CAPÍTULO 9

9.1 Estimador de Variable Indep. en RS

- En un modelo $Y_i = \beta_0 + \beta_1 X_i + \varepsilon_i$, si X_i está correlacionada con ε_i el estimador $\hat{\beta}_1$ estará sesgado. Decimos que se da la condición de endogeneidad cuando $cov(X_i, \varepsilon_i) \neq 0$
- Una solución es el uso de variable instrumentales. Queremos encontrar una variable Z tal que:

  * $cov(Z, \varepsilon) = 0$ ← condición de exogeneidad
  * $cov(Z, X) \neq 0$ ← condición de relevancia

Se crea un estimador MC2E, de mínimos cuadrados en dos etapas:

1. $X_i = \pi_0 + \pi_1 Z_i + u_i \rightsquigarrow \hat{X}_i$
2. $Y_i = \beta_0 + \beta_1 \hat{X}_i + \varepsilon_i \rightsquigarrow \hat{Y}_i, \hat{\beta}_1, \hat{\beta}_0$

- Entonces $\hat{\beta}_1^{MC2E} \xrightarrow{P} \beta_1$.

9.2 Modelo general de regresión con VI

Modelo:

$$Y_{0i} = \underbrace{\beta_0 + \beta_1 X_{1i} + \dots + \beta_r X_{ri}}_{\text{variables exógenas}} + \underbrace{\beta_{(r+1)} Y_{(1)i} + \dots + \beta_{(r+k)} Y_{(k)i}}_{\text{variables endógenas}} + \varepsilon_i$$

$$Y_{(j)i} = \pi_{0j} + \pi_{1j} X_{1i} + \dots + \pi_{rj} X_{ri} + \pi_{(r+1)j} Z_{(1)i} + \dots + u_{ji}$$

* Condición de orden: al menos k variables instrumental correlacionadas con $Y_{(t)i}$
* Condición de rango: La matriz

$$\begin{pmatrix} \pi_{(r+1)1} & \cdots & \pi_{(r+m)1} \\ \vdots & & \\ \pi_{(r+1)k} & \cdots & \pi_{(r+m)k} \end{pmatrix}$$

debe tener rango $= k$.

$Y = \Pi_1 X + \Pi_2 Z + u$

Tenemos dos condiciones sobre los instrumentos:

* Relevancia del instrumento: Los vectores $(1, X_{1i}, \dots, X_{ri}, \hat{Y}^*_{1i}, \dots, \hat{Y}^*_{ki})$ no pueden ser colineales.
* Exogeneidad del inst: $corr(Z_{ji}, \varepsilon_i) = 0 \ \forall j$.

9.3 La regresión VI para resolver la endogenidad

- Problemas: * Omisión de una variable explicativa da lugar a endogenidad. Si encontramos un instrumento para la variable omitida podemos solucionarlo.
* Problema de medición, i.e. medimos $X_i^* = X_i + w_i$. Da lugar a $\mathbb{E}[X_{1i}\varepsilon_i] \neq 0$. Solución: encontrar Z_i correlacionado con X_i^* pero no con w_i. Entonces $\mathbb{E}[Z_i \varepsilon_i] = 0$
* Simultaneidad (Y afecta a X también).
- Test de endogenidad: contraste de especificación de Hausman. Contrastamos los estimadores MCO y MC2E. Si son muy diferentes concluimos que la variable es endógena.

9.4 Validez de los instrumentos

- $\hat{\beta}_1^{MC2E} = \beta_1 + \dfrac{\rho_{Z\varepsilon} \cdot \sigma_\varepsilon}{\rho_{ZX} \ \sigma_X}$ por lo que necesitamos que $cov(Z, X)$ sea grande para que el resultado sea fiable. Llamamos a esto la relevancia del instrumento. La segunda condición es la exogeneidad del instrumento, i.e. $cov(Z, \varepsilon) = 0$. Si no, no hay convergencia $\hat{\beta}_1 \to \beta_1$.

9.5 Expresión matricial

- $Y = X\beta + \varepsilon$, estimamos $\hat{X} = P_Z X = Z(Z^T Z)^{-1} Z^T X$. Después miramos $Y = \hat{X}\beta + u$. Por lo tanto

$$\hat{\beta}^{MC2E} = (\hat{X}^T \hat{X})^{-1} \hat{X}^T Y$$
$$= \beta + (X^T P_Z X)^{-1} X^T P_Z \varepsilon$$

$$\Rightarrow \sqrt{n}(\hat{\beta}^{MC2E} - \beta) \xrightarrow{d} N(0, V^{MC2E})$$

complicado.

CAPÍTULO 10

→ Datos fusionados : Y_t , t aleatorio
→ Datos de panel : Y_{it}

10.1 Datos fusionados de sección cruzada

- Realizamos encuestas de tamaño N en T diferentes momentos (= NT puntos). Esto dificulta la suposición idd.

$$y_{it} = \alpha + x_{it}^T \beta + \varepsilon_{it} \quad , \quad x_{it} \in \mathbb{R}^{k \times 1}$$

10.2 Datos de panel

- Si intentamos controlar por efectos no observados que son constantes en el tiempo pero varían entre unidades podemos recurrir al método de efectos fijos :

$$Y_{it} = \beta_1 X_{it} + \alpha_i + \varepsilon_{it}$$

3 métodos para efectos fijos.

* Se puede estimar utilizando mco + variables binarias , ie. $\alpha_1 \mathbb{1}\{i=1\} + \cdots + \alpha_n \mathbb{1}\{i=n\}$.

* Otra manera de hacerlo es $\overline{Y}_i = \sum_{t=1}^{T} Y_{it}$, y similar para $\overline{X}_i$, $\overline{\varepsilon}_i$ tal que $\overline{Y}_i = \beta_1 \overline{X}_i + \alpha_i + \overline{\varepsilon}_i$. Entonces,

$Y_{it} - \overline{Y}_i = \beta_1 (X_{it} - \overline{X}_i) + (\varepsilon_{it} - \overline{\varepsilon}_i)$. Estimamos es mco , lo cual nos conduce al mismo estimador. Después, $\hat{\alpha}_i = \overline{Y}_i - \hat{\beta}_1 \overline{X}_i$ se llama estimador intragrupos o "within".

* Método de diferencias :
$\Delta Y_i = Y_{it} - Y_{i0} = \beta_1 \Delta X_i + \Delta \varepsilon_i \rightsquigarrow$ encontrar $\hat{\beta}_1$

- ¿Qué pasa con el modelo $Y_{it} = \beta_1 X_{it} + \alpha_i + \mu_t + \varepsilon_{it}$? Podemos usar variables binarias $\mathbb{1}\{i=\cdot\}, \mathbb{1}\{t=\cdot\}$ otra vez. También podemos hacer

$$Y_{it} - \overline{Y}_{i.} - \overline{Y}_{.t} + \overline{Y}_{..} = \beta_1 (X_{it_1} - \overline{X}_{i.1} - \overline{X}_{.t1} + \overline{X}_{..}) + \beta_2 (\cdots) + \cdots + (\varepsilon_{it} - \overline{\varepsilon}_{i.} - \overline{\varepsilon}_{.t} + \overline{\varepsilon}_{..})$$

donde $\overline{Y}_{..} = \frac{1}{(nT)} \sum_i \sum_t y_{it}$.

- Datos de panel con efecto aleatorio

Modelo $Y_{it} = X_{it}^T \beta + \alpha_i + \varepsilon_{it}$ pero ahora $\alpha_i \sim [\alpha, \sigma_\alpha^2]$, ie. no es constante. Asumimos que $\mathbb{E}[\alpha_i | X_{i1}, \ldots, X_{iT}] = 0$. Let $v_{it} = \alpha_i + \varepsilon_{it}$, entonces $corr(v_{it}, v_{is}) \neq 0$, hay autocorrelación por lo que los errores estándar no serán correctos en la estimación mco.

Chapter 6 →

⚠ Ante la duda utilizar efectos fijos , ya que éstos son también consistentes bajo los supuestos de efectos aleatorios.

Capítulo 11

11.1 Introducción

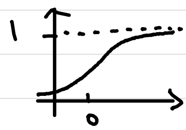
- Estudiamos modelos en los que la variable dependiente es cualitativa

11.2 Modelo lineal de probabilidad

- Si $Y_i \in \{0,1\}$ la regresión $Y_i = \alpha + X_i^T \beta + \varepsilon_i$ da lugar al modelo lineal de prob. Nótese que $\mathbb{P}(Y_i = 1) = \mathbb{E}[Y_i] = \alpha + X_i^T \beta$
- Problemas: no tenemos necesariamente que $0 \leq \alpha + X_i^T \beta \leq 1$, $var(\varepsilon_i)$ depende de X_i por lo que no hay homocedasticidad, ε_i es binomial (no normal) ...

11.3 El modelo Logit

- $f(x) = \dfrac{1}{1+e^{-x}}$ $\rightsquigarrow$



Modelo: $Y_i = \dfrac{1}{1+\exp\{-(\alpha + X_i^T \beta)\}}$ no lineal

- Si $\mathbb{P}(Y_i = 1) \geq 0.5$ entonces $\hat{Y}_i = 1$. La proporción de aciertos nos da una manera de medir la bondad.
- Aquí $\partial Y_i / \partial \beta_j \neq \beta_j$ así que la interpretación de β_j es más complicada que en el antiguo modelo 11.2.

11.4 El modelo Probit

- La función de densidad acumulada de una normal, ϕ, también tiene la forma de S. El modelo es $\mathbb{P}(Y_i = 1) = \phi(\alpha + X_i^T \beta) = \mathbb{P}(z \leq \alpha + X_i^T \beta)$, no lineal.

11.5 Estimación de modelos logit y probit

- Estos dos modelos se estiman con estimadores MV (máxima verosimilitud). Si F es f o ϕ entonces la función de verosimilitud conjunta es

$$\ell = \prod_{i=1}^{n} F(X\beta)^{y_i} [1 - F[X\beta]]^{1-y_i} \qquad y_i \in \{0,1\}$$

Trabajamos con $\log(\ell)$. La condición $\frac{\partial \ln(\ell)}{\partial \beta} = 0$ nos deja estimar β, y la Hessiana $\partial^2 \ln(\ell) / \partial \beta^2$ estima asintóticamente $var(\hat{\beta})$.

Tema 13

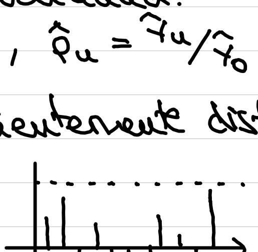
13.1 Procesos estocásticos

- Sea $\{z_t : 1 \leq t \leq T\}$ un proceso estocástico con esperanza marginal $\mathbb{E}[z_t] = \mu_t$, varianza marginal $var(z_t) = \sigma_t^2$. Además tenemos:
  - * Función de autocovarianza $\gamma(t, t+u) = cov(z_t, z_{t+u})$.
  - * Función de autocorrelación $\rho(t, t+u) = \frac{cov(z_t, z_{t+u})}{\sigma_t \, \sigma_{t+u}}$ con valores entre -1 y 1.
- Es estacionario en sentido estricto si las funciones de distⁿ son idénticas, ie. $F(z_t) = F(z_s)$.
- Es estacionario en sentido débil si
  - * $\mu_t = \mu$, $\sigma_t^2 = \sigma^2$, $\gamma(t, t+u) = \gamma_u = \gamma_{-u}$
- **Ej:** Ruido blanco ε_t con $\mathbb{E}[z_t] = 0$, $var(z_t) = \sigma^2$ y $cov(z_t, z_{t+u}) = 0 \quad \forall u \neq 0$. Gaussiano si $z_t \sim N(0, \sigma^2)$.

13.2 Estimación de los momentos en procesos estacionarios

- media muestral: $\bar{z} = \frac{\sum_{t=1}^{T} z_t}{T}$. Nótese que $\mathbb{E}[\bar{z}] = \mu$. Llamamos al proceso ergódico si $var(\bar{z}) \to 0$ cuando $T \to \infty$. Estacionario $\not\Rightarrow$ ergódico, ej: $z_t = z_1 \ \forall t$. Si el proceso es ergódico, $z_t \xrightarrow{P} \mu$
- Nótese que $var(\bar{z}) = \frac{1}{T}[\sigma^2 + 2\sum_{t=1}^{T-1}(1 - i/T)\gamma_i]$. Si existe una dependencia débil en covarianza, ie. $\gamma_u \to 0$ si $u \to \infty$, tendremos ergodicidad.
- $\hat{\gamma}_u = \frac{1}{T}\sum_{t=u+1}^{T}(z_t - \bar{z})(z_{t-u} - \bar{z})$, $\hat{\rho}_u = \hat{\gamma}_u / \hat{\gamma}_0$.

Para determinar si $\hat{\rho}_u$ es suficientemente distinto de 0 dibujamos un correlograma y comparamos con un proceso de ruido blanco ($\mathbb{E}[\hat{\rho}_u] = 0$ y $var(\hat{\rho}_u) = 1/T$). De hecho aquí tendremos $\hat{\rho}_u \sim N(0, 1/T)$ e indep. Podemos pues realizar un contraste de hipótesis nula $H_0 : \rho_u = 0$, $|\hat{\rho}_u - 0 / 1/\sqrt{T}| > 1.96$.



13.3 Procesos integrados

Ej. en 468 muy bueno

- Muchas veces los procesos no son estacionarios en media y/o varianza. Hay dos transformaciones que pueden ayudar:
  - * Diferenciación: $\Delta X_t = X_t - X_{t-1}$, $\Delta^i = \Delta \circ \Delta \cdots \circ \Delta$.
  - * Dif. estacional: $x_t = X_t - X_{t-12}$ (por ejemplo).

Un proceso $I(d, D)$ será integrado de orden d regular y orden D estacional.

- También está la transformación de Box-Cox. Si $\bar{x}_i$ y s_{x_i} es la media y desviación típica hacemos un MCO $\ln s_{x_i} = c + (1-\lambda)\ln \bar{x}_i$. Calculamos λ y hacemos la transformación $y_t = x_t^{\lambda} - 1 / \lambda$.

13.4 Procesos autorregresivos

- Un proceso autorregresivo de orden p, $AR(p)$, es
  (*) $z_t = \phi_0 + \phi_1 z_{t-1} + \cdots + \phi_p z_{t-p} + \varepsilon_t$, donde ε_t es ruido blanco $\Leftrightarrow$
  $(1 - B\phi_1 - \cdots - B^p\phi_p)z_t = \varepsilon_t \Leftrightarrow AR(B)z_t = \varepsilon_t$.

Es fácil ver porqué en AR(1).

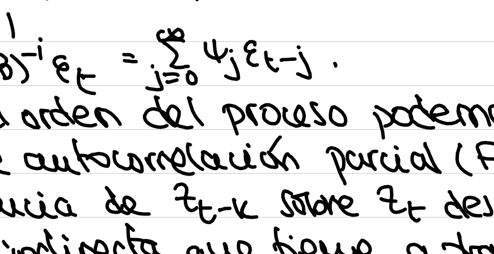
- El proceso es estacionario si todas las raíces de $AR(B)$ tienen $|\cdot| > 1$.
- Nótese que $z_t = AR(B)^{-1}\varepsilon_t = \sum_{j=0}^{\infty}\psi_j \varepsilon_{t-j}$.
- Para determinar el orden del proceso podemos mirar la función de autocorrelación parcial (FAP), que mide la influencia de z_{t-k} sobre z_t descontando la influencia indirecta que tiene a través de $z_{t-1}, \ldots, z_{t-k+1}$. Llamamos al coef. ϕ_{kk}. eg:

$$\left.\begin{array}{ll} z_t = \phi_1 z_{t-1} & \rightsquigarrow \phi_{11} = \phi_1 \\ z_t = \varphi_1 z_{t-1} + \varphi_2 z_{t-2} & \rightsquigarrow \phi_{22} = \varphi_2 \\ \vdots & \quad \vdots \end{array}\right\} \text{Estimación con MCO.}$$

También podemos estimar ϕ_{uu} con las ecuaciones de Yule-Walker: multiplicamos (*) por z_{t-u} y aplicamos $\mathbb{E}[\cdot]$ $\rightsquigarrow \rho_u = \phi_1 \rho_{u-1} + \cdots + \phi_p \rho_{u-p}$

$$\Leftrightarrow \begin{pmatrix} \phi_{11} \\ \vdots \\ \phi_{uu} \end{pmatrix} = \begin{pmatrix} \rho_0 & \rho_1 & \cdots & \rho_{u-1} \\ \rho_1 & \rho_0 & \cdots & \rho_1 \\ \rho_{u-1} & \cdots & \rho_1 & \rho_0 \end{pmatrix}^{-1} \begin{pmatrix} \rho_1 \\ \vdots \\ \rho_u \end{pmatrix}.$$

- La FAP sólo es significativa en procesos AR(p) hasta ϕ_{pp} ej:



Proceso AR(2), $\phi_1, \phi_2 > 0$

13.5 Procesos de medias móviles

- Modelo $z_t = c + \varepsilon_t - \theta_1 \varepsilon_{t-1} - \cdots - \theta_q \varepsilon_{t-q}$
  $= MA(B)\varepsilon_t \Leftrightarrow \varepsilon_t = MA(B)^{-1} z_t$.

Si las raíces son $|\cdot| > 1$, convergerá, por lo que escribimos $z_t = \sum_{j=1}^{\infty}\pi_j z_{t-j} + \varepsilon_t$, ie. es un proceso $AR(\infty)$. La FAP decrece geométricamente y la FAT es $\neq 0$ en los primeros q términos, ie. la FAT y FAP se intercambian roles.

en MA(1) si $|\theta_1| < 1$.

- Un proceso es invertible si se puede escribir como $AR(\infty)$.

13.6 Procesos ARMA

- Modelo $z_t = \sum_{j=1}^{p}\phi_j z_{t-j} - \sum_{i=0}^{q}\theta_i \varepsilon_{t-i}$ $\in ARMA(p, q)$

$\Leftrightarrow AR(B)z_t = MA(B)\varepsilon_t$. Invertible si las raíces de MA son $|\cdot| > 1$ y estacionario si lo mismo con las raíces de AR (si no comparten raíces). Entonces se puede expresar como $AR(\infty)$ o $MA(\infty)$.

- Elegimos el modelo en un rango razonable para p y q, via la maximización de verosimilitud, que minimice la BIC o AIC.

13.7, 13.8 Procesos ARIMA

- w_t es un proceso $ARIMA(p, d, q)$ si $\Delta^d w_t$ es $ARMA(p, q)$, donde $\Delta w_t = w_t - w_{t-1}$.
- Los modelos SARIMA, o ARIMA estacionales, son aquellos con retardos estacionales $s, 2s, 3s \ldots$, ej: $SAR(2)$ es $z_t = \phi_1 z_{t-s} + \phi_2 z_{t-2s} + \varepsilon_t$.

13.9 Identificación y Validación

- Procedimiento para ajustar un modelo:
  (1) Transformarlo en estacionaria
  (2) Identificar el modelo $ARMA(p, q)$
  (3) Estimarlo y (4) Validarlo.
- Para validar estudiamos si $\hat{\varepsilon}_t$ se comporta como ruido blanco. Contrastamos $|\hat{\rho}_u| \leq 1.96/\sqrt{T}$ uno a uno, pero como algunos pueden sobrepasar el umbral por azar, hacemos un contraste global:

$$Q(K) = T(T+2)\sum_{i=1}^{K}\frac{\hat{\rho}_i^2}{T-i} \sim \chi^2_{n-K} \text{ bajo } H_0 \quad \leftarrow \text{estadístico Q de Ljung-Box.}$$

13.10 Predicción

- Si $\underline{z}_T = \{z_1, \ldots, z_T, \varepsilon_1, \ldots, \varepsilon_T, \phi_1, \ldots, \phi_p, \theta_1, \ldots, \theta_q\}$ entonces predecimos $\hat{z}_{t+h} = \mathbb{E}[z_{t+h} \mid \underline{z}_T]$.
- Ejemplo: AR(1), $z_t = c + \phi_1 z_{t-1} + \varepsilon_t$. Entonces, $\hat{z}_{t+1} = \mathbb{E}[c + \phi_1 z_t + \varepsilon_{t+1} \mid \underline{z}_t] = c + \phi_1 z_t$ y el error es $err(\hat{z}_{t+1}) = z_{t+1} - \hat{z}_{t+1} = c + \phi_1 z_t + \varepsilon_{t+1} - c - \phi_1 z_t = \varepsilon_{t+1}$.

En general $\hat{z}_{t+h} = c(1 + \phi_1 + \cdots + \phi_1^{h-1}) + \phi_1^h z_t \to c/1-\phi_1$, ie. as $h \to \infty$ la predicción tiende a la media.

13.11 Modelos autorregresivos con predictores

- Modelo $y_t = \beta_0 + \sum_{i=1}^{p}\beta_i y_{t-i} + \sum_{j=1}^{q}\delta_j x_{t-j} + \varepsilon_t$.
  ↳ modelo autorregresivo de retardos distribuidos, ARD(p,q).

Tema 16

- Tratamos efectos causales dinámicos, ej:
  $Y_t = f(X_t) + \varepsilon_t$ donde $\varepsilon_t = g(\varepsilon_{t-1})$
  o $Y_t = f(X_t, X_{t-1}, \dots) + \varepsilon_t$. modelos de retardos distribuidos (RD).
- Los autoregresivos RD, ie. ARD, son tal que
  $Y_t = f(Y_{t-1}, X_t, X_{t-1}, \dots) + \varepsilon_t$.
- En los modelos RD asumimos exogeneidad, ie. $\mathbb{E}[\varepsilon_t | X_t, X_{t-1}, \dots] = 0$, para estimar los coeficientes MCO consistentemente.
  ↳ $Y_t = \beta_0 + \beta_1 X_t + \dots + \beta_{k+1} X_{t-k} + \varepsilon_t$.

  También asumimos que Y_t, X_t son estacionarios.
- Un modelo dinámico completo incluiría $\underline{X}_t$, ie. un vector de regresores en cada momento.
- Modelo ARD(p,q):

$$y_t = \sum_{i=1}^{p} \gamma_i y_{t-i} + \sum_{i=0}^{q} \beta_i x_{t-i} + \varepsilon_t$$

B = backward operator

$$\Leftrightarrow \quad C(B) y_t = A(B) x_t + \varepsilon_t$$

$$\Leftrightarrow \quad y_t = \frac{A(B)}{C(B)} x_t + n_t$$

Tema 17

17.1 Tendencias

- Existen tendencias deterministas (ej. $y_t = \beta_0 + \beta_1 t + \varepsilon_t$) donde la media cambia por una función. Las tendencias estocásticas tienen una media cambiante pero se debe a sumas de variables aleatorias causadas por, por ejemplo, raíces unitarias:
  - Paseo aleatorio $y_t = y_{t-1} + \varepsilon_t$ (ie. $\phi_1 = 1$)
  $\Rightarrow y_t = y_0 + \sum_{i=0}^{t-1} \varepsilon_i$.
  La media es estacionaria ($\mathbb{E}y_t = \mathbb{E}y_0$) pero la varianza no ($\operatorname{var} y_t = \sum \operatorname{var}\varepsilon_i = t\sigma_\varepsilon^2$).

17.2 Regresiones espurias

- Si y_t, x_t son paseos aleatorios independientes y regresamos $y_t = \alpha + \beta_0 x_t + u_t$. Al hacer MCO, observamos que rechazamos $H_0: \beta_0 = 0$ con prob 1 cuando $t \to \infty$. ¿Por qué? Porque aunque H_1 es mala, H_0 también lo es y $\operatorname{var}(y_t) \to \infty$ pero $\operatorname{var}(\alpha + u_t) = \sigma_u^2 \ \forall t$.
- Por eso si regresemos $\text{consumo}_{ESP} = \beta \, PIB_{ARG} + \alpha$ vemos que β es significativo: no hay relación casual pero al menos hay variación.
- Si $DW < R^2$ entonces hay riesgo de reg. espuria.

17.3 Contraste de raíces unitarias

- Para contrastar un AR(1) $y_t = \rho y_{t-1} + \varepsilon_t$ si queremos ver si hay raíz unitaria ($H_0: \rho = 1$) hacemos un contraste de Dickey-Fuller (DF), ie. $\frac{\hat{\rho} - 1}{ee(\hat{\rho})} \to DF_i$ $\left(\begin{array}{ll} i=0 & \text{si no hay deriva} \\ i=1 & \alpha \\ i=2 & \alpha + \beta t \end{array} \right)$.
- Para $\Delta y_t = \delta y_{t-1} + \sum_{i=1}^{q} \gamma_i \Delta y_{t-i} + \varepsilon_t$ hacemos un contraste aumentado de Dickey-Fuller (ADF).

17.4 Orientaciones en la modelación

- Intenta diferenciar cuando el proceso sea $I(i)$ para conseguir una serie estacionaria.

Tema 18

- Hay procesos (ej. activo bursátil) donde no hay homocedasticidad. Los modelos ARCH (autorregresivo heterocedástico condicionado) son tal que los residuos estimados tengan una estructura AR, i.e.

$$\hat{\varepsilon}_t^2 = \alpha_0 + \alpha_1 \hat{\varepsilon}_{t-1}^2 + \dots + \alpha_q \hat{\varepsilon}_{t-q}^2 + v_t \leftarrow \text{ruido blanco}$$

- Una alternativa son los de perturbación multiplicativa:
$\varepsilon_t = v_t \sqrt{\alpha_0 + \alpha_1 \varepsilon_{t-1}^2}$. Entonces,

$$\mathbb{E}[\varepsilon_t^2 \mid \varepsilon_{t-1}, \dots] = \alpha_0 + \alpha_1 \varepsilon_{t-1}^2, \text{ i.e. ARCH(1)}.$$

- ARCH(q) son $\varepsilon_t = v_t \sqrt{\alpha_0 + \sum_{i=1}^{q} \alpha_i \varepsilon_{t-i}^2}$

- Los modelos GARCH son una generalización, permitiendo que la varianza condicionada sea un proceso ARMA, i.e.: $\varepsilon_t = v_t \sqrt{h_t}$ donde $h_t = \alpha_0 + \sum_{i=1}^{q} \alpha_i \varepsilon_{t-i}^2 + \sum_{i=1}^{p} \beta_i h_{t-i}$.

- Modelización
  (1) Estimar $\{Y_t\}$ usando un ARMA para estimar $\hat{\varepsilon}_t^2$.
  (2) Dibujar autocorrelaciones muestrales de $\hat{\varepsilon}_t^2$ (i.e. $\hat{\rho}_{\hat{\varepsilon}^2}$) para saber si GARCH es apropiado.
  (3) Utilizar MCOs para estimar GARCH.

Tema 19

19.1 Introducción

- Modelo de vectores autorregresivos (VAR):

$$x_t = \alpha_{10} + \alpha_{11} x_{t-1} + \dots + \alpha_{1p} x_{t-p} + \beta_{11} y_{t-1} + \dots + u_{1t}$$
$$y_t = \alpha_{20} + \alpha_{21} x_{t-1} + \dots + \beta_{21} y_{t-1} + \dots + u_{2t}$$

$$\Leftrightarrow \underline{w}_t = \underline{A}_0 + \underline{A}_1 \underline{w}_{t-1} + \dots + \underline{A}_p \underline{w}_{t-p} + \underline{u}_t$$

19.2 Estimación y orden del VAR

- Si asumimos errores homocedásticos y no autocorrelados, los estimadores MCO son consistentes y asintóticamente normales.
- selección entre VAR(p) y VAR(q): estimamos matrices de covarianzas $\hat{\Sigma}_p, \hat{\Sigma}_q$ y entonces $T[\ln(\det \hat{\Sigma}_p) - \ln(\det \hat{\Sigma}_q)] \sim \chi^2_r$. También podemos comparar con AIC.

19.3 Diferentes formas del VAR

- La VAR estructural (SVAR) incluye los valores actuales, ie.

$$\begin{pmatrix} x_t \\ y_t \end{pmatrix} = \begin{pmatrix} 0 & \gamma_{12} \\ \gamma_{21} & 0 \end{pmatrix} \begin{pmatrix} x_t \\ y_t \end{pmatrix} + \begin{pmatrix} \alpha_{11} & \beta_{11} \\ \alpha_{21} & \beta_{21} \end{pmatrix} \begin{pmatrix} x_{t-1} \\ y_{t-1} \end{pmatrix} + \begin{pmatrix} \varepsilon_{1t} \\ \varepsilon_{2t} \end{pmatrix}.$$

Se puede manipular para que sea un VAR, pero después no se puede retornar al modelo. ver pg 707,

19.4 Predicción

$$\mathbb{E}[w_{t+1} | w_t, \dots, w_1] = \hat{w}_{t+1} = A_0 + A_1 w_t \quad \text{y}$$
$$\hat{w}_{t+h} = (I + A_1 + \dots + A_1^{h-1}) A_0 + A_1^h w_t.$$

19.5 Causalidad de Granger

$$x_t = \alpha_{11} x_{t-1} + \dots + \beta_{11} y_{t-1} + \dots + u_{1t}$$
$$y_t = \alpha_{21} x_{t-1} + \dots + \beta_{21} y_{t-1} + \dots + u_{2t}$$

- Si los coef β_{1i} son significativos pero α_{2i} no, diremos que hay causalidad de Granger de y a x, and so on...
- Podemos expresar también los modelos VAR como vectores MA(∞), ie:

$$\begin{pmatrix} x_t \\ y_t \end{pmatrix} = \sum_{i=0}^{\infty} \begin{pmatrix} \varphi_{11}(i) & \varphi_{12}(i) \\ \varphi_{21}(i) & \varphi_{22}(i) \end{pmatrix} \begin{pmatrix} u_{1(t-i)} \\ u_{2(t-i)} \end{pmatrix}.$$ Se llaman funciones ($\varphi(i)$) de respuesta al impulso (FRI).

Tema 20

20.1 Introducción

- En series univariantes $I(1)$ puede darse regresiones espurias, las cuales solucionamos diferenciando para obtener series estacionarias.
- En modelos VAR con la misma problemática podemos utilizar la cointegración. X_t e Y_t están cointegradas si ambas son $I(1)$ pero existe una combinación lineal estacionaria.

20.2 Cointegración

- X_t e Y_t están cointegradas de orden d, b, $d \geq b$ (y escribimos $CI(d,b)$) si ambas son $I(d)$ y existen β_1, β_2 tal que $\beta_1 Y_t + \beta_2 X_t$ es $I(d-b)$.

20.3 Contraste (Engle y Granger)

(1) Utilizar test ADF para contrastar si ambas series son $I(1)$.

(2) Estimamos $Y_t = \alpha + \beta X_t + \varepsilon_t$ con MCO y conseguimos $e_t = \hat{\varepsilon}_t$. Debe ser estacionario al largo plazo si hay cointegración.

(3) Contrastamos la estacionalidad de e_t con un nuevo test ADF, ie. $\Delta e_t = \delta e_{t-1} + \sum \alpha_i \Delta e_{t-i} + \eta_t$ y $H_0: \delta = 0$.