Econometrics II - EXAM Outline Solutions

All questions have 25pts

Answer each question in separate sheets

1. *Consider the two linear simultaneous equations* $(G = 2)$ *with two exogeneous variables* $(K = 2)$,

$$\begin{aligned} y_1\gamma_{11} + y_2\gamma_{12} + x_1\delta_{11} + x_2\delta_{12} &= u_1 \\ y_1\gamma_{21} + y_2\gamma_{22} + x_1\delta_{21} + x_2\delta_{22} &= u_2 \end{aligned}$$

where, $u = (u_1, u_2)'$,

$$E[\mathbf{uu}'] = \Sigma = \begin{bmatrix} \sigma_1^2 & \sigma_{12} \\ \sigma_{21} & \sigma_2^2 \end{bmatrix}, \quad \Gamma = \begin{bmatrix} \gamma_{11} & \gamma_{21} \\ \gamma_{12} & \gamma_{22} \end{bmatrix}.$$

(a) *Using the standard normalization, write the general form of the order and rank conditions for single equation identification and for system identification. Then, stating the restrictions on the system parameters, check the identification of the above system in the following cases.*

Imposing the normalization $\gamma_{11} = \gamma_{22} = 1$, these are:

Single equation: rank$[\mathbf{R}_1\mathbf{B}] = G - 1$, with $\mathbf{R}_1\boldsymbol{\beta}_1 = \mathbf{0}$. Order condition: rank$[\mathbf{R}_1] \geq G - 1$.

System: rank$[\mathbf{R}(\mathbf{I}_G \otimes \mathbf{B})] = G(G-1)$, with restrictions $\mathbf{R}\boldsymbol{\beta} = \mathbf{0}$, $\boldsymbol{\beta} =$vec$[\mathbf{B}]$. Order condition: rank$[\mathbf{R}] \geq G(G-1)$.

(b) *SUR model.*

$\gamma_{12} = \gamma_{21} = 0$. System always (just) identified, taking for granted that $E[\mathbf{xx}']$ is full rank and $E[u_i\mathbf{x}] = 0$, $i = 1, 2$, $\mathbf{x} = (x_1, x_2)'$.

For equation 1:

$$\mathbf{R}_1\mathbf{B} = \begin{pmatrix} 0 & 1 & 0 & 0 \end{pmatrix} \begin{pmatrix} \gamma_{11} & \gamma_{21} \\ \gamma_{12} & \gamma_{22} \\ \delta_{11} & \delta_{21} \\ \delta_{12} & \delta_{22} \end{pmatrix} = \begin{pmatrix} \gamma_{12} & \gamma_{22} \end{pmatrix} = \begin{pmatrix} 0 & 1 \end{pmatrix}$$

using the normalization and the exclusion restrictions, which is of rank $G - 1 = 1$.

(c) *A different exogeneous variable is omitted from each structural equation.*

For example, $\delta_{12} = \delta_{21} = 0$. Each equation (just) identified if these variables are in the other equations, i.e. if $\delta_{22} \neq 0$ (for eq. 1) and $\delta_{11} \neq 0$ (for eq.2).

(d) *The variable* x_2 *does not appear in the system.*

$\delta_{12} = \delta_{22} = 0$. Rank conditions fail: no single equation identified.

(e) *Neither* x_1 *nor* x_2 *appear in the first equation.*

$\delta_{11} = \delta_{12} = 0$. First equation overidentified if $\delta_{21}\delta_{22} \neq 0$ (just identified if $\delta_{21} \neq 0$ or and $\delta_{22} \neq 0$). Second equation not identified.

(f) Γ *is constrained to be symmetric and the coefficient in* x_1 *is the same in both equations.*

$\gamma_{12} = \gamma_{21},\ \ \delta_{11} = \delta_{21}$. This corresponds to

$$\mathbf{R} = \begin{pmatrix} 0 & 1 & 0 & 0 & -1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & -1 & 0 \end{pmatrix},$$

so, with $\boldsymbol{\beta} = (\gamma_{11}, \gamma_{12}, \delta_{11}, \delta_{12}, \gamma_{21}, \gamma_{22}, \delta_{21}, \delta_{22})'$, we obtain that

$$\mathbf{R}\left(\mathbf{I_G} \otimes \mathbf{B}\right) = \begin{pmatrix} \gamma_{12} & \gamma_{22} & -\gamma_{11} & -\gamma_{21} \\ \delta_{11} & \delta_{21} & -\delta_{11} & -\delta_{21} \end{pmatrix} = \begin{pmatrix} \gamma_{12} & -1 & 1 & -\gamma_{12} \\ \delta_{11} & \delta_{11} & -\delta_{11} & -\delta_{11} \end{pmatrix}$$

which is of rank $2\,(2-1) = 2$, if any two vectors are linearly independent, ie. if $\delta_{11} \neq 0$ and $\gamma_{12} \neq -1$, so the system would be (just) identified.

(g) Γ *is constrained to be lower triangular with diagonal elements equal to 1 and* Σ *is diagonal. In this case explain how you would estimate the structural form parameters from the estimation of the reduced form. Are these estimates efficient in general? And if you further assume the restrictions on (a)?*

$\gamma_{12} = 0$: this is a triangular system: (just) identified (since y_1 is exogeneous in the second equation).

We can estimate the reduced form

$$\mathbf{y} = \mathbf{\Pi}'\mathbf{x} + \mathbf{v},$$

where $\mathbf{\Pi} = \mathbf{\Delta}\mathbf{\Gamma}^{-1}$, by OLS and the covariance matrix $\mathbb{E}\left[\mathbf{v}\mathbf{v}'\right] = \mathbf{\Lambda} = \mathbf{\Gamma}'^{-1}\mathbf{\Sigma}\mathbf{\Gamma}^{-1}$ using OLS residuals. In this case we have that

$$\Gamma = \begin{bmatrix} \gamma_{11} & \gamma_{21} \\ \gamma_{12} & \gamma_{22} \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ \gamma_{12} & 1 \end{bmatrix}, \quad \text{so} \quad \Gamma^{-1} = \begin{bmatrix} 1 & 0 \\ -\gamma_{12} & 1 \end{bmatrix},$$

and

$$\mathbf{\Sigma} = \begin{bmatrix} \sigma_1^2 & \sigma_{12} \\ \sigma_{21} & \sigma_2^2 \end{bmatrix} = \begin{bmatrix} \sigma_1^2 & 0 \\ 0 & \sigma_2^2 \end{bmatrix}$$

and therefore from $\mathbf{\Pi} = \mathbf{\Delta}\mathbf{\Gamma}^{-1}$ we obtain 4 equations and from $\mathbf{\Lambda} = \mathbf{\Gamma}'^{-1}\mathbf{\Sigma}\mathbf{\Gamma}^{-1}$ we obtain another 3 equations (because of symmetry), and we have 7 unknowns (1 element in $\mathbf{\Gamma}$, 4 in $\mathbf{\Delta}$. and 2 in $\mathbf{\Sigma}$). In particular note that

$$\begin{aligned} \mathbf{\Gamma}'^{-1}\mathbf{\Sigma}\mathbf{\Gamma}^{-1} &= \begin{bmatrix} 1 & -\gamma_{12} \\ 0 & 1 \end{bmatrix} \begin{bmatrix} \sigma_1^2 & 0 \\ 0 & \sigma_2^2 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ -\gamma_{12} & 1 \end{bmatrix} \\ &= \begin{bmatrix} \sigma_1^2 & -\gamma_{12}\sigma_2^2 \\ 0 & \sigma_2^2 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ -\gamma_{12} & 1 \end{bmatrix} = \begin{bmatrix} \sigma_1^2\left(1-\gamma_{12}^2\right) & -\gamma_{12}\sigma_2^2 \\ -\gamma_{12}\sigma_2^2 & \sigma_2^2 \end{bmatrix}, \end{aligned}$$

we obtain three equations for three unknowns $\left(\sigma_1^2, \gamma_{12}, \sigma_2^2\right)$.

(h) *Which is your recommended estimation method in each case?*

When the system is identified, and if we can not assume some form of conditional homoskedasticity, i.e. $E\left[\mathbf{u}\ \mathbf{u}'|\mathbf{x}\right] = E\left[\mathbf{u}\ \mathbf{u}'\right] = \Sigma$, then we should rely on system efficient system GMM (chi-square) estimates using $\hat{W}_n = \hat{E}\left[\mathbf{X}\mathbf{u}\ \mathbf{u}'\mathbf{X}'\right]^{-1}$ with $\mathbf{X} = (I_2 \otimes \mathbf{x})$ and the corresponding restrictions. Otherwise we could also use 3SLS.

In case (e) we can only use efficient single equation GMM for the first one.

In the just identified cases we have that system GMM is equal to System IV (and 3SLS to 2SLS).

In (b) all estimates are equal to SOLS, and therefore to equation by equation OLS (as in (g) if we impose diagonal $\mathbf{\Sigma}$).

2. *Consider a nonlinear SUR system,*

$$E[y_g|\mathbf{z}] = E[y_g|\mathbf{z}_g] = m_g(\mathbf{z}_g, \theta_{0g}), \qquad g = 1, \ldots, G.$$

(a) *Propose a single equation (consistent) estimate $\tilde{\theta}_{n,g}$ for the parameters θ_{0g} of each equation given a random sample of n observations. State the additional regularity conditions you may need.*

NLS equation by equation,

$$\tilde{\theta}_{n,g} = \arg\min_{\theta_g \in \Theta_g} Q_n(\theta_g) = \arg\min_{\theta_g \in \Theta_g} \frac{1}{2} E_n\left[(y_g - m_g(\mathbf{z}_g, \theta_g))^2\right].$$

Given identification of θ_{0g} (with conditions such as $\theta_{0g} = \arg\min_{\theta_g \in \Theta_g} E\left[(y_g - m_g(\mathbf{z}_g, \theta_g))^2\right]$) and consistency of $\tilde{\theta}_{n,g}$ (so $Q_n(\theta_g)$ converges uniformly in Θ_g to $E\left[(y_g - m_g(\mathbf{z}_g, \theta_g))^2\right]$) to study the asymptotic distribution of $\tilde{\theta}_{n,g}$ we need that m_g is smooth (with second continuous derivatives) and that $u_g \nabla_\theta m_g(\mathbf{z}_g, \theta_{0g})$ satisfies a CLT, i.e.,

$$n^{1/2} \frac{\partial}{\partial \theta'} Q_n(\theta_{0g}) = -n^{1/2} E_n[u_g \nabla_\theta m_g(\mathbf{z}_g, \theta_{0g})] \rightarrow_d N(0, \mathbf{D}_g),$$

where $\mathbf{D}_g = V[u_g \nabla_\theta m_g(\mathbf{z}_g, \theta_{0g})]$, $u_g = y_g - E[y_g|\mathbf{z}]$, and that

$$p\lim \frac{\partial^2}{\partial \theta \partial \theta'} Q_n\left(\hat{\theta}_{ng}\right) = \mathbf{E}_g, \quad \hat{\theta}_{ng} \rightarrow_p \theta_{0g},$$

where $\mathbf{E}_g = E\left[\nabla_\theta m_g(\mathbf{z}_g, \theta_{0g}) \nabla_\theta m_g(\mathbf{z}_g, \theta_{0g})'\right]$ is not singular, so we obtain that

$$n^{1/2}\left(\tilde{\theta}_{n,g} - \theta_{0g}\right) \rightarrow_d N\left(0, \mathbf{E}_g^{-1} \mathbf{D}_g \mathbf{E}_g^{-1}\right).$$

In case of homoskedasticity we have that $\mathbf{E}_g^{-1} \mathbf{D}_g \mathbf{E}_g^{-1} = \sigma_g^2 \mathbf{E}_g^{-1}$, where $\sigma_g^2 = E\left[u_g^2\right]$.

(b) *Suppose that $V[\mathbf{y}|\mathbf{z}] = \Omega_0 > 0$, $y = (y_1, \ldots, y_G)'$. How can Ω_0 be estimated consistently?*

Using the single equation residuals in $\hat{\Omega}_n = E_n\left[\tilde{\mathbf{u}}_n\ \tilde{\mathbf{u}}_n'\right]$, $\tilde{u}_{gn} = y_g - m_g\left(\mathbf{z}_g, \tilde{\theta}_{ng}\right)$.

(c) *Let $\hat{\theta}_n$ be the nonlinear SUR estimate that solves the problem*

$$\min_\theta \mathbb{E}_n\left[(\mathbf{y} - \mathbf{m}(\mathbf{z}, \theta))' \hat{\mathbf{\Omega}}_n^{-1} (\mathbf{y} - \mathbf{m}(\mathbf{z}, \theta))\right]$$

where $\hat{\Omega}_n$ is a consistent estimate of Ω_0 and $m(\mathbf{z}, \theta)$ is the $G \times 1$ vector of conditional expectations $m_g(\mathbf{x}_g, \theta_{0g})$, $g = 1, \ldots, G$. Find the asymptotic distribution of $\hat{\theta}_n$.

Using the same arguments as in single equation NLS, we find that the first order conditions are

$$\frac{\partial}{\partial \boldsymbol{\theta}} Q_n(\hat{\theta}_n) = \mathbb{E}_n\left[\nabla_\theta \mathbf{m}'\left(\mathbf{z}, \hat{\theta}_n\right) \hat{\mathbf{\Omega}}_n^{-1}\left(\mathbf{y} - \mathbf{m}\left(\mathbf{z}, \hat{\theta}_n\right)\right)\right] = 0$$

where $\nabla_\theta \mathbf{m}'(\mathbf{z}, \theta)$ means $\nabla_\theta\left\{\mathbf{m}(\mathbf{z}, \theta)'\right\}$

$$\begin{aligned}
\sqrt{n} \frac{\partial}{\partial \boldsymbol{\theta}} Q_n(\boldsymbol{\theta}_0) &= -\sqrt{n} 2 \mathbb{E}_n\left[\nabla_\theta \mathbf{m}'(\mathbf{z}, \theta_0) \hat{\mathbf{\Omega}}_n^{-1} \mathbf{u}\right] \\
&\rightarrow_p -\sqrt{n} 2 \mathbb{E}_n\left[\nabla_\theta \mathbf{m}'(\mathbf{z}, \theta_0) \mathbf{\Omega}_0^{-1} \mathbf{u}\right] \\
&\rightarrow_d N\left(0, 4\bar{\mathbf{E}}\right),
\end{aligned}$$

where $\bar{\mathbf{E}} = E\left[\nabla_\theta \mathbf{m}'(\mathbf{z},\theta_0)\,\boldsymbol{\Omega}_0^{-1}\nabla_\theta \mathbf{m}'(\mathbf{z},\theta_0)'\right]$ and $\mathbf{u} = \mathbf{y} - \mathbf{m}(\mathbf{z},\theta_0)$. Similarly, we have that the second order derivative of the objective function satisfies that

$$\begin{aligned}\frac{\partial^2}{\partial\boldsymbol{\theta}\partial\boldsymbol{\theta}'}Q_n(\theta_0) &= 2\mathbb{E}_n\left[\nabla_\theta \mathbf{m}'(\mathbf{z},\theta_0)\,\hat{\boldsymbol{\Omega}}_n^{-1}\nabla_\theta \mathbf{m}'(\mathbf{z},\theta_0)'\right] \\ &\quad -2\mathbb{E}_n\left[\nabla_{\theta'}\left\{\nabla_\theta \mathbf{m}'(\mathbf{z},\theta_0)\right\}\hat{\boldsymbol{\Omega}}_n^{-1}\left(\mathbf{y} - \mathbf{m}(\mathbf{z},\theta_0)\right)\right] \\ &\to_p 2\bar{\mathbf{E}} + \mathbf{0}\end{aligned}$$

($\nabla_{\theta'}\{\}$ indicates the terms involving second derivatives of $\mathbf{m}$) so

$$n^{1/2}\left(\hat{\theta}_n - \theta_0\right) \to_d N\left(0, \bar{\mathbf{E}}^{-1}\right).$$

(d) *Describe a numerical method to obtain the estimate $\hat{\theta}_n$ and how you would estimate $AVar\left[\hat{\theta}_n\right]$?*

We could use Newton-Raphson, but preferably Gauss Newton, dropping the second term in the second derivative of the objective function which has zero expectation, so the s-iteration is

$$\hat{\theta}_{n(s)} = \hat{\theta}_{n(s-1)} - \hat{\bar{\mathbf{E}}}_{(s)}^{-1}\frac{\partial}{\partial\boldsymbol{\theta}}Q_n(\hat{\theta}_{n(s-1)}, \hat{\boldsymbol{\Omega}}_{n(s-1)}^{-1})$$

where for example

$$\hat{\bar{\mathbf{E}}}_{(s)} = E_n\left[\nabla_\theta \mathbf{m}'\left(\mathbf{z},\hat{\theta}_{n(s-1)}\right)\hat{\boldsymbol{\Omega}}_{n(s-1)}^{-1}\nabla_\theta \mathbf{m}'\left(\mathbf{z},\hat{\theta}_{n(s-1)}\right)'\right]$$

where in the first iteration $\hat{\theta}_{n(s-1)}$ and $\hat{\boldsymbol{\Omega}}_{n(s-1)}$ are based on the single equation estimates $\tilde{\theta}_{n,g}$, $g = 1,\ldots,G$. The last value of $\hat{\bar{\mathbf{E}}}_{(s)}^{-1}$ can be used to approximate the AVar.

(e) *If Ω_0 is diagonal and if the assumptions stated previously hold, which estimation method would be preferred?*

Noting that, if there are no common parameters among the different equations (SUR), $\nabla_\theta \mathbf{m}'\left(\mathbf{z},\hat{\theta}_n\right)$ is block-diagonal and $\boldsymbol{\Omega}_0^{-1} = diag\left\{\sigma_1^2,\ldots,\sigma_G^2\right\}$, we can obtain that

$$\bar{\mathbf{E}} = diag\left\{E_1,\ldots,E_G\right\}$$

so single equation NLS is asymptotically as efficient as system estimation (but not numerically).

3. *Consider the econometric model which relates y_t to the expected value of x_t, x_t^*, where the expectation is based on all observed information at time $t-1$,*

$$y_t = \alpha_0 + \alpha_1 x_t^* + u_t.$$

A natural assumption on $\{u_t\}$ is that $E[u_t|I_{t-1}] = 0$, where I_{t-1} denotes all information on $\{y_t\}$ and $\{x_t\}$ observed at time $t-1$.

To complete the model, we assume that x_t^* is formed according to

$$x_t^* - x_{t-1}^* = \lambda\left(x_{t-1} - x_{t-1}^*\right),$$

where $\lambda \in (0,1)$. This equation implies that the change in expectations reacts to whether last period's realized value was above or below its expectation. The assumption $0 < \lambda < 1$ implies that the change in expectations is a fraction of last period's error.

(a) *Show that the two equations imply that*

$$y_t = \delta_0 + \delta_1 y_{t-1} + \delta_2 x_{t-1} + v_t \tag{1}$$

and give appropriate definitions of $\delta = (\delta_0, \delta_1, \delta_2)'$ *and of* v_t *in terms of* $\{u_t\}$. *What are the dynamic properties of the errors* v_t*?*

We have that $x_t^* - x_{t-1}^* = \lambda \left(x_{t-1} - x_{t-1}^*\right)$ implies

$$\begin{aligned} x_t^* &= x_{t-1}^* + \lambda \left(x_{t-1} - x_{t-1}^*\right) \\ &= \lambda x_{t-1} + (1-\lambda) x_{t-1}^* \\ &= \frac{\lambda}{1-(1-\lambda)L} x_{t-1}, \end{aligned}$$

because $|-(1-\lambda)| < 1$, so

$$\begin{aligned} y_t &= \alpha_0 + \alpha_1 x_t^* + u_t \\ &= \alpha_0 + \frac{\alpha_1 \lambda}{1-(1-\lambda)L} x_{t-1} + u_t \\ &= \alpha_0 \{1-(1-\lambda)\} + (1-\lambda) y_{t-1} + \alpha_1 \lambda x_{t-1} + (1-(1-\lambda)L) u_t \\ &= \underbrace{\alpha_0 \lambda}_{\delta_0} + \underbrace{(1-\lambda)}_{\delta_1} y_{t-1} + \underbrace{\alpha_1 \lambda}_{\delta_2} x_{t-1} + \underbrace{(1-(1-\lambda)L) u_t}_{v_t} \end{aligned}$$

where

$$v_t = (1-(1-\lambda)L) u_t = u_t - (1-\lambda) u_{t-1} = u_t - \delta_1 u_{t-1}$$

is a MA(1), invertible because $|-\delta_1| = |-(1-\lambda)| < 1$.

(b) *How would you estimate the vector* δ *consistently? Study the asymptotic properties of your estimate.*

OLS estimates are inconsistent because of the correlation between v_t and y_{t-1}, which is a function of u_{t-1} as v_t. A possibly solution is to use IV, with instrument x_{t-2} for y_{t-1} (adding possibly more lags in $\mathbf{x}_t = (x_{t-2}, \ldots, x_{t-q})'$ to improve efficiency using 2SLS). In this case we have that

$$\hat{\delta}_T = \left(\sum_{t=1}^T \mathbf{z}_t \mathbf{w}_t'\right)^{-1} \sum_{t=1}^T \mathbf{z}_t y_t$$

where $\mathbf{z}_t = (1, x_{t-1}, x_{t-2})'$, $\mathbf{w}_t = (1, y_{t-1}, x_{t-1})$. We have that, using that $E[\mathbf{z}_t v_t] = E[E[\mathbf{z}_t v_t | I_{t-1}]] = E[\mathbf{z}_t E[v_t | I_{t-1}]] = 0$, since $\mathbf{z}_t \in I_{t-1}$ depends only on past observations of x_t,

$$\begin{aligned} T^{1/2}\left(\hat{\delta}_T - \delta\right) &= \left(T^{-1} \sum_{t=1}^T \mathbf{z}_t \mathbf{w}_t'\right)^{-1} T^{-1/2} \sum_{t=1}^T \mathbf{z}_t v_t \\ &\to_d N\left(0, E[\mathbf{z}_t \mathbf{w}_t']^{-1} \mathbf{S} E[\mathbf{w}_t \mathbf{z}_t']^{-1}\right), \end{aligned}$$

assuming (weak) stationarity and $E[\mathbf{z}_t \mathbf{w}_t']$ full rank, where

$$\mathbf{S} = \sum_{j=-\infty}^{\infty} \Gamma(j), \quad \Gamma(j) = E\left[v_t v_{t+j} \mathbf{z}_t \mathbf{z}_{t+j}'\right].$$

Note that if we further assume some sort of dynamic conditional homoskedasticity we have that

$$\begin{aligned} E\left[v_t v_{t+j} \mathbf{z}_t \mathbf{z}_{t+j}'\right] &= E\left[E\left[v_t v_{t+j} \mathbf{z}_t \mathbf{z}_{t+j}' | I_{t-1}\right]\right] = E\left[\mathbf{z}_t \mathbf{z}_{t+j}' E\left[v_t v_{t+j} | I_{t-1}\right]\right] = E\left[\mathbf{z}_t \mathbf{z}_{t+j}'\right] E\left[v_t v_{t+j}\right] \\ &= E\left[\mathbf{z}_t \mathbf{z}_t'\right] E\left[v_t^2\right] = E\left[\mathbf{z}_t \mathbf{z}_t'\right] \sigma_v^2, \quad j = 0 \\ &= E\left[\mathbf{z}_t \mathbf{z}_{t+1}'\right] E\left[v_t v_{t+1}\right] = E\left[\mathbf{z}_t \mathbf{z}_{t+1}'\right] \gamma_v(1), \quad j = 1 \\ &= 0, \quad |j| > 1, \end{aligned}$$

so

$$\mathbf{S} = \left\{E\left[\mathbf{z}_t \mathbf{z}_{t+1}'\right] + E\left[\mathbf{z}_{t+1} \mathbf{z}_t'\right]\right\} \gamma_v(1) + E\left[\mathbf{z}_t \mathbf{z}_t'\right] \sigma_v^2,$$

and estimates of $\mathbf{S}$ could use this structure. Otherwise we have to use general estimates with a bandwidth $m \to \infty$.

Then, we can think on alternative, more efficient solutions, such as GLS, transforming the noise into an uncorrelated sequence and simultaneously orthogonal with the regressors. In this case the solution is not so neat as with AR(p) disturbances, because Ω is not band diagonal and we require additionally (at all lags) strong exogeneity of the $x_t's$. Note also that Ω only depends on δ_1 (or λ) up to scale.

(c) *Given a consistent estimate of δ, how would you estimate α_1 and λ?*

Just use the definition of δ in terms of $(\alpha_0, \alpha_1, \lambda)$, and find the solution. Then use the the delta method to obtain the asymptotic distribution.

(d) *If you estimate (1) by OLS, how would you check the correct specification of the model, i.e. how can you test the assumptions required for consistency of the OLS estimates.*

Apart from usual checkings using alternative estimates, as Hausman tests (since we can not check directly the orthogonality condition $E\left[y_{t-1} v_t\right] = 0$ with OLS estimation), and given the above structure, with a lagged dependent variable on the rhs of the model, a key condition for OLS consistency is that the errors should be uncorrelated. This can be checked using OLS residuals, either by means of LM or Portmanteau statistics (using an exogeneity assumption, or correcting for possible endogeneity, which is not ruled out by the conditions given). (Using IV estimation we can also directly check that $\delta_1 = 0$ (or $\lambda = 1$ using part (c))).

4. *Consider the linear regression model*

$$y_t = \mathbf{z}_t' \beta + v_t$$

where $E\left[v_t | I_t\right] = 0$, $E\left[v_t^2 | I_t\right] := \sigma_t^2 = \sigma^2 + \delta v_{t-1}^2$, where I_t denotes the information set of current and past observations of z_t and past observations of v_t. Assume that v_t and z_t are stationary.

(a) *Are the v_t independent? Uncorrelated? And/or form a martingale difference?*

Not independent (at least second order conditional moments are not equal to marginal ones) and is a MD sequence because $E\left[v_t | v_{t-1}, \ldots\right] = E\left[E\left[v_t | I_t\right] | v_{t-1}, \ldots\right] = 0$, since $\{v_{t-1}, \ldots\} \subseteq I_t$.

(b) *Find the asymptotic properties of the OLS estimate of β and propose an estimate of its asymptotic covariance matrix if consistent.*

We have that

$$T^{1/2}\left(\beta_n - \beta\right) \to_d N\left(0, E\left[\mathbf{z}_t \mathbf{z}_t'\right]^{-1} E\left[v_t^2 \mathbf{z}_t \mathbf{z}_t'\right] E\left[\mathbf{z}_t \mathbf{z}_t'\right]^{-1}\right)$$

assuming that

$$\begin{aligned}
\frac{1}{T}\sum_{t=1}^{T}\mathbf{z}_t\mathbf{z}_t' &\to_p \quad E\left[\mathbf{z}_t\mathbf{z}_t'\right]>0\\
\frac{1}{T^{1/2}}\sum_{t=1}^{T}\mathbf{z}_t v_t &\to_d \quad N\left(0,E\left[v_t^2\mathbf{z}_t\mathbf{z}_t'\right]\right)
\end{aligned}$$

because $\mathbf{z}_t v_t$ is zero mean and serially uncorrelated.

(c) *Using the additional assumption that the $v_t|I_t$ are normally distributed, study the first order conditions for ML estimation of all the parameters of the model, $\theta=\left(\delta,\sigma^2,\beta\right)'$, and give some approximation for the asymptotic distribution of $\hat{\theta}_T$.*

In this case we have that the v_t are (conditionally) independent, so the likelihood is

$$\begin{aligned}
L_T\left(\delta,\sigma^2,\beta\right) &= \frac{1}{T}\sum_{t=1}^{T}l_t\left(\delta,\sigma^2,\beta\right)\\
&= c-\frac{1}{2T}\sum_{t=1}^{T}\log\sigma_t^2\left(\theta\right)-\frac{1}{2T}\sum_{t=1}^{T}\frac{\left(y_t-\mathbf{z}_t'\beta\right)^2}{\sigma_t^2\left(\theta\right)}
\end{aligned}$$

with $\sigma_t^2\left(\theta\right)=\sigma^2+\delta v_{t-1}^2\left(\beta\right)=\sigma^2+\delta\left(y_{t-1}-\mathbf{z}_{t-1}'\beta\right)^2$, $v_t\left(\beta\right)=y_t-\mathbf{z}_t'\beta$ (and pretending we have initial conditions on the variables required), the score is given by

$$\begin{aligned}
s_T\left(\delta,\beta,\sigma^2\right) &= \frac{\partial L_T}{\partial\theta}\left(\theta\right)\\
&= \frac{1}{T}\sum_{t=1}^{T}\begin{pmatrix}
-\frac{1}{2}\frac{v_{t-1}^2(\beta)}{\sigma_t^2(\theta)}+\frac{1}{2}v_{t-1}^2\left(\beta\right)\frac{\left(y_t-\mathbf{z}_t'\beta\right)^2}{\sigma_t^4(\theta)}\\
\frac{\delta v_{t-1}(\beta)\mathbf{z}_{t-1}}{\sigma_t^2(\theta)}-\delta v_{t-1}\left(\beta\right)\mathbf{z}_{t-1}\frac{\left(y_t-\mathbf{z}_t'\beta\right)^2}{\sigma_t^4(\theta)}+\frac{\left(y_t-\mathbf{z}_t'\beta\right)\mathbf{z}_t}{\sigma_t^2(\theta)}\\
-\frac{1}{2}\frac{1}{\sigma_t^2(\theta)}+\frac{1}{2}\frac{\left(y_t-\mathbf{z}_t'\beta\right)^2}{\sigma_t^4(\theta)}
\end{pmatrix}\\
&\approx \frac{1}{T}\sum_{t=1}^{T}\begin{pmatrix}
\frac{1}{2}\frac{v_{t-1}^2(\beta)}{\sigma_t^2(\theta)}\left(\frac{v_t^2(\beta)}{\sigma_t^2(\theta)}-1\right)\\
\frac{v_t(\beta)\mathbf{z}_t}{\sigma_t^2(\theta)}\left(1-\delta\frac{\sigma_t^2(\theta)}{\sigma_{t+1}^2(\theta)}\left(\frac{v_{t+1}^2(\beta)}{\sigma_{t+1}^2(\theta)}-1\right)\right)\\
\frac{1}{2}\frac{1}{\sigma_t^2(\theta)}\left(\frac{v_t^2(\beta)}{\sigma_t^2(\theta)}-1\right)
\end{pmatrix},
\end{aligned}$$

(the last step ignoring end effects) since

$$\frac{\partial}{\partial\beta}\sigma_t^2\left(\theta\right)=\frac{\partial}{\partial\beta}\delta\left(y_{t-1}-\mathbf{z}_{t-1}'\beta\right)^2=-2\delta\left(y_{t-1}-\mathbf{z}_{t-1}'\beta\right)\mathbf{z}_{t-1}=-2\delta v_{t-1}\left(\beta\right)\mathbf{z}_{t-1}$$

which shows that

$$s_T\left(\theta_0\right)=\frac{1}{T}\sum_{t=1}^{T}s_t(\theta_0)=\frac{1}{T}\sum_{t=1}^{T}\begin{pmatrix}
\frac{1}{2}\frac{v_{t-1}^2}{\sigma_t^2}\left(\frac{v_t^2}{\sigma_t^2}-1\right)\\
\frac{v_t\mathbf{z}_t}{\sigma_t^2}\left(1-\delta\frac{\sigma_t^2}{\sigma_{t+1}^2}\left(\frac{v_{t+1}^2}{\sigma_{t+1}^2}-1\right)\right)\\
\frac{1}{2}\frac{1}{\sigma_t^2}\left(\frac{v_t^2}{\sigma_t^2}-1\right)
\end{pmatrix},$$

so the estimation equation is $E[s_t(\theta_0)]=E[s_t(\theta_0)|I_t]=0$ and under the above conditions on v_t the elements of the sum are a zero mean (conditionally) independent sequence (so a MD even if Gaussianity is dropped), so $T^{1/2}s_T\left(\theta_0\right)$ is asymptotically normal under standard conditions. Then, using the usual Taylor-expansion argument,

$$T^{1/2}\left(\hat{\theta}-\theta_0\right)\to_d N\left(0,I^{-1}\right),$$

where
$$I := -E\left[\frac{\partial^2 L_T}{\partial\theta\partial\theta'}(\theta_0)\right] := -E\left[H_T(\theta_0)\right] = AsyVar\left[T^{1/2}s_T(\theta_0)\right]$$
is the information matrix, and (ignoring end effects), we have for example,
$$\begin{aligned}
H_T\left(\delta,\beta,\sigma^2\right) &= \frac{\partial^2 L_T}{\partial\theta\partial\theta'}(\theta) \\
&\approx \frac{1}{T}\sum_{t=1}^{T}\frac{\partial}{\partial\theta'}\begin{pmatrix} \frac{1}{2}\frac{v_{t-1}^2(\beta)}{\sigma_t^2(\theta)}\left(\frac{v_t^2(\beta)}{\sigma_t^2(\theta)}-1\right) \\ \frac{v_t(\beta)\mathbf{z}_t}{\sigma_t^2(\theta)}\left(1-\delta\frac{\sigma_t^2(\theta)}{\sigma_{t+1}^2(\theta)}\left(\frac{v_{t+1}^2(\beta)}{\sigma_{t+1}^2(\theta)}-1\right)\right) \\ \frac{1}{2}\frac{1}{\sigma_t^2(\theta)}\left(\frac{v_t^2(\beta)}{\sigma_t^2(\theta)}-1\right)\end{pmatrix} \\
&= \frac{1}{T}\sum_{t=1}^{T}\begin{pmatrix} -\frac{1}{2}\frac{v_{t-1}^2(\beta)}{\sigma_t^2(\theta)}\frac{v_{t-1}^2(\beta)}{\sigma_t^2(\theta)}\left(\frac{v_t^2(\beta)}{\sigma_t^2(\theta)}-1\right)-\frac{1}{2}\frac{v_t^2(\beta)}{\sigma_t^2(\theta)}\left(\frac{v_{t-1}^2(\beta)}{\sigma_t^2(\theta)}\right)^2 & \cdots & \\ \vdots & \ddots & \\ & & \ddots \end{pmatrix}.
\end{aligned}$$
Note that the conditional expectation of the $(1,1)$ term is zero, so we have that
$$I_{11} = \frac{1}{2T}\sum_{t=1}^{T}E\left[\frac{v_t^2(\beta)}{\sigma_t^2(\theta)}\left(\frac{v_{t-1}^2(\beta)}{\sigma_t^2(\theta)}\right)^2\right] = \frac{1}{2T}\sum_{t=1}^{T}E\left[\left(\frac{v_{t-1}^2(\beta)}{\sigma_t^2(\theta)}\right)^2\right]$$
because $E\left[v_t^2(\beta)/\sigma_t^2(\theta)\,|I_t\right] = 1$.

(d) *Propose an LM test for the lack of ARCH(1) effects in v_t and describe its asymptotic properties.*

The restricted estimates when $\delta = 0$ are given by $s_T\left(0,\tilde\beta_T,\tilde\sigma_T^2\right) = 0$, and are equivalent to the OLS coefficient and the residual variance,
$$\begin{aligned}
\tilde\beta_T &= \left(\sum_{t=1}^{T}\mathbf{z}_t\mathbf{z}_t'\right)^{-1}\sum_{t=1}^{T}\mathbf{z}_t y_t \\
\tilde\sigma_T^2 &= \frac{1}{T}\sum_{t=1}^{T}\hat v_t^2, \quad \hat v_t = v_t\left(\tilde\beta_T\right) = y_t - \mathbf{z}_t'\tilde\beta_T,
\end{aligned}$$
so the restricted score is equal to
$$s_T\left(0,\tilde\beta_T,\tilde\sigma_T^2\right) = \frac{1}{T}\sum_{t=1}^{T}\begin{pmatrix}\frac{1}{2}\frac{v_{t-1}^2(\tilde\beta_T)}{\tilde\sigma_T^2}\left\{\frac{v_t^2(\tilde\beta_T)}{\tilde\sigma_T^2}-1\right\} \\ 0 \\ 0\end{pmatrix}.$$
Note that under the null $\sigma_t^2 = \sigma^2$ and
$$s_T(\theta_0) = \frac{1}{T}\sum_{t=1}^{T}\begin{pmatrix}\frac{1}{2}\frac{v_{t-1}^2}{\sigma^2}\left(\frac{v_t^2}{\sigma^2}-1\right) \\ \frac{v_t\mathbf{z}_t}{\sigma^2} \\ \frac{1}{2}\frac{1}{\sigma^2}\left(\frac{v_t^2}{\sigma^2}-1\right)\end{pmatrix}$$
which is asymptotically normal with Asymptotic Var-Cov matrix I_0, which can be estimated by $-H_T\left(0,\tilde\beta_T,\tilde\sigma_T^2\right)$ (which can be showed to be diagonal under the null). In this case we can use
$$-\hat H_{T,11}\left(0,\tilde\beta_T,\tilde\sigma_T^2\right) = \frac{1}{2T}\sum_{t=1}^{T}\left(\frac{v_{t-1}^2\left(\tilde\beta_T\right)}{\tilde\sigma_T^2}\right)^2$$

Therefore the LM statistic is

$$\begin{aligned}
LM_T &= T^{-1} s_T\left(0,\tilde{\beta}_T,\tilde{\sigma}_T^2\right)'\left(-\hat{H}_T\left(0,\tilde{\beta}_T,\tilde{\sigma}_T^2\right)\right)^{-1} s_T\left(0,\tilde{\beta}_T,\tilde{\sigma}_T^2\right) \\
&= T^{-1} s_T^{(1)}\left(0,\tilde{\beta}_T,\tilde{\sigma}_T^2\right)^2\left(-\hat{H}_{T,11}\left(0,\tilde{\beta}_T,\tilde{\sigma}_T^2\right)\right)^{-1} \\
&\to_p \frac{T}{2}\left(\frac{1}{T}\sum_{t=1}^{T}\left\{v_t^2\left(\tilde{\beta}_T\right)-\tilde{\sigma}_T^2\right\}\left\{v_{t-1}^2\left(\tilde{\beta}_T\right)-\tilde{\sigma}_T^2\right\}\right)^2\left(\frac{1}{T}\sum_{t=1}^{T}v_{t-1}^2\left(\tilde{\beta}_T\right)^2\right)^{-1} \\
&\to_d {}^{H_0}\chi_1^2,
\end{aligned}$$

which checks the first order auto-correlation between the OLS-squared-residuals $v_t^2\left(\tilde{\beta}_T\right)$. Using the further approximation,

$$\frac{1}{T}\sum_{t=1}^{T}\left\{\frac{v_t^2\left(\tilde{\beta}_T\right)}{\tilde{\sigma}_T^2}-1\right\}^2 \to_p 2 = E\left[Z^4\right]-E\left[Z^2\right]^2, \quad Z\sim N(0,1)$$

we have that

$$\begin{aligned}
LM_T &\to_p T\left(\frac{1}{T}\sum_{t=1}^{T}\left\{v_t^2\left(\tilde{\beta}_T\right)-\tilde{\sigma}_T^2\right\}\left\{v_{t-1}^2\left(\tilde{\beta}_T\right)-\tilde{\sigma}_T^2\right\}\right)^2\left(\frac{1}{T}\sum_{t=1}^{T}v_{t-1}^2\left(\tilde{\beta}_T\right)^2\right)^{-1}\left(\frac{1}{T}\sum_{t=1}^{T}v_t^2\left(\tilde{\beta}_T\right)^2\right)^{-1} \\
&= T\hat{\rho}_{T,v_t^2(\tilde{\beta}_T)}(1)^2
\end{aligned}$$

which can also be interpreted as usual as a TR^2 statistic.

(e) *Dropping the Gaussianity assumption, consider the moment conditions*

$$\begin{aligned}
\mathbb{E}\left[\left(y_t-\boldsymbol{\beta}'\mathbf{z}_t\right)\mathbf{z}_t\right] &= \mathbf{0} \\
\mathbb{E}\left[\left(v_t^2-\sigma_t^2\right)\mathbf{x}_t(\boldsymbol{\beta})\right] &= \mathbf{0}
\end{aligned}$$

where $x_t(\beta)=\left[1,\left(y_{t-1}-\boldsymbol{\beta}'\mathbf{z}_{t-1}\right)^2\right]'$, *to define GMM estimates.*

Compare GMM estimation (and GMM LM test for the same hypothesis) with ML estimation.

The basic difference is that the first moment condition (which identifies $\boldsymbol{\beta}$) does not use the information on the conditional variance (to set up a GLS-type estimate for $\boldsymbol{\beta}$ as ML estimation does) as is equivalent to OLS estimation. (Note also that we are in a just identified case, so weighting is not relevant.)

Therefore, under the null of no ARCH effects both estimates are the same, OLS for β (no GLS correction) and the LM test is the same.

(f) *Set up an iterative procedure to approximate the value of the GMM estimates from an initial point.*

We can use the linearized GMM estimate,

$$\hat{\boldsymbol{\theta}}_{GMM}^{iter} = \tilde{\boldsymbol{\theta}}_n - \left\{\bar{\boldsymbol{\xi}}_{\boldsymbol{\theta},T}(\tilde{\boldsymbol{\theta}}_T)'\bar{\boldsymbol{\xi}}_{\boldsymbol{\theta},T}(\tilde{\boldsymbol{\theta}}_T)\right\}^{-1}\bar{\boldsymbol{\xi}}_{\boldsymbol{\theta},T}(\tilde{\boldsymbol{\theta}}_n)'\bar{\boldsymbol{\xi}}_T(\tilde{\boldsymbol{\theta}}_T),$$

where $\tilde{\boldsymbol{\theta}}_T$ is an initial estimate, where

$$\bar{\xi}_T(\theta) = \frac{1}{T}\sum_{t=1}^{T}\begin{pmatrix}\left(y_t-\boldsymbol{\beta}'\mathbf{z}_t\right)\mathbf{z}_t \\ \left(v_t^2-\sigma_t^2\right)\mathbf{x}_t(\boldsymbol{\beta})\end{pmatrix} = \frac{1}{T}\sum_{t=1}^{T}\begin{pmatrix} v_t(\boldsymbol{\beta})\mathbf{z}_t \\ v_t^2(\boldsymbol{\beta})-\sigma_t^2(\boldsymbol{\theta}) \\ \left(v_t^2(\boldsymbol{\beta})-\sigma_t^2(\boldsymbol{\theta})\right)v_{t-1}^2(\boldsymbol{\beta})\end{pmatrix},$$

and, $\theta = \left(\beta', \sigma^2, \delta\right)'$,

$$\begin{aligned}
\bar{\boldsymbol{\xi}}_{\boldsymbol{\theta},n}(\theta) &= \frac{\partial}{\partial \theta'} \bar{\xi}_T\left(\theta\right) \\
&= \frac{1}{T}\sum_{t=1}^{T} \begin{pmatrix} -\mathbf{z}_t\mathbf{z}_t' & -2v_t\left(\boldsymbol{\beta}\right)\mathbf{z}_t - 2\delta v_{t-1}\left(\beta\right)\mathbf{z}_{t-1} & \begin{array}{c} -2\left(v_t\left(\boldsymbol{\beta}\right)\mathbf{z}_t + \delta v_{t-1}\left(\beta\right)\mathbf{z}_{t-1}\right)v_{t-1}^2\left(\boldsymbol{\beta}\right) \\ -\left(v_t^2\left(\boldsymbol{\beta}\right) - \sigma_t^2\left(\boldsymbol{\theta}\right)\right)2v_{t-1}\left(\boldsymbol{\beta}\right)\mathbf{z}_{t-1} \end{array} \\ 0 & -1 & -v_{t-1}^2\left(\boldsymbol{\beta}\right) \\ 0 & -v_{t-1}^2\left(\boldsymbol{\beta}\right) & -v_{t-1}^4\left(\boldsymbol{\beta}\right) \end{pmatrix},
\end{aligned}$$

(from where you can show that, under appropriate conditions, $\Xi_0 = p\lim \bar{\boldsymbol{\xi}}_{\boldsymbol{\theta},n}(\theta_0)$ is block diagonal, so asymptotic properties are very easy to establish.)