Financial Econometrics A | Final Exam | January 6th, 2017

Solution Key

Question A:

Consider the following log-linear Realized GARCH model given by

$$x_t = \sigma_t z_t, \tag{A.1}$$

with $z_t \sim i.i.d.N\left(0,1\right)$, and

$$\log(\sigma_t^2) = 1 + \alpha \log(y_{t-1}), \tag{A.2}$$

$$\log(y_t) = \gamma + \phi \log(\sigma_t^2) + u_t, \tag{A.3}$$

with $u_t \sim i.i.d.N\left(0,1\right)$ and $\alpha, \gamma, \phi \in \mathbb{R}$. It is assumed that the processes (z_t) and (u_t) are independent. Here y_t is some *observed* positive exogenous covariate as for example the realized volatility.

Question A.1: Use the drift criterion to show that $\log(y_t)$ is weakly mixing with $E[(\log(y_t))^2] < \infty$, if $|\alpha\phi| < 1$.
Given that $\log(y_t)$ is weakly mixing we do also have that the joint process $(x_t, \log(y_t))$ is weakly mixing.

Solution: Substituting (A.2) into (A.3) yields $\log(y_t) = \gamma+\phi+\phi\alpha \log(y_{t-1})+ u_t$. Hence, $\log(y_t)$ is an AR(1) process with an intercept and a Gaussian (i.i.d.) error term. The drift criterion, with drift function $\delta(x) = 1 + x^2$, is established via standard arguments for the AR(1) process. It should be mentioned that $\log(y_t)$ has a positive and continuous conditional density. Derivations should be included.

Question A.2: Let $\theta = (\alpha, \gamma, \phi)$ denote the model parameters. Given a sample $(x_t, \log(y_t))$, $t = 0, 1, ..., T$, the joint log-likelihood is (up to a constant term and a scaling factor)

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

$$l_t(\theta) = -\log(\sigma_t^2(\theta)) - \frac{x_t^2}{\sigma_t^2(\theta)} - \left[\log(y_t) - \gamma - \phi \log(\sigma_t^2(\theta))\right]^2,$$

where $\log(\sigma_t^2(\theta)) = 1 + \alpha \log(y_{t-1})$.
Show that

$$\frac{\partial l_t(\theta)}{\partial \alpha} = \left\{ \frac{x_t^2}{\sigma_t^2(\theta)} - 1 + 2\phi \left[\log(y_t) - \gamma - \phi \log(\sigma_t^2(\theta))\right] \right\} \log(y_{t-1}).$$

Hint: You may want to use that

$$\frac{\partial l_t(\theta)}{\partial \alpha} = \frac{\partial l_t(\theta)}{\partial \log(\sigma_t^2(\theta))} \frac{\partial \log(\sigma_t^2(\theta))}{\partial \alpha}.$$

Solution: Using the hint, the result follows directly by observing that

$$\frac{\partial l_t(\theta)}{\partial \log(\sigma_t^2(\theta))} = -1 + \frac{x_t^2}{\sigma_t^2(\theta)} + 2\phi \left[\log(y_t) - \gamma - \phi \log(\sigma_t^2(\theta))\right]$$

and

$$\frac{\partial \log(\sigma_t^2(\theta))}{\partial \alpha} = \log(y_{t-1}).$$

Question A.3: Let $\theta_0 = (\alpha_0, \gamma_0, \phi_0)$ denote the vector of true parameter values. Define $S_T(\theta) = \partial L_T(\theta)/\partial \alpha$.
Assume that $(x_t, \log(y_t))$ is weakly mixing and satisfies the drift criterion such that $E[(\log(y_{t-1}))^2] < \infty$. Show that

$$\frac{1}{\sqrt{T}} S_T(\theta_0) \xrightarrow{d} N(0, v), \tag{A.4}$$

where $v = (2 + 4\phi_0^2) E[(\log(y_{t-1}))^2]$.
Explain briefly what the property (A.4) can be used for.

Hint: Use that $\log(y_t) - \gamma_0 - \phi_0 \log(\sigma_t^2(\theta_0)) = u_t$. Moreover, you may want to recall that $E[z_t^4] = 3$.

Solution: The result is established by verifying the conditions of the CLT for weakly mixing processes (Theorem II.1 from the lecture notes). It holds that $S_T(\theta_0) = \sum_{t=1}^T f(x_t, \log(y_t), x_{t-1}, \log(y_{t-1}))$, with

$$\begin{aligned} f(x_t, \log(y_t), x_{t-1}, \log(y_{t-1})) &= \left\{-1 + \frac{x_t^2}{\sigma_t^2(\theta_0)} + 2\phi_0 \left[\log(y_t) - \gamma_0 - \phi_0 \log(\sigma_t^2(\theta_0))\right]\right\} \log(y_{t-1}) \\ &= \left\{z_t^2 - 1 + 2\phi_0 u_t\right\} \log(y_{t-1}). \end{aligned}$$

Hence the CLT is satisfied if $E[\{z_t^2 - 1 + 2\phi_0 u_t\} \log(y_{t-1}) | x_{t-1}, \log(y_{t-1})] = 0$ and $E[| \{z_t^2 - 1 + 2\phi_0 u_t\} \log(y_{t-1})|^2] < \infty$. These conditions hold since (1) $E[\{z_t^2 - 1 + 2\phi_0 u_t\}] = 0$, (2) $\{z_t^2 - 1 + 2\phi_0 u_t\}$ and $(x_{t-1}, \log(y_{t-1}))$ are independent, (3) $E[\{z_t^2 - 1 + 2\phi_0 u_t\}^2] = 2 + 4\phi_0^2 < \infty$, and (4) $E[(\log(y_{t-1}))^2] < \infty$. Details and derivations should be provided.

The property (A.4) is important for obtaining the asymptotic distribution of the maximum likelihood estimator, and hence for testing hypotheses about θ. The very good answer would mention the remaining regularity conditions of Theorem III.2, related to the second- and third-order derivatives of the log-likelihood function.

Question A.4: For the model (A.1)-(A.3), the one-period VaR at risk level κ, $\text{VaR}_{T,1}^{\kappa}$, is defined as

$$P_T(x_{T+1} < -\text{VaR}_{T,1}^{\kappa}) = \kappa, \quad \kappa \in (0,1),$$

where $P_T(\cdot)$ denotes the conditional distribution of x_{T+1}. It can be shown (but do not do so) that

$$\text{VaR}_{T,1}^{\kappa} = -\sigma_{T+1}\Phi^{-1}(\kappa),$$

where $\Phi^{-1}(\cdot)$ denotes the inverse cdf of the standard normal distribution. Explain briefly how you would compute an estimate of $\text{VaR}_{T,1}^{\kappa}$.

Solution: Given an estimate of $\theta = (\alpha, \gamma, \phi)$, denoted $\hat{\theta} = (\hat{\alpha}, \hat{\gamma}, \hat{\phi})$, obtained by maximum likelihood (or some other method), an estimate of σ_{T+1} is given by

$$\hat{\sigma}_{T+1} = \sqrt{\exp[1 + \hat{\alpha}\log(y_T)]},$$

where y_T is part of the data set. For given $\kappa \in (0,1)$, $\Phi^{-1}(\kappa)$ is known, since $\Phi^{-1}(\cdot)$ denotes the inverse cdf of the standard normal distribution. An estimate of $\text{VaR}_{T,1}^{\kappa}$ is thus computed as $-\hat{\sigma}_{T+1}\Phi^{-1}(\kappa)$. One might observe that we would only need $\hat{\alpha}$ in order to obtain an estimate of $\text{VaR}_{T,1}^{\kappa}$. This means that one can ignore modelling the dynamics of y_t.

Question B:

Consider the following switching model given by

$$x_t = \mu 1_{(s_t=1)} + \varepsilon_t, \tag{B.1}$$

where μ is an $\mathbb{R}$-valued constant and s_t can take value 1 or 2. Moreover, $\varepsilon_t \sim i.i.d.N(0, \sigma^2)$, and we assume that the processes (s_t) and (ε_t) are independent. Suppose that s_t is a two-state Markov chain with transition probabilities $P(s_t = j|s_{t-1} = i) = p_{ij}$, $i, j = 1, 2$.
Note that $1_{(s_t=1)} = 1$ if $s_t = 1$ and $1_{(s_t=1)} = 0$ if $s_t = 2$.

Question B.1: Suppose that $\mu = 0$. Explain if x_t is weakly mixing.
What should hold for p_{11} and p_{22} for s_t to be weakly mixing?

Solution: If $\mu = 0$, $x_t = \varepsilon_t \sim i.i.d.N(0, \sigma^2)$, and hence x_t is weakly mixing. The Markov chain s_t is weakly mixing if p_{11}, $p_{22} < 1$ and $p_{11} + p_{22} > 0$.

Question B.2: Next, assume that s_t is *observed*. Moreover, suppose that the transition probabilities satisfy $p_{11} = (1 - p_{22}) = p \in (0, 1)$ such that s_t is and *i.i.d.* process with $P(s_t = 1) = p$ and $P(s_t = 2) = 1 - p$.
Show that for $t \geq 1$, the joint conditional density of (x_t, s_t) is

$$\begin{aligned} f(x_t, s_t|x_{t-1}, s_{t-1}, ..., x_0, s_0) = &\left[\frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(x_t - \mu)^2}{2\sigma^2}\right) p\right]^{1_{(s_t=1)}} \\ &\times \left[\frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{{x_t}^2}{2\sigma^2}\right) (1 - p)\right]^{1_{(s_t=2)}}. \end{aligned}$$

Solution: Since s_t is i.i.d., (x_t, s_t) and $(x_{t-1}, s_{t-1}, ..., x_0, s_0)$ are independent. Hence,

$$\begin{aligned} f(x_t, s_t|x_{t-1}, s_{t-1}, ..., x_0, s_0) &= f(x_t, s_t) \\ &= f(x_t|s_t)f(s_t) \end{aligned}$$

We have that

$$f(s_t) = p^{1_{(s_t=1)}}(1 - p)^{1_{(s_t=2)}}.$$

Moreover, using that $\varepsilon_t \sim N(0, \sigma^2)$,

$$\begin{aligned} f(x_t|s_t = 1) &= \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(x_t - \mu)^2}{2\sigma^2}\right), \\ f(x_t|s_t = 2) &= \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{{x_t}^2}{2\sigma^2}\right), \end{aligned}$$

and hence that

$$f(x_t|s_t) = \left[\frac{1}{\sqrt{2\pi\sigma^2}}\exp\left(-\frac{(x_t-\mu)^2}{2\sigma^2}\right)\right]^{1_{(s_t=1)}} \times \left[\frac{1}{\sqrt{2\pi\sigma^2}}\exp\left(-\frac{{x_t}^2}{2\sigma^2}\right)\right]^{1_{(s_t=2)}}.$$

Question B.3: Maintaining the assumptions from Question B.2, let $\theta = (\mu, \sigma^2, p)$ denote the model parameters. The log-likelihood function is

$$L_T(\theta) = \sum_{t=1}^{T}\left\{\log(p) - \frac{1}{2}\log(2\pi\sigma^2) - \frac{(x_t-\mu)^2}{2\sigma^2}\right\}1_{(s_t=1)} + \sum_{t=1}^{T}\left\{\log(1-p) - \frac{1}{2}\log(2\pi\sigma^2) - \frac{{x_t}^2}{2\sigma^2}\right\}1_{(s_t=2)}.$$

Let $\hat{\mu}$ denote the maximum likelihood estimator for μ.
Show that

$$\hat{\mu} = \frac{\sum_{t=1}^{T} x_t 1_{(s_t=1)}}{\sum_{t=1}^{T} 1_{(s_t=1)}}.$$

Moreover, let $\hat{p}$ denote the maximum likelihood estimator for p. Derive $\hat{p}$ and argue that $\hat{p} \xrightarrow{P} p$ as $T \to \infty$.

Solution: The expression for $\hat{\mu}$ is obtained by solving the F.O.C. for the maximization of $L_T(\theta)$ with respect to μ, i.e. by solving $\partial L_T(\theta)/\partial\mu = 0$ for μ. Derivations should be included.
Likewise, solving $\partial L_T(\theta)/\partial p = 0$ for p, and using that $1_{(s_t=2)} = 1 - 1_{(s_t=1)}$, yields

$$\hat{p} = \frac{1}{T}\sum_{t=1}^{T} 1_{(s_t=1)}.$$

Since s_t is *i.i.d.*, we have that $1_{(s_t=1)}$ is *i.i.d.* with $E[1_{(s_t=1)}] = P(s_t = 1) = p < \infty$, it holds by the LLN for *i.i.d.* processes that

$$\hat{p} = \frac{1}{T}\sum_{t=1}^{T} 1_{(s_t=1)} \xrightarrow{P} E[1_{(s_t=1)}].$$

Derivations should be included.

Question B.4: Suppose that the process (s_t) is *unobserved*, but does still satisfy the *i.i.d.* assumption, i.e. $p_{11} = (1 - p_{22}) = p \in (0, 1)$. Then the estimators derived in Question B.3 are infeasible. Instead we may introduce

$$\tilde{L}_T(\theta) = E[L_T(\theta)|x_1, ...x_T].$$

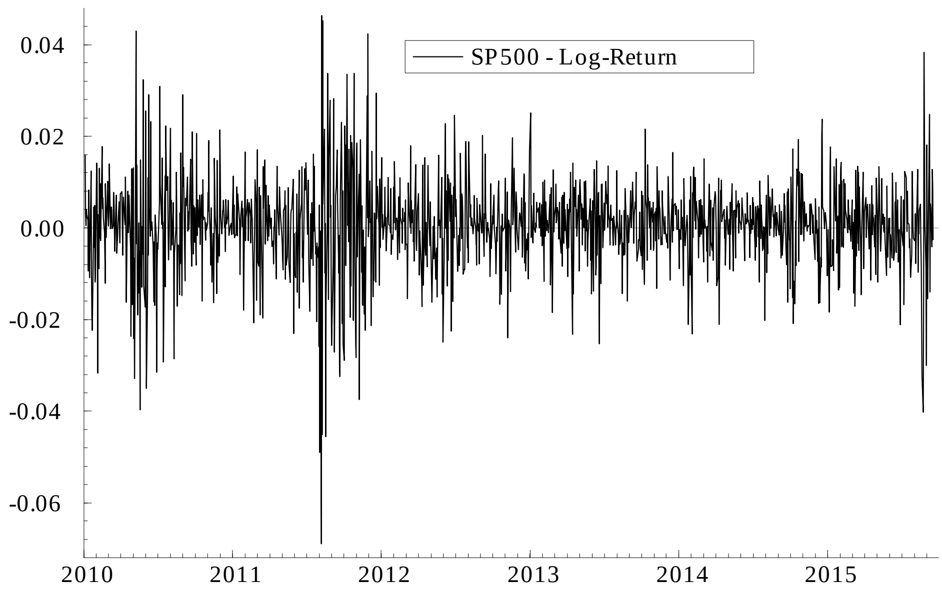
It holds that

$$\tilde{L}_T(\theta) = \sum_{t=!}^{T} \left\{ \log(p) - \frac{1}{2}\log(2\pi\sigma^2) - \frac{(x_t-\mu)^2}{2\sigma^2} \right\} P_t^\star(1)$$
$$+ \sum_{t=1}^{T} \left\{ \log(1-p) - \frac{1}{2}\log(2\pi\sigma^2) - \frac{{x_t}^2}{2\sigma^2} \right\} (1 - P_t^\star(1)),$$

where $P_t^\star(1) = P(s_t = 1|x_t)$.
Explain briefly the role of $\tilde{L}_T(\theta)$ for the estimation of θ.

Solution: This question is about the EM algorithm. Given $P_t^\star(1)$ (E-step), an estimate of θ is obtained by maximizing $\tilde{L}_T(\theta)$ (M-step). It should be noted that the computation of $P_t^\star(1)$ relies on an initial guess of θ, say $\tilde{\theta}$. Clearly, the estimate of θ will depend on $\tilde{\theta}$ through $P_t^\star(1)$. Hence, one may apply the EM algorithm iteratively, by using the estimate of θ for the computation of $P_t^\star(1)$, and then obtain a new estimate of θ. Ideally, a brief outline of this should be included.

Question B.5: The following figure shows the daily log-returns of the S&P 500 index for the period January 4, 2010 to September 17, 2015.



Discuss briefly whether the switching model in (B.1) is adequate for modelling the main features of the log-returns. Would another type of Markov switching model be more suitable?

Solution: By visual inspection of the series, it appears that the returns are heteroskedastic. Hence a model for a switching level, as (B.1), does not appear to be appropriate. Instead, as studied during the lectures and problem sessions, a model for switching variance may be more useful. Specifically, a switching volatility model is given by

$$
\begin{aligned}
x_t &= \varepsilon_t, \\
\varepsilon_t &= \sigma_t z_t, \quad z_t \sim i.i.d.N(0,1) \\
\sigma_t^2 &= \tilde{\sigma}_1^2 1_{(s_t=1)} + \tilde{\sigma}_2^2 1_{(s_t=2)},
\end{aligned}
$$

with $\tilde{\sigma}_1^2$ and $\tilde{\sigma}_2^2$ positive constants, (s_t) a two-state Markov chain, and with (s_t) and (z_t) independent. Alternatively, a switching ARCH model may also serve as a good model for the returns, as studied by Cai (1994, JBES).