Inference for Duration Models using Estimating Functions

Julieta Frank* You Liang† A. Thavaneswaran‡ Melody Ghahramani§

Abstract
A class of martingale estimating functions provides a more convenient framework for studying inference for nonlinear time series models relative to other widely used methods such as maximum likelihood estimation. For example, the estimating function approach does not assume any particular distribution for the innovation. Liang et al. (2011) have recently shown that quadratic estimating functions are more informative than linear estimating functions for Random Coefficient Autoregresive (RCA) models. Duration models are commonly used to model the behaviour of irregularly time-spaced financial data. The method is used to study the inference for the parameters of a new class of multiplicative Random Coefficient Autoregressive Conditional Duration (RCACD) models.

Key Words: Estimating Functions, Random Coefficient Autoregressive Conditional Duration (RCACD) model, Nonlinear Time Series, Information

1. Introduction

Godambe (1985) proposed the estimating function approach to study inference for discrete time stochastic processes. Thavaneswaran and Abraham (1988) showed that estimating functions provide a desirable framework for inference in nonlinear time series models. Bera et al. (2006) provide a survey of the theory of estimating functions and illustrate applications to various time series models such as the Random Coefficient Autoregressive (RCA). Recently, Ghahramani and Thavaneswaran (2012) use combined estimating functions to obtain optimal nonlinear recursive estimation for RCA models.

The RCA model proposed by Nicholls and Quinn (1982) has been widely used in the financial literature. Nicholls and Quinn (1982) discuss maximum likelihood and conditional least squares estimation methods for the RCA model. Thavaneswaran and Abraham (1988) and Chandra and Taniguchi (2001) discuss parameter estimation using estimating functions. Liang et al. (2011) show that, for the RCA model, the optimal combined martingale estimating function is more informative than the component estimating functions.

Duration models, first introduced by Engle and Russell (1998), are commonly used to model the behavior of irregularly time-spaced financial data. Such data has become increasingly important as real-time, high frequency, transaction and quote data are now readily available. Estimation of the standard Autoregressive Conditional Duration (ACD) model is usually performed via maximum likelihood. However, inference for duration models using the estimating function approach has been little explored.

*Department of Agribusiness and Agricultural Economics, University of Manitoba, 376 Agriculture Bldg., Winnipeg, Manitoba R3T 2N2, Canada
†Department of Statistics, University of Manitoba, 356 Machray Hall, Winnipeg, Manitoba R3T 2N2, Canada
‡Department of Statistics, University of Manitoba, 326 Machray Hall, Winnipeg, Manitoba R3T 2N2, Canada
§Department of Mathematics and Statistics, University of Winnipeg, 515 Portage Avenue, Winnipeg, Manitoba R3B 2E9, Canada

In this paper we first propose a new class of Multiplicative Random Coefficient Autoregressive Conditional Duration (RCACD) model and then we derive a general expression for the optimal estimating function, the quadratic estimating function for model parameters, and the corresponding information matrix for the new model. The information gain associated with the proposed method is briefly discussed.

2. Estimating Function Approach

2.1 The basics

Let $\{\mathbf{y}_t\}, t = 1, \ldots, n$ be a realization of a discrete-time stochastic process with its distribution depending on a vector parameter $\boldsymbol{\theta}$, $\mathscr{F}_t^y$ be the σ-field generated by $\{\mathbf{y}_t\}$, and $\mathbf{h}_t = \mathbf{h}_t(\mathbf{y}_1, \ldots, \mathbf{y}_t, \boldsymbol{\theta})$ be specified q-dimensional vectors that are martingales. For the class $\mathscr{M}$ of zero mean and square integrable p-dimensional martingale estimating functions of the form $\mathscr{M} = \{\mathbf{g}_n(\boldsymbol{\theta}) : \mathbf{g}_n(\boldsymbol{\theta}) = \sum_{t=1}^n \mathbf{a}_{t-1}\mathbf{h}_t\}$, where $\mathbf{a}_{t-1}$ are $p \times q$ matrices depending on $\mathbf{y}_t, \ldots, \mathbf{y}_{t-1}$, the optimal EF which maximizes the information matrix

$$\mathbf{I}_{\mathbf{g}_n}(\boldsymbol{\theta}) = \left(\mathrm{E}\left[\frac{\partial \mathbf{g}_n(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}}\middle| \mathscr{F}_{n-1}^y\right]\right)' \left(\mathrm{E}[\mathbf{g}_n(\boldsymbol{\theta})\mathbf{g}_n(\boldsymbol{\theta})'|\mathscr{F}_{n-1}^y]\right)^{-1} \left(\mathrm{E}\left[\frac{\partial \mathbf{g}_n(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}}\middle| \mathscr{F}_{n-1}^y\right]\right)$$

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

and the corresponding optimal information reduces to $\mathrm{E}[\mathbf{g}_n^*(\boldsymbol{\theta})\mathbf{g}_n^*(\boldsymbol{\theta})'|\mathscr{F}_{n-1}^y]$.

2.2 Parameter estimation for a general model

For the discrete-time stochastic process $\{y_t\}$ with conditional moments

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

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

$$\gamma_t(\boldsymbol{\theta}) = \mathrm{E}\left[(y_t - \mu_t(\boldsymbol{\theta}))^3\,|\mathscr{F}_{(t-1)}^y\right], \text{and}$$

$$\kappa_t(\boldsymbol{\theta}) = \mathrm{E}\left[(y_t - \mu_t(\boldsymbol{\theta}))^4\,|\mathscr{F}_{(t-1)}^y\right],$$

define martingale differences: $\{m_t(\boldsymbol{\theta}) = y_t - \mu_t(\boldsymbol{\theta})\}$ and $\{M_t(\boldsymbol{\theta}) = m_t^2(\boldsymbol{\theta}) - \sigma_t^2(\boldsymbol{\theta})\}$ with quadratic variations $\langle m\rangle_t = \mathrm{E}[m_t^2|\mathscr{F}_{t-1}^y] = \sigma_t^2$, $\langle M\rangle_t = \mathrm{E}[M_t^2|\mathscr{F}_{t-1}^y] = \kappa_t - \sigma_t^4$, and quadratic covariation $\langle m, M\rangle_t = \mathrm{E}[m_t M_t|\mathscr{F}_{t-1}^y] = \gamma_t$. Then, the optimal estimating functions $\mathbf{g}_m^*(\boldsymbol{\theta})$ and $\mathbf{g}_M^*(\boldsymbol{\theta})$ based on m_t and M_t, respectively, are:

$$\mathbf{g}_m^*(\boldsymbol{\theta}) = -\sum_{t=1}^n \frac{\partial \mu_t}{\partial \boldsymbol{\theta}}\frac{m_t}{\langle m\rangle_t} \quad \text{and} \quad \mathbf{g}_M^*(\boldsymbol{\theta}) = -\sum_{t=1}^n \frac{\partial \sigma_t^2}{\partial \boldsymbol{\theta}}\frac{M_t}{\langle M\rangle_t};$$

and the corresponding information is:

$$\mathbf{I}_{\mathbf{g}_m^*}(\boldsymbol{\theta}) = \sum_{t=1}^n \frac{\partial \mu_t}{\partial \boldsymbol{\theta}}\frac{\partial \mu_t}{\partial \boldsymbol{\theta}'}\frac{1}{\langle m\rangle_t} \quad \text{and} \quad \mathbf{I}_{\mathbf{g}_M^*}(\boldsymbol{\theta}) = \sum_{t=1}^n \frac{\partial \sigma_t^2}{\partial \boldsymbol{\theta}}\frac{\partial \sigma_t^2}{\partial \boldsymbol{\theta}'}\frac{1}{\langle M\rangle_t}.$$

The optimal combined estimating function and its corresponding information are:

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

$$\text{where } \mathbf{a}_{t-1}^* = \left(1 - \frac{\langle m, M\rangle_t^2}{\langle m\rangle_t \langle M\rangle_t}\right)^{-1} \left(-\frac{\partial \mu_t}{\partial \boldsymbol{\theta}} \frac{1}{\langle m\rangle_t} + \frac{\partial \sigma_t^2}{\partial \boldsymbol{\theta}} \frac{\langle m, M\rangle_t}{\langle m\rangle_t \langle M\rangle_t}\right) \text{ and}$$

$$\mathbf{b}_{t-1}^* = \left(1 - \frac{\langle m, M\rangle_t^2}{\langle m\rangle_t \langle M\rangle_t}\right)^{-1} \left(\frac{\partial \mu_t}{\partial \boldsymbol{\theta}} \frac{\langle m, M\rangle_t}{\langle m\rangle_t \langle M\rangle_t} - \frac{\partial \sigma_t^2}{\partial \boldsymbol{\theta}} \frac{1}{\langle M\rangle_t}\right);$$

$$\begin{aligned}\mathbf{I}_{\mathbf{g}_C^*}(\boldsymbol{\theta}) &= \sum_{t=1}^{n} \left(1 - \frac{\langle m, M\rangle_t^2}{\langle m\rangle_t \langle M\rangle_t}\right)^{-1} \left(\frac{\partial \mu_t}{\partial \boldsymbol{\theta}} \frac{\partial \mu_t}{\partial \boldsymbol{\theta}'} \frac{1}{\langle m\rangle_t} + \frac{\partial \sigma_t^2}{\partial \boldsymbol{\theta}} \frac{\partial \sigma_t^2}{\partial \boldsymbol{\theta}'} \frac{1}{\langle M\rangle_t}\right. \\ &\quad \left. - \left(\frac{\partial \mu_t}{\partial \boldsymbol{\theta}} \frac{\partial \sigma_t^2}{\partial \boldsymbol{\theta}'} + \frac{\partial \sigma_t^2}{\partial \boldsymbol{\theta}} \frac{\partial \mu_t}{\partial \boldsymbol{\theta}'}\right) \frac{\langle m, M\rangle_t}{\langle m\rangle_t \langle M\rangle_t}\right).\end{aligned}$$

.

3. Multiplicative Random Coefficient Autoregressive Conditional Duration (RCACD) Model

In this section we introduce a new class of Multiplicative Random Coefficient ACD (1, 1) models of the form

$$x_t = (b_t + \psi_t)\varepsilon_t, \tag{1}$$

$$\psi_t = \omega + \alpha_1 x_{t-1} + \beta_1 \psi_{t-1}, \tag{2}$$

where $\omega > 0$, $\alpha_1 > 0$, $\beta_1 > 0$, and $\alpha_1 + \beta_1 < 1$. We assume that ε_t's are iid non-negative random variables with mean μ_ε, variance σ_ε^2, skewness γ_ε, and excess kurtosis κ_ε. Further assume that b_t's are iid non-negative random variables with mean μ_b, variance σ_b^2, skewness γ_b, and excess kurtosis κ_b. Also, ε_t and b_t are mutually independent. In order to estimate the parameter vector $\boldsymbol{\theta} = (\mu_b, \sigma_b^2, \omega, \alpha_1, \beta_1)'$, we use the estimating function approach. For this model, the conditional moments can be calculated as

$$\begin{aligned}\mu_t &= \mu_\varepsilon(\mu_b + \psi_t), \\ \sigma_t^2 &= \sigma_\varepsilon^2(\mu_b + \psi_t)^2 + \sigma_b^2(\mu_\varepsilon^2 + \sigma_\varepsilon^2), \\ \gamma_t &= \gamma_\varepsilon(\mu_b + \psi_t)^3 + (3\sigma_b^2\gamma_\varepsilon + 6\mu_\varepsilon\sigma_b^2\sigma_\varepsilon^2)(\mu_b + \psi_t) + \gamma_b(\mu_\varepsilon^3 + 3\mu_\varepsilon\sigma_\varepsilon^2 + \gamma_\varepsilon), \\ \kappa_t &= \kappa_\varepsilon(\mu_b + \psi_t)^4 + 6\sigma_b^2(\mu_\varepsilon^2\sigma_\varepsilon^2 + 2\mu_\varepsilon\gamma_\varepsilon + \kappa_\varepsilon)(\mu_b + \psi_t)^2 \\ &\quad + 4\gamma_b(3\mu_\varepsilon^2\sigma_b^2 + 3\mu_\varepsilon\gamma_\varepsilon + \kappa_\varepsilon)(\mu_b + \psi_t) + \kappa_b(\mu_\varepsilon^4 + 4\mu_\varepsilon\gamma_\varepsilon + 6\mu_\varepsilon^2\sigma_\varepsilon^2 + \kappa_\varepsilon).\end{aligned}$$

Let $m_t = x_t - \mu_t$ and $M_t = m_t^2 - \sigma_t^2$ be the sequences of martingale differences such that

$$\begin{aligned}\langle m\rangle_t &= \sigma_\varepsilon^2(\mu_b + \psi_t)^2 + \sigma_b^2(\mu_\varepsilon^2 + \sigma_\varepsilon^2) \\ \langle M\rangle_t &= (\kappa_\varepsilon - \sigma_\varepsilon^4)(\mu_b + \psi_t)^4 + 2\sigma_b^2(2\mu_\varepsilon^2\sigma_\varepsilon^2 + 6\mu_\varepsilon\gamma_\varepsilon + 3\kappa_\varepsilon - \sigma_\varepsilon^4)(\mu_b + \psi_t)^2 \\ &\quad + 4\gamma_b(3\mu_\varepsilon^2\sigma_b^2 + 3\mu_\varepsilon\gamma_\varepsilon + \kappa_\varepsilon)(\mu_b + \psi_t) + \kappa_b(\mu_\varepsilon^4 + 4\mu_\varepsilon\gamma_\varepsilon + 6\mu_\varepsilon^2\sigma_\varepsilon^2 + \kappa_\varepsilon) \\ &\quad - \sigma_b^4(\mu_\varepsilon^4 + 2\mu_\varepsilon^2\sigma_\varepsilon^2 + \sigma_\varepsilon^4), \\ \langle m, M\rangle_t &= \gamma_\varepsilon(\mu_b + \psi_t)^3 + (3\sigma_b^2\gamma_\varepsilon + 6\mu_\varepsilon\sigma_b^2\sigma_\varepsilon^2)(\mu_b + \psi_t) + \gamma_b(\mu_\varepsilon^3 + 3\mu_\varepsilon\sigma_\varepsilon^2 + \gamma_\varepsilon).\end{aligned}$$

Theorem 1. *For the model (1) - (2), in the class of all quadratic estimating functions of the form* $\mathscr{G}_C = \{\mathbf{g}_C(\boldsymbol{\theta}) : \mathbf{g}_C(\boldsymbol{\theta}) = \sum_{t=1}^{n} (\mathbf{a}_{t-1} m_t + \mathbf{b}_{t-1} M_t)\}$,
(a) the optimal estimating function is given by $\mathbf{g}_C^*(\boldsymbol{\theta}) = \sum_{t=1}^{n} \left(\mathbf{a}_{t-1}^* m_t + \mathbf{b}_{t-1}^* M_t\right)$, *where*

$$\begin{aligned}\mathbf{a}_{t-1}^* = \frac{1}{\langle m\rangle_t \langle M\rangle_t - \langle m, M\rangle_t^2} & \left(-\mu_\varepsilon \langle M\rangle_t + 2\sigma_\varepsilon^2(\mu_b + \psi_t)\langle m, M\rangle_t, (\mu_\varepsilon^2 + \sigma_\varepsilon^2)\langle m, M\rangle_t,\right. \\ & - \mu_\varepsilon \langle M\rangle_t + 2\sigma_\varepsilon^2(\mu_b + \psi_t)\langle m, M\rangle_t, \left(-\mu_\varepsilon \langle M\rangle_t + 2\sigma_\varepsilon^2(\mu_b + \psi_t)\langle m, M\rangle_t\right) x_{t-1}, \\ & \left.\left(-\mu_\varepsilon \langle M\rangle_t + 2\sigma_\varepsilon^2(\mu_b + \psi_t)\langle m, M\rangle_t\right) \psi_{t-1}\right)'\end{aligned}$$

and

$$\begin{aligned}\mathbf{b}_{t-1}^* = \frac{1}{\langle m\rangle_t \langle M\rangle_t - \langle m, M\rangle_t^2} & \left(\mu_\varepsilon \langle m, M\rangle_t - 2\sigma_\varepsilon^2(\mu_b + \psi_t)\langle m\rangle_t, - \left(\mu_\varepsilon^2 + \sigma_\varepsilon^2\right) \langle m\rangle_t,\right. \\ & \mu_\varepsilon \langle m, M\rangle_t - 2\sigma_\varepsilon^2(\mu_b + \psi_t)\langle m\rangle_t, \left(\mu_\varepsilon \langle m, M\rangle_t - 2\sigma_\varepsilon^2(\mu_b + \psi_t)\right) x_{t-1}, \\ & \left.\left(\mu_\varepsilon \langle m, M\rangle_t - 2\sigma_\varepsilon^2(\mu_b + \psi_t)\right) \psi_{t-1}\right)';\end{aligned}$$

(b) the quadratic estimating function for each component of $\boldsymbol{\theta}$ *is given by*

$$\begin{aligned}
g_C^*(\mu_b) &= \sum_{t=1}^{n} \frac{1}{\langle m\rangle_t \langle M\rangle_t - \langle m, M\rangle_t^2} \left(\left(-\mu_\varepsilon \langle M\rangle_t + 2\sigma_\varepsilon^2(\mu_b + \psi_t)\langle m, M\rangle_t\right) m_t \right. \\
&\quad \left. + \left(\mu_\varepsilon \langle m, M\rangle_t - 2\sigma_\varepsilon^2(\mu_b + \psi_t)\langle m\rangle_t\right) M_t\right), \\
g_C^*(\sigma_b^2) &= \sum_{t=1}^{n} \frac{1}{\langle m\rangle_t \langle M\rangle_t - \langle m, M\rangle_t^2} \left((\mu_\varepsilon^2 + \sigma_\varepsilon^2)\langle m, M\rangle_t m_t - \left(\mu_\varepsilon^2 + \sigma_\varepsilon^2\right) \langle m\rangle_t M_t\right), \\
g_C^*(\omega) &= \sum_{t=1}^{n} \frac{1}{\langle m\rangle_t \langle M\rangle_t - \langle m, M\rangle_t^2} \left(\left(-\mu_\varepsilon \langle M\rangle_t + 2\sigma_\varepsilon^2(\mu_b + \psi_t)\langle m, M\rangle_t\right) m_t \right. \\
&\quad \left. + \left(\mu_\varepsilon \langle m, M\rangle_t - 2\sigma_\varepsilon^2(\mu_b + \psi_t)\langle m\rangle_t\right) M_t\right), \\
g_C^*(\alpha_1) &= \sum_{t=1}^{n} \frac{1}{\langle m\rangle_t \langle M\rangle_t - \langle m, M\rangle_t^2} \left(\left(-\mu_\varepsilon \langle M\rangle_t + 2\sigma_\varepsilon^2(\mu_b + \psi_t)\langle m, M\rangle_t\right) x_{t-1} m_t \right. \\
&\quad \left. + \left(\mu_\varepsilon \langle m, M\rangle_t - 2\sigma_\varepsilon^2(\mu_b + \psi_t)\langle m\rangle_t\right) x_{t-1} M_t\right), \\
g_C^*(\beta_1) &= \sum_{t=1}^{n} \frac{1}{\langle m\rangle_t \langle M\rangle_t - \langle m, M\rangle_t^2} \left(\left(-\mu_\varepsilon \langle M\rangle_t + 2\sigma_\varepsilon^2(\mu_b + \psi_t)\langle m, M\rangle_t\right) \psi_{t-1} m_t \right. \\
&\quad \left. + \left(\mu_\varepsilon \langle m, M\rangle_t - 2\sigma_\varepsilon^2(\mu_b + \psi_t)\langle m\rangle_t\right) \psi_{t-1} M_t\right);
\end{aligned}$$

(c) the information matrix of the optimal quadratic estimating function for $\boldsymbol{\theta}$ *is given by*

$$\mathbf{I}_{\mathbf{g}_C^*}(\boldsymbol{\theta}) = \begin{pmatrix} I_{\mu_b\mu_b} & I_{\mu_b\sigma_b^2} & I_{\mu_b\omega} & I_{\mu_b\alpha_1} & I_{\mu_b\beta_1} \\ I_{\mu_b\sigma_b^2} & I_{\sigma_b^2\sigma_b^2} & I_{\sigma_b^2\omega} & I_{\sigma_b^2\alpha_1} & I_{\sigma_b^2\beta_1} \\ I_{\mu_b\omega} & I_{\sigma_b^2\omega} & I_{\omega\omega} & I_{\omega\alpha_1} & I_{\omega\beta_1} \\ I_{\mu_b\alpha_1} & I_{\sigma_b^2\alpha_1} & I_{\omega\alpha_1} & I_{\alpha_1\alpha_1} & I_{\alpha_1\beta_1} \\ I_{\mu_b\beta_1} & I_{\sigma_b^2\beta_1} & I_{\omega\beta_1} & I_{\alpha_1\beta_1} & I_{\beta_1\beta_1} \end{pmatrix},$$

where

$$I_{\mu_b\mu_b} = \sum_{t=1}^{n} \left(1 - \frac{\langle m, M\rangle_t^2}{\langle m\rangle_t \langle M\rangle_t}\right)^{-1} \left(\frac{\mu_\varepsilon^2}{\langle m\rangle_t} + \frac{4\sigma_\varepsilon^2(\mu_b + \psi_t)^2}{\langle M\rangle_t} - \frac{4\mu_\varepsilon \sigma_\varepsilon^2(\mu_b + \psi_t)\langle m, M\rangle_t}{\langle m\rangle_t \langle M\rangle_t}\right),$$

$$
\begin{aligned}
I_{\mu_b\sigma_b^2} &= = \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{2\sigma_\varepsilon^2(\mu_\varepsilon^2+\sigma_\varepsilon^2)}{\langle M\rangle_t}-\frac{\mu_\varepsilon(\mu_\varepsilon^2+\sigma_\varepsilon^2)\langle m,M\rangle_t}{\langle m\rangle_t\langle M\rangle_t}\right),\\
I_{\sigma_b^2\sigma_b^2} &= \sum_{t=1}^{n}\left(1-\frac{\langle m,M\rangle_t^2}{\langle m\rangle_t\langle M\rangle_t}\right)^{-1}\frac{(\mu_\varepsilon^2+\sigma_\varepsilon^2)^2}{\langle M\rangle_t},\\
I_{\mu_b\omega} &= \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{\mu_\varepsilon^2}{\langle m\rangle_t}+\frac{4\sigma_\varepsilon^2(\mu_b+\psi_t)^2}{\langle M\rangle_t}-\frac{4\mu_\varepsilon\sigma_\varepsilon^2(\mu_b+\psi_t)\langle m,M\rangle_t}{\langle m\rangle_t\langle M\rangle_t}\right),\\
I_{\sigma_b^2\omega} &= \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{4\sigma_\varepsilon^2(\mu_b+\psi_t)^2}{\langle M\rangle_t}-\frac{4\mu_\varepsilon\sigma_\varepsilon^2(\mu_b+\psi_t)\langle m,M\rangle_t}{\langle m\rangle_t\langle M\rangle_t}\right),\\
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}\left(\frac{\mu_\varepsilon^2}{\langle m\rangle_t}+\frac{4\sigma_\varepsilon^2(\mu_b+\psi_t)^2}{\langle M\rangle_t}-\frac{4\mu_\varepsilon\sigma_\varepsilon^2(\mu_b+\psi_t)\langle m,M\rangle_t}{\langle m\rangle_t\langle M\rangle_t}\right),\\
I_{\mu_b\alpha_1} &= \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{\mu_\varepsilon^2}{\langle m\rangle_t}+\frac{4\sigma_\varepsilon^2(\mu_b+\psi_t)^2}{\langle M\rangle_t}-\frac{4\mu_\varepsilon\sigma_\varepsilon^2(\mu_b+\psi_t)\langle m,M\rangle_t}{\langle m\rangle_t\langle M\rangle_t}\right)x_{t-1},\\
I_{\sigma_b^2\alpha_1} &= \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{2\sigma_\varepsilon^2(\mu_\varepsilon^2+\sigma_b^2)(\mu_b+\psi_t)}{\langle M\rangle_t}-\frac{\mu_\varepsilon(\mu_\varepsilon^2+\sigma_b^2)\langle m,M\rangle_t}{\langle m\rangle_t\langle M\rangle_t}\right)x_{t-1},\\
I_{\omega\alpha_1} &= \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{\mu_\varepsilon^2}{\langle m\rangle_t}+\frac{4\sigma_\varepsilon^2(\mu_b+\psi_t)^2}{\langle M\rangle_t}-\frac{4\mu_\varepsilon\sigma_\varepsilon^2(\mu_b+\psi_t)\langle m,M\rangle_t}{\langle m\rangle_t\langle M\rangle_t}\right)x_{t-1},\\
I_{\alpha_1\alpha_1} &= \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{\mu_\varepsilon^2}{\langle m\rangle_t}+\frac{4\sigma_\varepsilon^2(\mu_b+\psi_t)^2}{\langle M\rangle_t}-\frac{4\mu_\varepsilon\sigma_\varepsilon^2(\mu_b+\psi_t)\langle m,M\rangle_t}{\langle m\rangle_t\langle M\rangle_t}\right)x_{t-1}^2,\\
I_{\mu_b\beta_1} &= \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{\mu_\varepsilon^2\psi_{t-j}}{\langle m\rangle_t}+\frac{4\sigma_\varepsilon^2(\mu_b+\psi_t)^2}{\langle M\rangle_t}-\frac{4\mu_\varepsilon\sigma_\varepsilon^2(\mu_b+\psi_t)\langle m,M\rangle_t}{\langle m\rangle_t\langle M\rangle_t}\right)\psi_{t-1},\\
I_{\sigma_b^2\beta_1} &= \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{2\sigma_\varepsilon^2(\mu_\varepsilon^2+\sigma_b^2)(\mu_b+\psi_t)}{\langle M\rangle_t}-\frac{\mu_\varepsilon(\mu_\varepsilon^2+\sigma_b^2)\langle m,M\rangle_t}{\langle m\rangle_t\langle M\rangle_t}\right)\psi_{t-1},\\
I_{\omega\beta_1} &= \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{\mu_\varepsilon^2}{\langle m\rangle_t}+\frac{4\sigma_\varepsilon^2(\mu_b+\psi_t)^2}{\langle M\rangle_t}-\frac{4\mu_\varepsilon\sigma_\varepsilon^2(\mu_b+\psi_t)\langle m,M\rangle_t}{\langle m\rangle_t\langle M\rangle_t}\right)\psi_{t-1},\\
I_{\alpha_1\beta_1} &= \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{\mu_\varepsilon^2}{\langle m\rangle_t}+\frac{4\sigma_\varepsilon^2(\mu_b+\psi_t)^2}{\langle M\rangle_t}-\frac{4\mu_\varepsilon\sigma_\varepsilon^2(\mu_b+\psi_t)\langle m,M\rangle_t}{\langle m\rangle_t\langle M\rangle_t}\right)x_{t-1}\psi_{t-1},\\
I_{\beta_1\beta_1} &= \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{\mu_\varepsilon^2}{\langle m\rangle_t}+\frac{4\sigma_\varepsilon^2(\mu_b+\psi_t)^2}{\langle M\rangle_t}-\frac{4\mu_\varepsilon\sigma_\varepsilon^2(\mu_b+\psi_t)\langle m,M\rangle_t}{\langle m\rangle_t\langle M\rangle_t}\right)\psi_{t-1}^2.
\end{aligned}
$$

Proof. Since

$$\frac{\partial\mu_t}{\partial\boldsymbol{\theta}} = (\mu_\varepsilon, 0, \mu_\varepsilon, \mu_\varepsilon x_{t-1}, \mu_\varepsilon\psi_{t-1})'$$

and

$$\frac{\partial\sigma_t^2}{\partial\boldsymbol{\theta}} = \left(2\sigma_\varepsilon^2(\mu_b+\psi_t), \mu_\varepsilon^2+\sigma_\varepsilon^2, 2\sigma_\varepsilon^2(\mu_b+\psi_t), 2\sigma_\varepsilon^2(\mu_b+\psi_t)x_{t-1}, 2\sigma_\varepsilon^2(\mu_b+\psi_t)\psi_{t-1}\right)',$$

the proof follows Theorem 2.1 in Liang et al. (2011).

3.1 Information gain

For an RCA model with GARCH errors, Thavaneswaran et al. (2012) have shown that the martingale optimal estimating functions are more informative than the conditional least squares. For our new model, following Liang et al. (2011) and using two orthogonal martingale differences m_t and $\Psi_t = M_t - \sigma_t\gamma_t m_t$, where $\langle\Psi\rangle_t = (\langle m\rangle_t\langle M\rangle_t - \langle m, M\rangle_t^2)/\langle m\rangle_t$, it can be shown that $\mathbf{I}_{\mathbf{g}_C^*}(\boldsymbol{\theta}) = \mathbf{I}_{\mathbf{g}_m^*}(\boldsymbol{\theta}) + \mathbf{I}_{\mathbf{g}_\Psi^*}(\boldsymbol{\theta})$ and therefore $\mathbf{I}_{\mathbf{g}_C^*}(\boldsymbol{\theta}) \geq \mathbf{I}_{\mathbf{g}_m^*}(\boldsymbol{\theta})$ and $\mathbf{I}_{\mathbf{g}_C^*}(\boldsymbol{\theta}) \geq \mathbf{I}_{\mathbf{g}_M^*}(\boldsymbol{\theta})$.

4. Conclusions

A general framework using the estimating function approach is developed for a new class of random coefficient autoregressive conditional duration models. The model may be appropriate for irregularly spaced time series commonly used in financial applications. We derive the quadratic estimating functions for model parameters and the corresponding information matrix. We show that when the information about the higher order conditional moments of the observed process becomes available, the combined optimal quadratic estimating function is more informative than the component estimating functions. Our proposed model can be extended to a more general form to account for seasonality and such model and estimation framework can be of interest for various financial applications including option pricing (e.g. Frank et al. 2011).

Acknowledgment

The authors gratefully acknowledge financial support from the Natural Sciences and Engineering Research Council of Canada.

REFERENCES

Bera, A., Bilias, Y., and Simlai, P. (2006), "Estimating functions and equations: An essay on historical developments with applications to econometrics," in: *Palgrave Handbook of Econometrics*, vol. 1., eds. T. Mills and K. Patterson, UK: Palgrave MacMillan.

Chandra, S., and Taniguchi, M. (2001), "Estimation functions for non-linear time series models," *Annals of the Institute of Statistical Mathematics*, 53(1), 125–141.

Engle, R., and Russell, J. (1998), "Autoregressive conditional duration: a new approach for irregularly spaced transaction data," *Econometrica*, 66, 1127–1162.

Frank, J., Ghahramani, M. , and Thavaneswaran, A. (2011), "Recent developments in seasonal volatility models," in: *Advances in Econometrics - Theory and Applications*, chapter 3, ed. M. Verbič, Croatia: InTech.

Ghahramani, M., and Thavaneswaran, A. (2012), "Nonlinear recursive estimation of volatility via estimating functions," *Journal of Statistical Planning & Inference*, 142(1), 171–180.

Godambe, V. (1985), "The foundations of finite sample estimation in stochastic processes," *Biometrika*, 72, 419–478.

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

Nicholls, D., and Quinn, B. (1982), *Random Coefficients Autoregressive Models: An introduction*, New York: Springer.

Thavaneswaran, A., and Abraham, B. (1988), "Estimation of nonlinear time series models using estimating functions," *Journal of Time Series Analysis*, 9, 99–108.

Thavaneswaran, A., Liang, Y., and Frank, J. (2012), "Inference for Random Coefficient Volatility Models," *Statistics and Probability Letters*, 82, 2086–2090.