \sigma_\varepsilon^2 \frac{\partial h_t}{\partial \omega},$$

$$I_{\delta\omega} = I_{\omega\delta} = \sum_{t=1}^{n} \left(1 - \frac{\langle m, M\rangle_t^2}{\langle m\rangle_t \langle M\rangle_t}\right)^{-1} \frac{y_{t-1}^2 \sigma_\varepsilon^2}{\langle M\rangle_t} \frac{\partial h_t}{\partial \omega},$$

$$I_{\omega\omega} = \sum_{t=1}^{n} \left(1 - \frac{\langle m, M\rangle_t^2}{\langle m\rangle_t \langle M\rangle_t}\right)^{-1} \frac{\sigma_\varepsilon^4}{\langle M\rangle_t} \left(\frac{\partial h_t}{\partial \omega}\right)^2,$$

$$\mathbf{I}_{\theta\alpha'} = \mathbf{I}'_{\alpha\theta} = \left(\sum_{t=1}^{n} \left(1 - \frac{\langle m, s\rangle_t^2}{\langle m\rangle_t \langle s\rangle_t}\right)^{-1} \left(\left(h_t \frac{\partial \sigma_\varepsilon^2}{\partial \theta} + \sigma_\varepsilon^2 \frac{\partial h_t}{\partial \theta}\right) \frac{1}{\langle M\rangle_t} - \frac{y_{t-1}\langle m, M\rangle_t}{\langle m\rangle_t \langle M\rangle_t}\right) \sigma_\varepsilon^2 \frac{\partial h_t}{\partial \alpha_i}\right)_{i=1,\cdots,p},$$

$$\mathbf{I}_{\delta\alpha'} = \mathbf{I}'_{\alpha\delta} = \left(\sum_{t=1}^{n} \left(1 - \frac{\langle m, M\rangle_t^2}{\langle m\rangle_t \langle M\rangle_t}\right)^{-1} \frac{y_{t-1}^2 \sigma_\varepsilon^2}{\langle M\rangle_t} \frac{\partial h_t}{\partial \alpha_i}\right)_{i=1,\cdots,p},$$

$$\mathbf{I}_{\omega\alpha'} = \mathbf{I}'_{\alpha\omega} = \left(\sum_{t=1}^{n} \left(1 - \frac{\langle m, M\rangle_t^2}{\langle m\rangle_t \langle M\rangle_t}\right)^{-1} \frac{\sigma_\varepsilon^4}{\langle M\rangle_t} \frac{\partial h_t}{\partial \omega} \frac{\partial h_t}{\partial \alpha_i}\right)_{i=1,\cdots,p},$$

$$\mathbf{I}_{\alpha\alpha'} = \left(\sum_{t=1}^{n} \left(1 - \frac{\langle m, M\rangle_t^2}{\langle m\rangle_t \langle M\rangle_t}\right)^{-1} \frac{\sigma_\varepsilon^4}{\langle M\rangle_t} \frac{\partial h_t}{\partial \alpha_i} \frac{\partial h_t}{\partial \alpha_j}\right)_{i,j=1,\cdots,p},$$

$$\mathbf{I}_{\theta\beta'} = \mathbf{I}'_{\beta\theta} = \left(\sum_{t=1}^{n} \left(1 - \frac{\langle m, M\rangle_t^2}{\langle m\rangle_t \langle M\rangle_t}\right)^{-1} \left(\left(h_t \frac{\partial \sigma_\varepsilon^2}{\partial \theta} + \sigma_\varepsilon^2 \frac{\partial h_t}{\partial \theta}\right) \frac{1}{\langle M\rangle_t} - \frac{y_{t-1}\langle m, M\rangle_t}{\langle m\rangle_t \langle M\rangle_t}\right) \sigma_\varepsilon^2 \frac{\partial h_t}{\partial \beta_j}\right)_{j=1,\cdots,q},$$

$$\mathbf{I}_{\delta\beta'} = \mathbf{I}'_{\beta\delta} = \left(\sum_{t=1}^{n} \left(1 - \frac{\langle m, M\rangle_t^2}{\langle m\rangle_t \langle M\rangle_t}\right)^{-1} \frac{y_{t-1}^2 \sigma_\varepsilon^2}{\langle M\rangle_t} \frac{\partial h_t}{\partial \beta_j}\right)_{j=1,\cdots,p},$$

$$\mathbf{I}_{\omega\beta'} = \mathbf{I}'_{\beta\omega} = \left(\sum_{t=1}^{n} \left(1 - \frac{\langle m, M\rangle_t^2}{\langle m\rangle_t \langle M\rangle_t}\right)^{-1} \frac{\sigma_\varepsilon^4}{\langle M\rangle_t} \frac{\partial h_t}{\partial \omega} \frac{\partial h_t}{\partial \beta_j}\right)_{j=1,\cdots,q},$$

$$\mathbf{I}_{\alpha\beta'} = \mathbf{I}'_{\beta\alpha'} = \left(\sum_{t=1}^{n} \left(1 - \frac{\langle m, M\rangle_t^2}{\langle m\rangle_t \langle M\rangle_t}\right)^{-1} \frac{\sigma_\varepsilon^4}{\langle M\rangle_t} \frac{\partial h_t}{\partial \alpha_i} \frac{\partial h_t}{\partial \beta_j}\right)_{i=1,\cdots,p, j=1,\cdots,q},$$

$$\mathbf{I}_{\beta\beta'} = \left(\sum_{t=1}^{n} \left(1 - \frac{\langle m, M\rangle_t^2}{\langle m\rangle_t \langle M\rangle_t}\right)^{-1} \frac{\sigma_\varepsilon^4}{\langle M\rangle_t} \frac{\partial h_t}{\partial \beta_i} \frac{\partial h_t}{\partial \beta_j}\right)_{i,j=1,\cdots,q}.$$

Acknowledgements

References

Bera, A. K., Bilias, Y., & Simlai, P. (2006).Estimating functions and equations: an essay on historical developments with applications to economics. In: Mills T. C., Patterson, K. (Eds.), *Palgrave Handbook of Econometrics, I,* 427 - 476.

Broto, C., & Ruiz, E. (2004). Estimation method for stochastic volatility models: a survey. *Journal of Economic Surveys,*

18, 613-637. http://dx.doi.org/10.1111/j.1467-6419.2004.00232.x

Ghahramani, M., & Thavaneswaran, A. (2009). Combining estimating functions for volatility. *Journal of Statistical Planning and Inference, 142*, 171 - 180. http://dx.doi.org/10.1016/j.jspi.2008.07.014

Ghahramani, M., & Thavaneswaran, A. (2012). Nonlinear recursive estimation of the volatility via estimating functions. *Journal of Statistical Planning and Inference, 139*, 1449 - 1461. http://dx.doi.org/10.1016/j.jspi.2011.07.006

Liang, Y., Thavaneswaran, A., & Abraham, B. (2011). Joint estimation using quadratic estimating functions. *Journal of Probability and Statistics,* Article ID 372512, 14 pages. http://dx.doi.org/10.1155/2011/372512

Merkouris, T. (2007). Transform martingale estimating functions. *The Annals of Statistics, 35*, 1975 - 2000. http://dx.doi.org/10.1214/009053607000000299

Naik-Nimbalkar, U. V., & Rajarshi, M. B. (1995). Filtering and smoothing via estimating functions. *Journal of the American Statistical Association, 90*, 301 - 306. http://dx.doi.org/10.1080/01621459.1995.10476513

Thavaneswaran, A., & Thompson, M. E. (1988). A criterion for filtering in semimartingale models. *Stochastic Process Appl., 28*, 259 - 265. http://dx.doi.org/10.1016/0304-4149(88)90099-3

Thavaneswaran, A., & Abraham, B. (1988). Estimation of nonlinear time series models using estimating functions. *Journal of Time Series Analysis, 9*, 99 - 108. http://dx.doi.org/10.1111/j.1467-9892.1988.tb00457.x

Thavaneswaran, A., & Heyde, C. C. (1999). Prediction via estimating functions. *Journal of Statistical Planning and Inference, 77*, 89 - 101. http://dx.doi.org/10.1016/S0378-3758(98)00179-7

Thavaneswaran, A. , Liang, Y., & Frank, J. (2012). Inference for random coeffient volatility models. *Statisitcs & Probability Letters, 82*, 2086 - 2090. http://dx.doi.org/10.1016/j.spl.2012.07.008

Thavaneswaran, A., Ravishanker, N., & Liang, Y. (2013). Inference for linear and nonlinear stable error processes via estimating functions. *Journal of Statistical Planning and Inference, 143*, 827 - 841. http://dx.doi.org/10.1016/j.jspi.2012.10.014

Thavaneswaran, & A. Ravishanker, N. (2015). Estimating functions for circular models. Technical Report # 2015-07, Department of Statistics, University of Connecticut, Storrs, USA.

Thavaneswaran, A., Ravishanker, N., & Liang, Y. (2015). Generalized duration models and optimal estimation using estimating functions. *Annals of the Institute of Statistical Mathematics, 67*, 129-156. http://dx.doi.org/10.1007/s10463-013-0442-9

Copyrights