A new GARCH model with a deterministic time-varying intercept

ESEM 2023, Barcelona

Niklas Ahlgren [1] **Alexander Back** [1] Timo Teräsvirta [2]

[1]Hanken School of Economics

[2]Aarhus University

31 August 2023

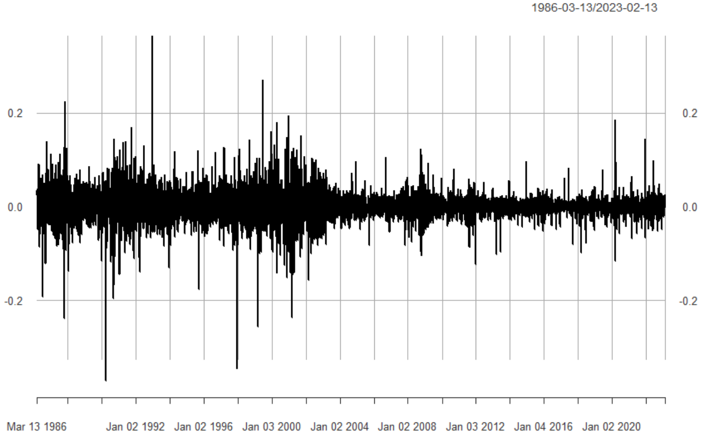
Outline

Introduction

- It is common for long financial time series to exhibit gradual change in the unconditional volatility.
- We propose a new GARCH model that captures this type of nonstationarity in a parsimonious way.
- The model, called the Additive Time-Varying (ATV-)GARCH model, augments the volatility equation by a deterministic time-varying intercept.

ORCL return series

Figure 1: ORCL stock returns 1986–2023.



Contribution

- The ATV-GARCH model is a parsimonious parameterization of a time-varying GARCH (tvGARCH) process.
- The parameters are estimated globally by QML, which leads to asymptotically normal, root-T-consistent estimators.
- The simplicity of the standard parametric likelihood framework of stationary and ergodic GARCH processes is preserved (with some mild strengthening of moment conditions).
- The ATV-GARCH model can be fitted by QML in a single step.

ATV-GARCH

The ATV-GARCH model is defined by augmenting the volatility equation of the GARCH model of Bollerslev (1986) by a deterministic time-varying intercept:

$$X_{t,T} = \sigma_{t,T}(\boldsymbol{\theta})\varepsilon_t, \;\; t = 1, \ldots, T,$$

where ε_t is $\mathrm{IID}(0,1)$, the volatility equation is given by

$$\sigma^2_{t,T}(\boldsymbol{\theta}) = \alpha_0(t/T; \boldsymbol{\theta}) + \alpha_1 X^2_{t-1,T} + \beta_1 \sigma^2_{t-1,T}(\boldsymbol{\theta})$$

and $\boldsymbol{\theta}$ is the vector of parameters.

ATV-GARCH

- The time-varying intercept $\alpha_0(t/T;\boldsymbol{\theta})$ is a Lipschitz continuous function.
- In applications, $X_{t,T}$ will typically be a log-return of a stock or stock index.
- The double subscript (t, T) is used to emphasise that we are working in rescaled time.

ATV-GARCH(p, q)

From now on the observations are assumed to come from an ATV-GARCH(p, q) process with volatility equation

$$\sigma^2_{t,T}(\boldsymbol{\theta}) = g(t/T; \boldsymbol{\theta}) + \sum_{i=1}^{p} \alpha_i X^2_{t-i,T} + \sum_{j=1}^{q} \beta_j \sigma^2_{t-j,T}(\boldsymbol{\theta}),$$

where $\alpha_1, \alpha_2, \ldots, \alpha_p \geq 0$, $\beta_1, \beta_2, \ldots, \beta_q \geq 0$, $g(t/T; \boldsymbol{\theta}) > 0$.

ATV-GARCH

We follow Amado and Teräsvirta (2013) and parameterise $g(t/T;\boldsymbol{\theta})$ as a linear combination of logistic transition functions

$$g(t/T;\boldsymbol{\theta}) = \alpha_0 + \sum_{l=1}^{L} \alpha_{0l} G_l\left(t/T;\gamma_l,\boldsymbol{c}_l\right),$$

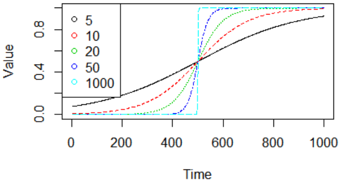
where

$$G(t/T;\gamma,\boldsymbol{c}) = \left(1+\exp\left\{-\gamma\prod_{k=1}^{K}(t/T-c_k)\right\}\right)^{-1}$$

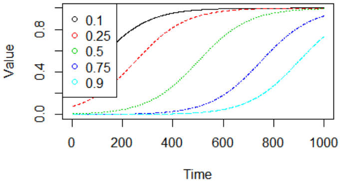
is the general logistic function with $\gamma > 0,\ c_1 < \ldots < c_K$.

Logistic transition functions

Figure 2: Logistic transition functions.



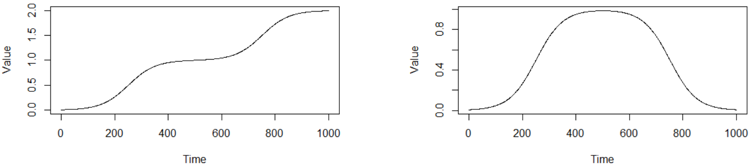
(a) γ varies, c is fixed at 0.5.



(b) c varies, γ is fixed at 10.

Logistic transition functions

Figure 3: Logistic transition functions



(a) Two transitions with $\gamma_1 = \gamma_2 = 20$, $c_1 = 0.25$, $c_2 = 0.75$, no sign change.

(b) Two transitions with $\gamma_1 = \gamma_2 = 20$, $c_1 = 0.25$, $c_2 = 0.75$, one sign change.

ATV-GARCH

- The model is unidentified if $\alpha_{01} = \cdots = \alpha_{0L} = 0$.
- The test of the null hypothesis $\alpha_{01} = \cdots = \alpha_{0L} = 0$, based on approximating the alternative by a Taylor series expansion around the null hypothesis as in Luukkonen, Saikkonen and Teräsvirta (1988), will be detailed in a separate paper.

Related work

- The multiplicative decomposition of Amado and Teräsvirta (2013) is asymptotically equivalent to the ATV-GARCH model.
- In the GARCH-X model of Han and Kristensen (2014) the volatility equation is augmented by a positive-valued function of an exogenous stochastic variable.
- C. W. S. Chen et al. (2014) propose a model where all parameters are time-varying.
- Medeiros and Vega (2009) propose a functional coefficient GARCH model.
- Dahlhaus and Subba Rao (2006) generalise the ARCH model to ARCH models with time-varying parameters.
- Subba Rao (2006) and Chen et al. (2014) consider GARCH models with time-varying parameters.

Locally stationary processes

- The process $\{X_{t,T}\}$ is globally nonstationary but locally stationary.
- Locally stationary processes are globally nonstationary processes that have (asymptotically) locally stationary behaviour.
- The process can locally be approximated by a stationary process.
- Locally stationary linear processes were introduced by Dahlhaus (1997, 2000).
- Locally stationary ARCH and GARCH processes were introduced in Dahlhaus and Subba Rao (2006), and Subba Rao (2006).

Local stationarity

Let $\{X_{t,T}\}$ be a locally stationary process and $\left\{\widetilde{X}_t(u)\right\}$ a stationary process at time point u, $u \in [0,1]$.

By the triangle inequality,

$$\left|X_{t,T} - \widetilde{X}_t(u)\right| \leq \left|X_{t,T} - \widetilde{X}_t(t/T)\right| + \left|\widetilde{X}_t(t/T) - \widetilde{X}_t(u)\right|.$$

Formally, a process $\{X_{t,T}\}$ is said to be locally stationary if (Dahlhaus and Subba Rao (2006)

$$X_{t,T} = \widetilde{X}_t(u) + O_p\left(|t/T - u| + \frac{1}{T}\right).$$

Local stationarity of the ATV-GARCH model

Theorem 2.1 of Subba Rao (2006) states conditions under which a nonstationary, nonlinear process with time-dependent parameters can be locally approximated by a stationary process.

Proposition

Assume that $\mathbb{E}(\varepsilon_t^2) = 1$, *the parameter space* Θ *is compact and* $\sum_{i=1}^{p} \alpha_i + \sum_{j=1}^{q} \beta_j < 1$. *Then the ATV-GARCH model is locally stationary.*

Proof.

Outline of proof. The logistic function is Lipschitz continuous. The result then follows from Subba Rao (2006). □

QML estimation of the ATV-GARCH model

The Gaussian log likelihood function is given by

$$L_T(\boldsymbol{\theta}) = \sum_{t=1}^{T} \ell_t(\boldsymbol{\theta}),$$

where

$$\ell_t(\boldsymbol{\theta}) = -\frac{1}{2}\left[\log h_t(\boldsymbol{\theta}) + \frac{X_t^2}{h_t(\boldsymbol{\theta})}\right],$$

and $h_t(\boldsymbol{\theta})$ is generated recursively as in Berkes, Horvath and Kokoszka (2003).

QML estimation of the ATV-GARCH model

The parameter vector of the ATV-GARCH$(1,1)$ model with one transition function and one c is

$$\boldsymbol{\theta} = (\alpha_0, \alpha_{01}, \gamma_1, c_1, \alpha_1, \beta_1)^T.$$

The QMLE is defined as

$$\widehat{\boldsymbol{\theta}}_T = \arg\max_{\boldsymbol{\theta}\in\Theta} \frac{1}{T}\sum_{t=1}^{T} \ell_t(\boldsymbol{\theta}).$$

Outline of the asymptotics

- We use results from Berkes, Horvath and Kokoszka (2003), and Francq and Zakoian (2004), who proved consistency and asympotic normality of the QMLE of the parameters of stationary GARCH models.
- Dahlhaus, Richter and Wu (2019) (DRW) developed a general theory for nonlinear locally stationary processes.
- To show that the QMLE of the parameters of the ATV-GARCH model is consistent, we mainly need the global law of large numbers of DRW (Theorem 2.7(i)).
- To prove asymptotic normality, we apply the global central limit theorem of DRW (Theorem 2.9) to the score.

Consistency

In order to show consistency of the QMLE, we make the following assumptions.

(A1) The random variables ε_t are IID with $\mathbb{E}(\varepsilon_0) = 0$ and $\mathbb{E}(\varepsilon_0^2) = 1$. $\mathbb{E}\left|\varepsilon_0^2\right|^{1+d} < \infty$, for some $d > 0$. The random variable ε_0^2 is non-degenerate and $\lim_{t\to 0} t^{-\mu}\mathbb{P}\{\varepsilon_0^2 \le t\} = 0$, for some $\mu > 0$.

(A2) The parameter space Θ is compact and $\boldsymbol{\theta}_0 \in \text{int}(\Theta)$.

(A3) $\alpha_0 > 0, \alpha_0 + \sum_{l=1}^{L} \alpha_{0l} G_l\left(u, \gamma_l, \boldsymbol{c}_l\right) > \inf_{\boldsymbol{\theta}\in\Theta} \alpha_0 \ \forall u \in [0,1]$. The functions G_l are non-constant for each l.

(A4) The polynomials $\mathcal{A} = \alpha_1 z + \alpha_2 z^2 + \ldots + \alpha_p z^p$ and $\mathcal{B} = \beta_1 z + \beta_2 z^2 + \ldots + \beta_q z^q$ are coprime on the set of polynomials with real coefficients.

(A5) $\sum_{i=1}^{p} \alpha_i + \sum_{j=1}^{q} \beta_j < 1$.

Consistency

Theorem

Under (A1)–(A5),

$$\widehat{\boldsymbol{\theta}}_T \xrightarrow{P} \boldsymbol{\theta}_0,$$

as $T \to \infty$.

Consistency

Proof outline.

We apply the global LLN (Theorem 2.7(i)) of DRW to the log likelihood function.

For this we need to show that a Lipschitz-type condition applies to $\ell_t(\boldsymbol{\theta})$ (Assumption 2.1(S1) in DRW is satisfied).

By Theorem 2.7(i) of DRW, pointwise,

$$\text{Ł}_T(\boldsymbol{\theta}) \xrightarrow{P} \int_0^1 \ell(u, \boldsymbol{\theta}) \, \mathrm{d}u = \mathbb{E}(\ell_T(\boldsymbol{\theta})).$$

Stochastic equicontinuity and compact parameter space are sufficient conditions for uniform convergence of $L_T(\boldsymbol{\theta})$.

Uniqueness follows from arguments similar to those in the stationary case.

□

Asymptotic normality

In order to show asymptotic normality of the QMLE, we make the following additional assumptions.

(A6) $\mathbb{E}|\varepsilon_t|^{4+d} < \infty$ for some $d > 0$.

(A7) $\lambda_{\text{spec}}[\mathbf{A}_t(u)]_2 < 1 - \delta$ for some $\delta > 0$, where $\mathbf{A}_t(u)$ is the coefficient matrix in the state space representation of a time-varying GARCH process.

Theorem

Under (A1)–(A7),

$$\sqrt{T}\left(\widehat{\boldsymbol{\theta}}_T - \boldsymbol{\theta}_0\right) \xrightarrow{D} N(\mathbf{0}, \mathbf{B}^{-1}\mathbf{A}\mathbf{B}^{-1}),$$

as $T \to \infty$.

Asymptotic normality

The expressions for $\mathbf{A}$ and $\mathbf{B}$ are integrals of the stationary approximation of the score and Hessian at time point u, $u \in [0,1]$, respectively,

$$\mathbf{A} = \int_0^1 \mathbb{E}(\mathbf{A}(u))\, \mathrm{d}u,$$

where

$$\mathbf{A}(u) = \sum_{k \in \mathbf{Z}} \mathrm{cov}(\widetilde{\mathbf{s}}_0(u,\theta_0), \widetilde{\mathbf{s}}_k(u,\theta_0))$$

and

$$\mathbf{B} = \int_0^1 \mathbb{E}(\widetilde{\mathbf{H}}(u,\theta_0))\, \mathrm{d}u.$$

Asymptotic normality

Proof outline.

We prove asymptotic normality of the QMLE based on a standard Taylor series expansion argument.

The idea is to apply the global CLT (Theorem 2.9) of DRW to the score and the global LLN (Theorem 2.7(i)) to the Hessian.

We need to show that the score $\mathbf{s}_{t,T}(\boldsymbol{\theta}_0)$ is locally stationary and Lipschitz continuous in the parameters in the two-norm sense (Assumption 2.1(S1), (S2) in DRW is satisfied) and that a mixing condition on the locally stationary approximation of the score $\widetilde{\mathbf{s}}_t(u, \boldsymbol{\theta})$ is fulfilled (Assumption 2.3(M1) in DRW is satisfied).

The arguments for the Hessian are the same as for the log likelihood function in the proof of consistency.

□

Simulation design

We consider three data-generating processes (DGPs)

$$X_{t,T} = \sigma_{t,T}(\boldsymbol{\theta})\varepsilon_t,$$

where

$$\sigma_{t,T}^2(\boldsymbol{\theta}) = g(t/T, \boldsymbol{\theta}) + 0.1X_{t-1,T}^2 + 0.8\sigma_{t-1,T}^2$$

and ε_t is NID(0,1) for all t.

DGP 1: $g(t/T, \boldsymbol{\theta}) = 0.05 + 0.15G_1\left(\frac{t}{T}, 20, 0.5\right).$

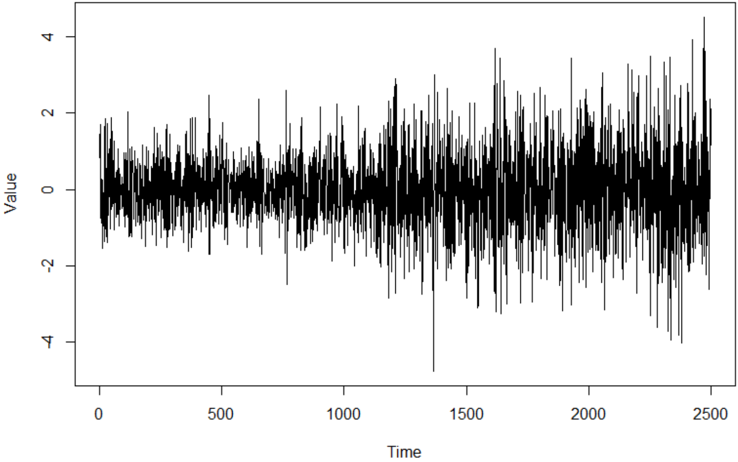
DGP 2: $g(t/T, \boldsymbol{\theta}) = 0.05 + 0.05G_1\left(\frac{t}{T}, 20, 0.5\right).$

DGP 3: $g(t/T, \boldsymbol{\theta}) = 0.05 + 0.05G_1\left(\frac{t}{T}, 10, 0.5\right).$

The time series lengths are $T = 1000$, 2500 and 5000.

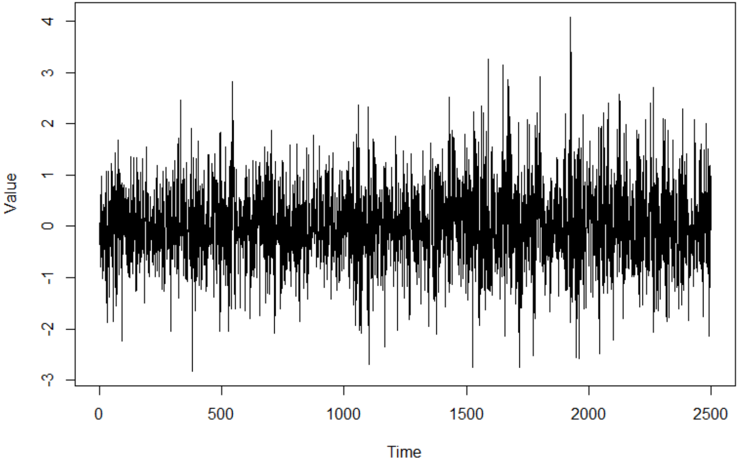
DGP 1

Figure 4: A realisation from DGP 1.



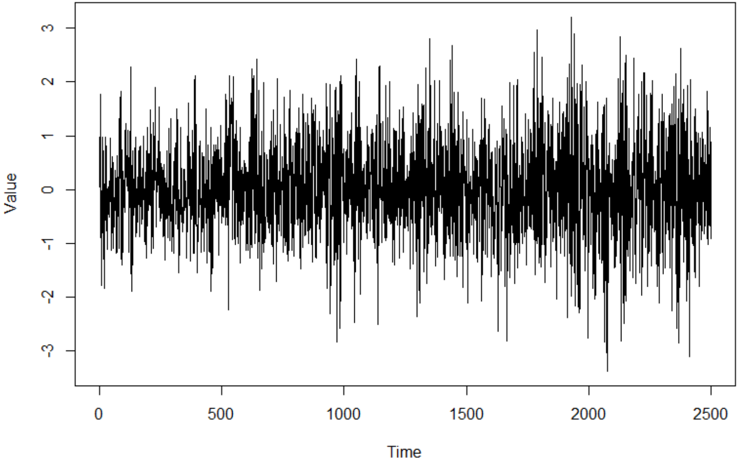
DGP 2

Figure 5: A realisation from DGP 2.



DGP 3

Figure 6: A realisation from DGP 3.



QML parameter estimates DGP 1

Table 1: True values and empirical means of the QML parameter estimates.

Parameter	α_0	α_1	β_1	γ_1	c_1	α_{01}
DGP 1						
True value	**0.05**	**0.1**	**0.8**	**20**	**0.5**	**0.15**
Mean 1000	0.073	0.101	0.747	19.921	0.510	0.241
Mean 2500	0.059	0.101	0.781	20.882	0.5	0.177
Mean 5000	0.054	0.1	0.792	20.277	0.5	0.163

QML parameter estimates DGP 2

Table 2: True values and empirical means of the QML parameter estimates.

Parameter	α_0	α_1	β_1	γ_1	c_1	α_{01}
DGP 2						
True value	**0.05**	**0.1**	**0.8**	**20**	**0.5**	**0.05**
Mean 1000	0.071	0.101	0.747	20.692	0.52	0.095
Mean 2500	0.057	0.101	0.781	19.408	0.508	0.065
Mean 5000	0.053	0.1	0.791	19.747	0.504	0.057

QML parameter estimates DGP 3

Table 3: True values and empirical means of the QML parameter estimates.

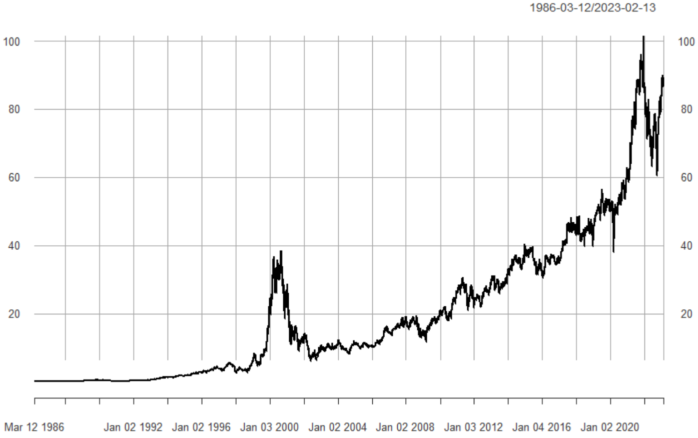
Parameter	α_0	α_1	β_1	γ_1	c_1	α_{01}
DGP 3						
True value	**0.05**	**0.1**	**0.8**	**10**	**0.5**	**0.05**
Mean 1000	0.07	0.101	0.746	12.369	0.538	0.105
Mean 2500	0.055	0.101	0.781	9.788	0.531	0.075
Mean 5000	0.052	0.1	0.791	9.582	0.519	0.063

Empirical application

- We illustrate the use of the ATV-GARCH model by a small empirical application to Oracle Corporation (ORCL) stock returns.
- The data consist of ORCL stock prices from the beginning of its listing in 1986 to February 2023.
- The time series of log returns contains $T = 9306$ observations.

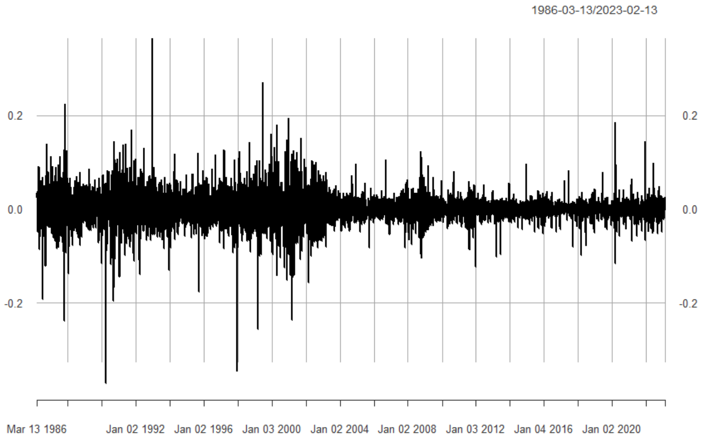
ORCL price series

Figure 7: ORCL stock price 1986–2023.



ORCL return series

Figure 8: ORCL stock returns 1986–2023.



Summary statistics and LM tests

Table 4: Summary statistics and LM tests for ORCL stock returns 1986–2023.

Mean	Sd	Med	Min	Max	Skew	Kurt
0.001	0.029	0.000	-0.372	0.364	-0.168	16.871

H_0	$\chi^2(3)$	p	$\chi^2(3)$	p	$\widehat{\alpha}_1 + \widehat{\beta}_1$
0	34.582	< 0.000	27.281	< 0.000	> 0.999
1	4.641	0.200	3.841	0.275	0.940

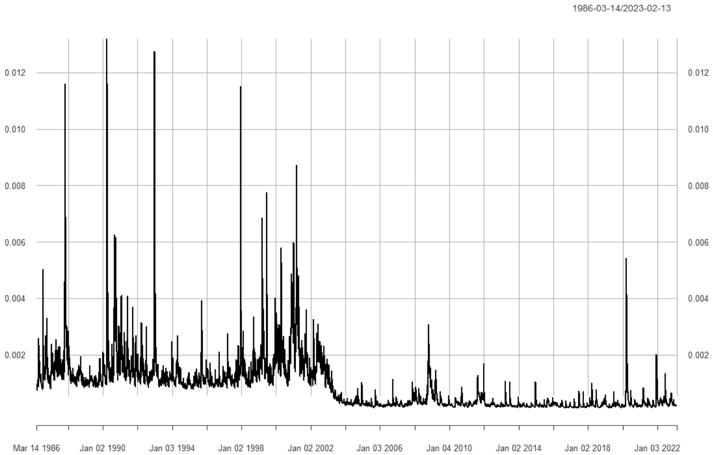
QML parameter estimates

Table 5: Parameter estimates of the ATV-GARCH$(1,1)$ model for ORCL.

Parameter	Estimate	se	Robust se
α_0	0.010	0.000	0.000
α_{01}	-0.008	0.000	0.000
γ_1	0.994	0.003	0.004
c_1	0.460	0.003	0.004
α_1	0.087	0.002	0.002
β_1	0.854	0.003	0.005

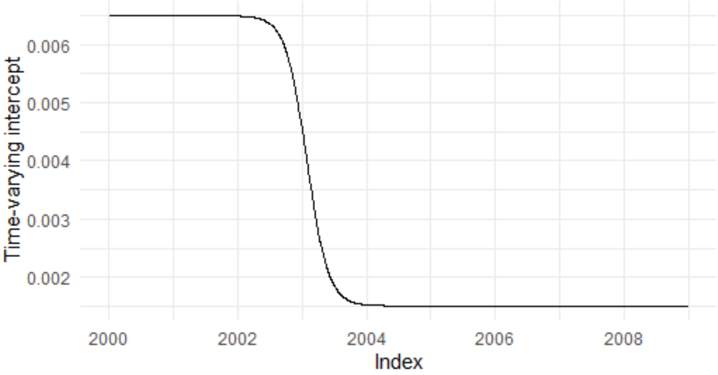
ORCL conditional variance

Figure 9: ORCL fitted conditional variance from the ATV-GARCH$(1,1)$ model.



ORCL transition function

Figure 10: ORCL fitted transition function from the ATV-GARCH$(1,1)$ model.



Summary

- We propose a GARCH model augmented by a time-varying deterministic intercept.
- We use the theory of locally stationary processes to prove consistency and asymptotic normality of the QMLE of the parameters.
- The results from a simulation study indicate that the parameters of the model can be estimated accurately for moderately large sample sizes.
- The results from an empirical application indicate that the persistence of standard GARCH models are reduced by including a time-varying intercept in the volatility equation.

Thank you!

Contact

Alexander Back
Hanken School of Economics
E-mail: alexander.back@hanken.fi