Semiparametric Continuous Time GARCH Models: An Estimation Function Approach

Alexander Szimayer

Department of Finance
School of Economics
University of Bonn

July 16, 2009

Introduction

This talk is concerned with:

- continuous time **GARCH** processes

The continuous time GARCH

- is based on the **discrete time** GARCH of Engle (1982) and Bollerslev (1986)
- reproduces the so-called **stylized facts** of financial data
  - volatility clustering
  - fat tails
  - volatility mean reversion

Introduction

Facts: about the discrete GARCH

- **popular** model choice by academics in finance and economics
- applied by finance and banking **industry** as well
- used for **option pricing**
- estimating volatility (as synonym for risk) for **risk management** purposes

Is there a **demand** for a continuous time version of the successful GARCH model?

Introduction

The **demand** for a continuous time version of GARCH stems from

- theory of **option pricing**:
  - study market for properties (no-arbitrage, completeness)
  - better formulation for pricing and hedging of derivatives
- **econometric theory**
  - continuous time model allows to study **unequally spaced** time series data in the GARCH framework (scaling rules)

Discrete Time GARCH

The **discrete time** GARCH(1,1)-model is defined by

$$\begin{aligned} X_t &= \mu_t + \sqrt{h_t}\,\varepsilon_t \\ h_t &= \omega + \alpha\,(X_{t-1} - \mu_{t-1})^2 + \beta\,h_{t-1} \end{aligned}$$

The parameters have the interpretation:

- ω is the base level variance;
- α quantifies the impact of shocks ;
- β quantifies the **persistence** of shocks.

Discrete Time GARCH

For the variance h_t we have the **representation**

$$h_t = \omega \sum_{s=0}^{t-1} \prod_{u=s+1}^{t-1} (\beta + \alpha\, \varepsilon_u^2) + h_0 \prod_{u=0}^{t-1} (\beta + \alpha\, \varepsilon_u^2) \,.$$

relating the stability ($h_t \xrightarrow{D} h_\infty$?) to **perpetuities**, see Goldie and Maller (1996).

The above representation inspired a **continuous time version** of the process, see Klüppelberg, Lindner and Maller (2004).

Continuous Time GARCH

Given:

- pure jump Lévy process L with characteristic triplet $(\gamma, 0, \Pi)$,
- parameter $\boldsymbol{\theta} = (\kappa, \overline{V}, \eta)'$, with $\kappa > 0, \overline{V} > 0, \eta > 0$.

Define the **variance process** $V = (V_t)_{t \geq 0}$ and the **integrated continuous time GARCH process** $E = (E_t)_{t \geq 0}$ by

$$\mathrm{d}V_t = \kappa \left(\overline{V} - V_{t-} \right) \mathrm{d}t + \eta \, V_{t-} \, \mathrm{d}[L, L]_t, \quad t > 0. \tag{1}$$

$$\mathrm{d}E_t = \sqrt{V_{t-}} \, \mathrm{d}L_t, \quad t \geq 0. \tag{2}$$

See Klüppelberg et al. (2004, 2006) for properties of E and V.

Continuous Time GARCH

Note:

- The term V_{t-} in the volatility of the variance equation

$$\mathrm{d}V_t = \kappa \left(\overline{V} - V_{t-}\right) \mathrm{d}t + \eta \, V_{t-} \, \mathrm{d}[L, L]_t, \; t > 0.$$

  leads to technically difficult **non-linearities**.

- E is **heteroscedastic noise**, hence we require $\mathbb{E}L_t = 0$ and unit variance rate, i.e. $\mathbb{E}L_t^2 = t$, or, equivalently, $\int x^2 \, \Pi(\mathrm{d}x) = 1$.

Continuous Time GARCH

Extension: The heteroscedastic noise E can serve as a basis for formulating continuous time processes of the form:

$$\mathrm{d}Y_t = f(Y_{t-}, \mathbf{X}_{t-}, V_{t-}; \boldsymbol{\theta})\,\mathrm{d}t + \mathrm{d}E_t\,.$$

Examples

- continuous time version of **linear regression**

$$\mathrm{d}Y_t = \beta' \mathbf{X}_t\,\mathrm{d}t + \mathrm{d}E_t\,.$$

- cumulative return process Y of **stock price** $S = S_0\,\mathcal{E}(Y)$:

$$\mathrm{d}Y_t = \mu(V_{t-})\,\mathrm{d}t + \mathrm{d}E_t\,.$$

The process has to be stopped at $\tau = \inf\{t \geq 0 : \Delta E \leq -1\}$.

Semiparametric Estimation

Question:

How can we **estimate** the parameter vector $\boldsymbol{\theta}$?

- i.e.: the **variance** parameters $(\kappa, \overline{V}, \eta)'$, and
- the structural parameters in the **mean equation**.

We do **not** want to estimate the Lévy characteristics!

Thus we are in a **semiparametric** setting!

Quasi Likelihood

For **discretely** observed data Haug, Klüppelberg, Lindner and Zapp (2007) apply the **method of moment** estimation.

We wish to estimate $\boldsymbol{\theta}$ from **continuous** time observations of E using **quasi likelihood**, see Hutton and Nelson (1986) and Heyde (1997).

Throughout, denote $\boldsymbol{\theta}_0$ the true parameter value.

Quasi Likelihood

Define the natural basis of the martingale estimating functions by

$$K_T(\boldsymbol{\theta}) := \int_0^T V_{t-}(\boldsymbol{\theta})\,\mathrm{d}t - [E,E]_T\,, \quad \text{for } t \geq 0\,.$$

We consider **martingale estimating functions**:

$$\mathcal{M} = \left\{ \mathbf{G}_T(\boldsymbol{\theta}) = \int_0^T \boldsymbol{\alpha}_t(\boldsymbol{\theta})\,\mathrm{d}K_t(\boldsymbol{\theta}) \right\}\,, \tag{3}$$

where $\boldsymbol{\alpha}(\boldsymbol{\theta}) = (\boldsymbol{\alpha}(\boldsymbol{\theta}))_{t\geq 0}$ is a predictable 3-dimensional processes that is twice continuously differentiable in $\boldsymbol{\theta}$.

Quasi Likelihood

Technical Assumptions:

- all elements $\mathbf{G}_T(\boldsymbol{\theta})$ of $\mathcal{M}$ are square integrable;
- $\mathbb{E}L_t^4 < \infty$, i.e., $\int_{|x|\geq 1} x^4\,\Pi(\mathrm{d}x) < \infty$, such that $[L, L]$ is square integrable (and hence $([L, L]_t - t)_{t\geq 0}$ is a square integrable martingale);
- $\langle \mathbf{G} \rangle_T(\boldsymbol{\theta})$ is nonsingular.

These assumptions are required to apply martingale convergence theorems leading to an asymptotic normal distribution of the estimators. Note that in general, these assumption are not necessary and can be relaxed by introducing appropriate transformations of the underlying martingale family, see Heyde (1997), Chapter 13.1.2.

Deriving the Optimal Estimator

In (3) we are given a large set of estimating functions.

The aim is to derive an **optimal estimating function** $\mathbf{G}^{\star}_{T}(\boldsymbol{\theta})$.

Optimality here is understood as outlined by, e.g., Heyde (1997), as the element $\mathbf{G}_T(\theta)$ that maximizes the martingale information

$$I_{\mathbf{G}}(\boldsymbol{\theta}) = \langle \mathbf{G}(\boldsymbol{\theta}) \rangle_T \,, \qquad (4)$$

leading to **minimal confidence zones**.

Deriving the Optimal Estimator

We prepare the derivation of optimal estimating function rewriting (1) as follows:

$$\mathrm{d}V_t(\boldsymbol{\theta}) = \kappa\,\left(\overline{V} - V_{t-}(\boldsymbol{\theta})\right)\,\mathrm{d}t + \eta\,\mathrm{d}[E,E]_t\,, \qquad t > 0, \tag{5}$$

where the dependence on the parameters $\boldsymbol{\theta}$ is made **explicit** here.

Noting that the expression is of **OU-type** with driver $\eta\,[E,E]$ yields the formal solution

$$V_t(\boldsymbol{\theta}) = \overline{V} + \left(V_0 - \overline{V}\right)\,e^{-\kappa\,t} + \eta\,e^{-\kappa\,t}\int_0^t e^{\kappa\,s}\mathrm{d}[E,E]_s\,, \qquad t \geq 0, \tag{6}$$

Deriving the Optimal Estimator

The construction of the optimal estimating function from the class of martingale estimating functions in (3) can be performed by

$$\mathbf{G}^{\star}_{kT}(\boldsymbol{\theta}) = -\int_0^T \frac{\mathrm{d}\bar{K}_{t-}(\boldsymbol{\theta})}{\mathrm{d}\langle K_t(\boldsymbol{\theta})\rangle_{t-}} \,\mathrm{d}K_t(\boldsymbol{\theta})\,,$$

see Hutton and Nelson (1986) and Heyde (1997), Chapter 2.5.

Deriving the Optimal Estimator

The first derivative and its compensator are

$$\mathrm{d}\dot{K}_t(\boldsymbol{\theta}) = \dot{V}_{t-}(\boldsymbol{\theta})\,\mathrm{d}t\,, \quad \text{and} \quad \mathrm{d}\bar{K}_t(\boldsymbol{\theta}) = \dot{V}_{t-}(\boldsymbol{\theta})\,\mathrm{d}t\,, \quad t > 0\,.$$

The predictable projection of the bracket process is

$$\mathrm{d}\langle K(\boldsymbol{\theta})\rangle_t = m_4\, V_{t-}^2(\boldsymbol{\theta})\,\mathrm{d}t\,,$$

where $m_4 = \int x^4\, \Pi(\mathrm{d}x) < \infty$, and m_4 is a nuisance parameter.

Deriving the Optimal Estimator

Then the **quasi score** is

$$G^{\star}_T(\boldsymbol{\theta}) = \frac{1}{m_4} \int_0^T \frac{\dot{V}_{t-}(\boldsymbol{\theta})}{V^2_{t-}(\boldsymbol{\theta})} \left(\mathrm{d}[E,E]_t - V_{t-}(\boldsymbol{\theta})\,\mathrm{d}t\right).$$

To obtain the **point estimate** $\hat{\boldsymbol{\theta}}_T$ solve

$$\mathbf{0} = \mathbf{G}^{\star}_T(\hat{\boldsymbol{\theta}}_T)\,.$$

The partial derivatives of $V_t(\boldsymbol{\theta})$ are straightforward

$$\frac{\partial V_t}{\partial \kappa}(\boldsymbol{\theta}) = -t\,(V_0 - \overline{V})\,e^{-\kappa\,t} - \eta \int_0^t (t-s)\,e^{-\kappa\,(t-s)}\,\mathrm{d}[E,E]_s\,,$$

$$\frac{\partial V_t}{\partial \overline{V}}(\boldsymbol{\theta}) = 1 - e^{-\kappa\,t}\,, \qquad \frac{\partial V_t}{\partial \eta}(\boldsymbol{\theta}) = \int_0^t e^{-\kappa\,(t-s)}\,\mathrm{d}[E,E]_s\,.$$

Deriving the Optimal Estimator

Standard Program in Statistics:

- Establish the **consistency** of estimator $\boldsymbol{\theta}_T$.
- Establish **local asymptotic normality** of estimator $\boldsymbol{\theta}_T$:
  - provide confidence zones
  - prepare hypothesis testing

Problem:

Non-linearity of volatility component of the variance process

$$\mathrm{d}V_t = \kappa \left(\overline{V} - V_{t-}\right) \mathrm{d}t + \eta\, V_{t-}\, \mathrm{d}[L, L]_t,\ \ t > 0.$$

Consistency

The formal condition for the **strong consistency** (possibly on an event $A \in \Omega$) is that for all sufficiently small $\delta > 0$,

$$\limsup_{T\to\infty} \left(\sup_{\|\boldsymbol{\theta}-\boldsymbol{\theta}_0\|=\delta} \mathscr{L}_T(\boldsymbol{\theta}) - \mathscr{L}_T(\boldsymbol{\theta}_0) \right) < 0 \, a.s. \, (\text{on } A) \, .$$

see Hutton and Nelson (1986) and Heyde (1997).

Here, the **quasi-likelihood** $\mathscr{L}_T(\boldsymbol{\theta})$ is

$$\mathscr{L}_T(\boldsymbol{\theta}) = -\frac{1}{m_4} \left(\int_0^T \log(V_{t-}(\boldsymbol{\theta}))\,\mathrm{d}t + \int_0^T \frac{1}{V_{t-}(\boldsymbol{\theta})}\,\mathrm{d}[E]_t \right) .$$

By differentiating with respect to $\boldsymbol{\theta}$ we verify

$$\frac{\partial \mathscr{L}_T}{\partial \boldsymbol{\theta}}(\boldsymbol{\theta}) = \mathbf{G}_T^{\star}(\boldsymbol{\theta}) \, .$$

Consistency

In the quasi-likelihood setting we can check the **stronger condition**

$$\limsup_{T\to\infty}\left(\sup_{\|\boldsymbol{\theta}-\boldsymbol{\theta}_0\|=\delta}\frac{\mathscr{L}_T(\boldsymbol{\theta})-\mathscr{L}_T(\boldsymbol{\theta}_0)}{T}\right)<0\,a.s.\,(\text{on } A)\,. \qquad (7)$$

For the rather **simple structure** of the quasi-likelihood, strong consistency can be shown by verifying (7) **directly**.

Consistency

Strategy

1. Define

$$l_T(\theta) = \frac{\mathscr{L}_T(\boldsymbol{\theta}) - \mathscr{L}_T(\boldsymbol{\theta}_0)}{T}\,.$$

2. For some real number $l_\infty(\theta) < 0$ show that

$$l_T(\theta) \to l_\infty(\theta) < 0 \text{ a.s.}$$

3. Establish equicontinuity by showing for some constant C

$$\left|\frac{\partial l_T(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}}\right| \leq C \text{ a.s.}\,.$$

Consistency

Lemma

Assume that $\mathbb{E}L_1^4 < \infty$. Then for any $\boldsymbol{\theta} \in \boldsymbol{\Theta}$ with $\boldsymbol{\theta} \neq \boldsymbol{\theta}_0$ the following limit behavior applies

$$\lim_{T \to \infty} \frac{\mathscr{L}_T(\boldsymbol{\theta}) - \mathscr{L}_T(\boldsymbol{\theta}_0)}{T} < 0 \ a.s. \tag{8}$$

Consistency

Sketch of Proof:

- Establish joint limit distribution of $V(\boldsymbol{\theta})$ and $V(\boldsymbol{\theta}_0)$:

$$(V_t(\boldsymbol{\theta}), V_t(\boldsymbol{\theta}_0)) \stackrel{D}{\rightarrow} (V_\infty(\boldsymbol{\theta}), V_\infty(\boldsymbol{\theta}_0)).$$

- Apply the ergodicity of $V_t(\boldsymbol{\theta})$ established by Fasen (2009).

- Prove that

$$V_\infty(\boldsymbol{\theta}) \neq V_\infty(\boldsymbol{\theta}_0), \text{ a.s.}$$

Consistency

Next the equicontinuity of $l_T(\boldsymbol{\theta}) = m_4 \frac{\mathscr{L}_T(\boldsymbol{\theta}) - \mathscr{L}_T(\boldsymbol{\theta}_0)}{T}$ is established

Lemma

Assume that $\mathbb{E}L_1^4 < \infty$. Then for some a.s. finite random time T_0 and some small $\delta_0 > 0$ and a positive constant $C > 0$ we have

$$\left\| \frac{\partial l_T}{\partial \boldsymbol{\theta}}(\boldsymbol{\theta}) \right\|_\infty \leq C \text{ a.s.}, \quad \text{for all } \boldsymbol{\theta} \in B_{\delta_0}(\boldsymbol{\theta}_0) \text{ and } T \geq T_0\,. \tag{9}$$

Consistency

Sketch of Proof:

- Make use of monotonicity of partials in all components of parameter θ.
- Apply same rationale as in previous lemma.

Consistency

Finally we obtain the desired consistency result.

Theorem

Assume that $\mathbb{E}L_1^4 < \infty$. Then the quasi-likelihood estimator $\hat{\boldsymbol{\theta}}_T$ is strongly consistent.

Consistency

The problem of showing **local asymptotic normality** remains **unsolved**.

The main technical problem is the **non-linearity** caused by the volatility of the variance process V.

Conclusion

Concluding Remarks:

- Estimation the continuous time GARCH model based on continuous observations is in principle **possible** (though maybe not realistic).
- Proving the usual consistency and LAN is technical and difficult, and moreover **incomplete**!
- In my **personal** opinion, the continuous GARCH is though interesting because:
  - study generalized Ornstein-Uhlenbeck (GOU) processes;
  - understand the unclear limit results obtained for the discrete GARCH;
  - get an idea of time-scaling for the discrete GARCH

Outlook

Outlook I: Extended the simple noise model E to

$$\mathrm{d}Y_t = f(Y_{t-}, \mathbf{X}_{t-}, V_{t-}; \boldsymbol{\theta})\,\mathrm{d}t + \mathrm{d}E_t \ .$$

The optimal quasi-likelihood estimators are

$$G_T^\star(\boldsymbol{\theta}) \;=\; \int_0^T \frac{\dot{f}_{t-}(\boldsymbol{\theta}) + f_{vt-}(\boldsymbol{\theta})\,\dot{V}_{t-}(\boldsymbol{\theta})}{V_{t-}(\boldsymbol{\theta})}\,\mathrm{d}G_{1t}(\boldsymbol{\theta})$$

$$+\int_0^T \frac{\dot{V}_{t-}(\boldsymbol{\theta})\left(1 - m_3\,V_{t-}^{1/2}(\boldsymbol{\theta}) f_{vt-}(\boldsymbol{\theta})\right) - m_3\,V_{t-}^{1/2}(\boldsymbol{\theta})\,\dot{f}_{t-}(\boldsymbol{\theta})}{(m_4 - m_3^2)\,V_{t-}^2(\boldsymbol{\theta})}\,\mathrm{d}G_{2t}(\boldsymbol{\theta}),$$

what is the continuous time version of (14) in Li and Turtle (2000).

Outlook

Outlook II: Study the model for the stock price $S = S_0\,\mathcal{E}(Y)$ with cumulative return process Y given by

$$\mathrm{d}Y_t = \mu(V_{t-})\,\mathrm{d}t + \mathrm{d}E_t\,.$$

The process has to be stopped at

$$\tau = \inf\{t \geq 0 : \Delta E_t \leq -1\} = \inf\{t \geq 0 : V_{t-}\,\Delta L_t \leq -1\}\,.$$

This can be seen as an interesting **boundary crossing problem**.