MODELOS CON VARIABLES REZAGADAS

I CONCEPTOS BASICOS

Un modelo con rezagos distribuidos toma la forma general:

$$Y_t = \alpha + \sum_{i=0}^{\infty} \beta_i X_{t-i} + u_t \qquad (1)$$

Bajo dicho modelo, un cambio de un período en X_t, en cualquier punto del tiempo, afectará a $E(Y_s|\mathbf{X})$, en cada período subsiguiente. Por ejemplo, supongamos s=t+1:

$$E(Y_{t+1}| \mathbf{X}) = \alpha + \sum_{i=0}^{\infty} \beta_i X_{t+1-i} = \alpha + \beta_0 X_{t+1} + \beta_1 X_t + \beta_2 X_{t-1} + ... \qquad (2)$$

con lo cual, $\dfrac{\partial E(Y_{t+1} | \mathbf{X})}{\partial X_t} = \beta_1$.

Supongamos un **estado estacionario** (*steady-state*), donde $E(Y_t|\mathbf{X}) = \overline{Y}$ y el nivel de X ha permanecido constante por muchos períodos en $\overline{X}$. Entonces:

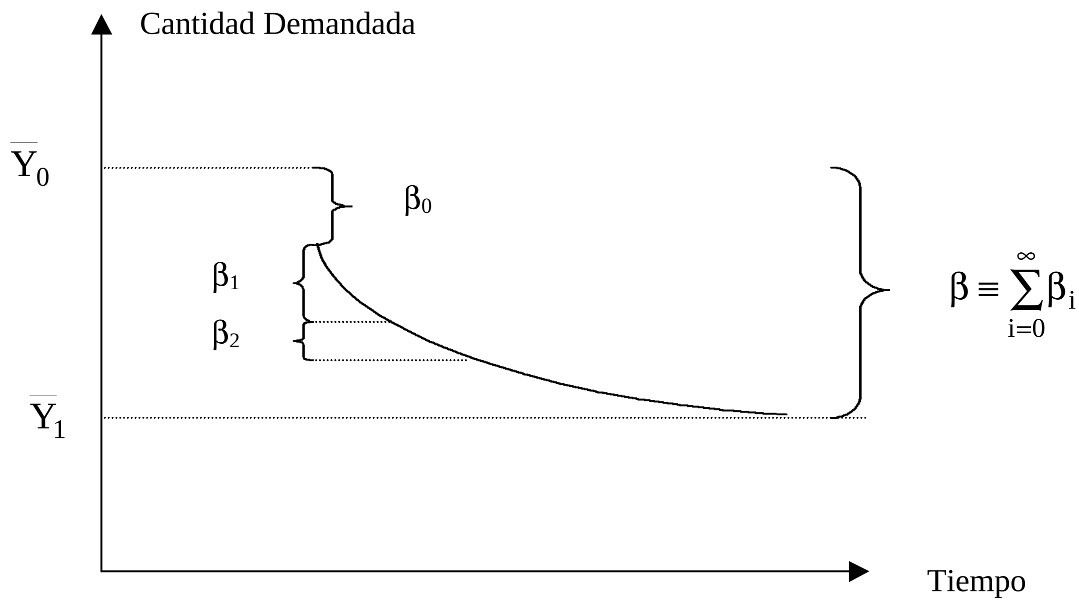
$$\overline{Y} = \alpha + \left(\sum_{i=0}^{\infty} \beta_i \right) \overline{X} \qquad (3)$$

donde $\left| \sum_{i=0}^{\infty} \beta_i \right| < \infty$, a fin de $\overline{Y}$ sea finito.

Multiplicadores: Se define β_0 como el multiplicador de corto plazo y $\beta \equiv \sum_{i=0}^{\infty} \beta_i$ como el multiplicador de equilibrio o de largo plazo.

Ejemplo

Supongamos que X presenta el precio e Y la cantidad demandada de un bien, y que nos encontramos en un estado estacionario $(\overline{X}_0, \overline{Y}_0)$. Al aumentar el precio de $\overline{X}_0$ a $\overline{X}_1$, pasamos a un nuevo estado estacionario, con una menor cantidad demandada $\overline{Y}_1$:



1.1 Modelo No Restringido con un Número de Rezagos Finitos

Un modelo no restringido toma la forma:

$$Y_t = \alpha + \sum_{i=0}^{q} \beta_i X_{t-i} + u_t \tag{4}$$

Si q es conocido y se satisfacen todos los supuestos del modelo lineal clásico, entonces la ecuación (4) es un modelo de regresión clásico. En la práctica, sin embargo, q es desconocido. Una forma de encontrar el q apropiado es a través del R^2 ajustado, $\overline{R}^2$. Es decir, se escoge aquel q que maximice $\overline{R}^2$. Otra medida alternativa, comúnmente utilizada, es el **Criterio de Información de Akaike** (AIC):

$$\boxed{AIC(q) = \ln\left(\frac{\hat{u}'\hat{u}}{T}\right) + \frac{2q}{T}} \qquad (5)$$

donde $\hat{u}$ es el error estimado y T es el tamaño de la muestra.

Si algún Q es conocido, tal que Q≥q, escogemos aquel q que **minimice** AIC(q). El criterio de información de Akaike sigue la misma lógica del $\bar{R}^2$, esto es, premia un buen ajuste pero penaliza la pérdida de grados de libertad sufrida al agregar un mayor número de regresores.

Una alternativa a las anteriores es utilizar test F secuenciales sobre los Q–q últimos parámetros. El número óptimo de rezagos es aquel para el cual se rechaza la hipótesis nula de que los Q–q últimos parámetros son cero en conjunto.

1.2 Modelos con Rezagos Distribuidos Polinomiales

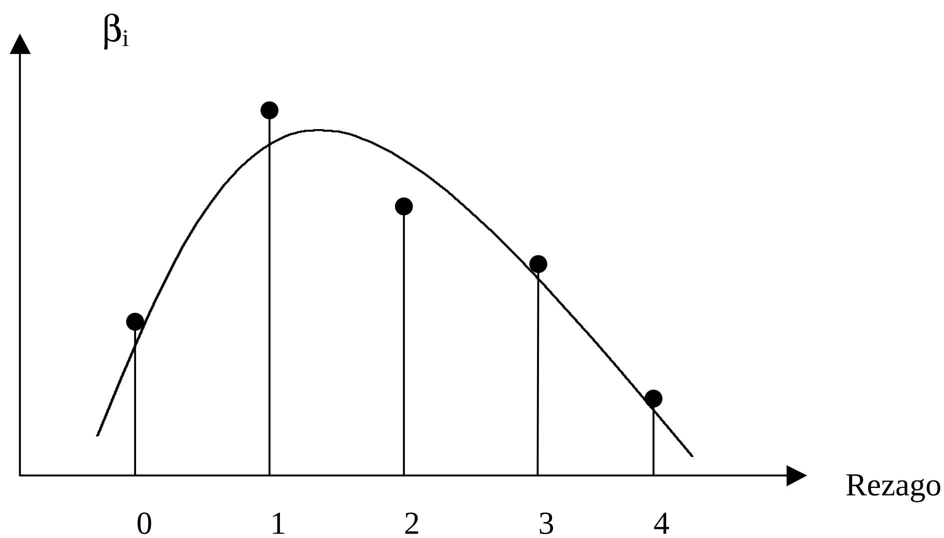
En algunos casos, el modelo contiene un número de rezagos muy grande. Un problema que ello puede traer es la existencia de multicolinealidad. En tales casos, es necesario imponer alguna estructura sobre la distribución de rezagos, a fin de reducir el número de parámetros a ser estimado.

Un modelo popular en la literatura es el de los rezagos distribuidos polinomiales de **Almon**. Este se basa en el supuesto de que el coeficiente del i-avo rezago se puede aproximar por un polinomio de un orden relativamente bajo:

$$\beta_i = \alpha_0 + \alpha_1\, i + \alpha_2\, i^2 + \ldots + \alpha_p\, i^p \qquad i=0, 1, 2, \ldots, q > p \qquad (5)$$

Generalmente, se escoge p=3 ó 4, como máximo.

Supongamos primero, por simplicidad, que tanto p y q son conocidos. De acuerdo a la ecuación (5) y después de re-agrupar términos, el modelo $Y_t = \alpha + \sum_{i=0}^{q} \beta_i X_{t-i} + u_t$ puede ser expresado como:



$$Y_t = \alpha + \alpha_0 (X_t + X_{t-1} + X_{t-2} + ... + X_{t-q}) + \alpha_1 (X_{t-1} + 2X_{t-2} + 3X_{t-3} + ... + qX_{t-q})$$

$$... + \alpha_p (X_{t-1} + 2^p X_{t-2} + 3^p X_{t-3} + ... + q^p X_{t-q}) + u_t$$

$$\equiv \alpha + \alpha_0 Z_{t0} + \alpha_1 Z_{t1} + ... + \alpha_p Z_{tp} + u_t \tag{6}$$

donde $Z_{t\,p} = X_{t-1} + 2^p X_{t-2} + 3^p X_{t-3} + ... + q^p X_{t-q}$ $\quad p \geq 1,\ t = q+1, ..., T$

Para t=q+1, ..., T se tiene:

$$\mathbf{X}^* = \begin{pmatrix} X_{q+1} & X_q & ... & X_1 \\ X_{q+2} & X_{q+1} & ... & X_2 \\ ... & ... & ... & ... \\ X_T & X_{T-1} & ... & X_{T-q} \end{pmatrix}_{(T-q)\text{ x }(q+1)}$$

Entonces,

$$\mathbf{Z} = \begin{pmatrix} Z_{q+1\,0} & Z_{q+1\,1} & ... & Z_{q+1\,p} \\ Z_{q+2\,0} & Z_{q+2\,1} & ... & Z_{q+2\,p} \\ ... & ... & ... & ... \\ Z_{T0} & Z_{T1} & ... & Z_{Tp} \end{pmatrix}_{(T-q)\text{ x }(p+1)} = \mathbf{X}^*\mathbf{H}$$

donde $\mathbf{H} = \begin{pmatrix} 1 & 0 & 0 & \dots & 0 \\ 1 & 1 & 1 & \dots & 1 \\ 1 & 2 & 4 & \dots & 2^p \\ \dots & \dots & \dots & \dots & \dots \\ 1 & q & q^2 & \dots & q^p \end{pmatrix}$

Entonces, en términos matriciales el modelo puede ser expresado como:

$$\mathbf{Y} = \alpha\, \boldsymbol{\iota} + \mathbf{Z}\, \boldsymbol{\delta} + \mathbf{u} \qquad (7)$$

donde $\boldsymbol{\iota} = \begin{pmatrix} 1 \\ 1 \\ \dots \\ 1 \end{pmatrix}_{(T-q)x1}$ $\qquad \boldsymbol{\delta} = \begin{pmatrix} \alpha_0 \\ \alpha_1 \\ \dots \\ \alpha_p \end{pmatrix}$

Asumiendo que u satisface todos los supuestos del modelo lineal clásico, se tiene que:

$$\hat{\boldsymbol{\delta}} = (\mathbf{Z}'\mathbf{M}_1\mathbf{Z})^{-1}(\mathbf{Z}'\mathbf{M}_1\mathbf{Y}) \qquad (8)$$

donde $\mathbf{M}_1 = \mathbf{I}_{T-q} - \boldsymbol{\iota}(\boldsymbol{\iota}'\boldsymbol{\iota})^{-1}\boldsymbol{\iota}'$

Notemos que $\beta_i = \begin{pmatrix} 1 & i & i^2 & \dots & i^p \end{pmatrix} \begin{pmatrix} \alpha_0 \\ \alpha_1 \\ \alpha_2 \\ \dots \\ \alpha_p \end{pmatrix}$ $\qquad$ i=0, ..., q

$$\boldsymbol{\beta} = \begin{pmatrix} \beta_0 \\ \beta_1 \\ \beta_2 \\ \dots \\ \beta_q \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 & \dots & 0 \\ 1 & 1 & 1 & \dots & 1 \\ 1 & 2 & 4 & \dots & 2^p \\ \dots & \dots & \dots & \dots & \dots \\ 1 & q & q^2 & \dots & q^p \end{pmatrix} \begin{pmatrix} \alpha_0 \\ \alpha_1 \\ \alpha_2 \\ \dots \\ \alpha_p \end{pmatrix} = \mathbf{H}\boldsymbol{\delta}$$

En consecuencia,

$$\boxed{\hat{\beta} = \mathbf{H}\,\hat{\delta}} \tag{9}$$

$$\hat{\mathrm{V}}\mathrm{ar}(\hat{\beta}) = \mathbf{H}\,\hat{\mathrm{V}}\mathrm{ar}(\hat{\delta})\mathbf{H}' = \hat{\sigma}^2\mathbf{H}(\mathbf{Z}'\mathbf{M}_1\mathbf{Z})^{-1}\mathbf{H}' \tag{10}$$

donde $\hat{\sigma}^2 = \dfrac{(\mathbf{Y} - \hat{\alpha}\iota - \mathbf{Z}\hat{\delta})'(\mathbf{Y} - \hat{\alpha}\iota - \mathbf{Z}\hat{\delta})}{T - q - (p+2)}$

❖ **¿Cómo Determinar el Grado del Polinomio**?

Si el número de rezagos, q , es conocido, entonces podemos contrastar un polinomio de grado p con otro de grado p* mediante un test F:

$$F(p^* - p, T - q - p^* - 2) = \frac{(\hat{u}_p'\hat{u}_p - \hat{u}_{p*}'\hat{u}_{p*})/(p^* - p)}{\hat{u}_{p*}'\hat{u}_{p*}/(T - q - p^* - 2)} \tag{11}$$

Se sugiere empezar con p* = q. El grado apropiado es aquel p mínimo para el cual el estadígrafo en (10) no supera al valor crítico de la tabla F con (p*–p, T–q–p*–2) grados de libertad, para un cierto nivel de significancia[1].

❖ ¿**Cómo Determinar el Largo del Rezago**?

Una vía es determinar en primer término q utilizando como criterio el $\bar{R}^2$ o AIC(q), y después determinar el grado del polinomio con el criterio antes señalado. Sin embargo, tal procedimiento secuencial no está libre de potenciales errores.

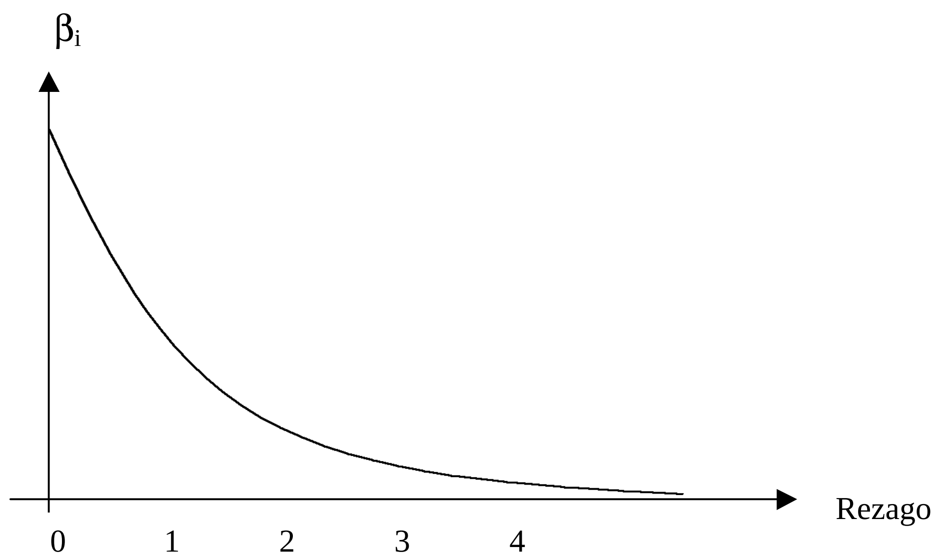
[1] Dado que los tests son secuenciales, en cada etapa existe la probabilidad de cometer el error tipo II (no rechazar H_0 cuando ésta es falsa). Por lo tanto, se ha sugerido utilizar un nivel de significancia corregido. Específicamente, si estamos en la etapa j, entonces deber usarse un nivel de significancia, $\varepsilon_j' = (1-\varepsilon)^j$, donde ε es el nivel de significancia originalmente escogido.

1.3 Modelos con Rezagos Geométricos

Este modelo—que se debe a Koyck—es uno de los más populares en la literatura de modelos de rezagos distribuidos. Este asume que:

$$\beta_i = \beta(1-\lambda)\lambda^i \qquad 0 \leq \lambda < 1 \qquad (12)$$

Gráficamente, se tiene que:



Esto es, la importancia de los rezagos declina geométricamente en el tiempo. Se tiene que el **multiplicador de corto plazo** viene dado por $\beta_0=\beta(1-\lambda)$, y el de **largo plazo** por $\sum_{i=0}^{\infty}\beta_i = \beta(1-\lambda)\sum_{i=0}^{\infty}\lambda^i = \beta$.

Operador de Rezagos: Se define el operador de rezagos como:

$$L^p X_t = X_{t-p}$$

Por ejemplo, $LX_t=X_{t-1}$, $L^2X_t=L(LX_t)=L(X_{t-1})=X_{t-2}$, etc. Por convención, se define $L^0X_t = X_t$.

1.3.1 Ejemplos de Modelos Económicos con Rezagos Geométricos

a) Modelo de Expectativas Adaptativas

$$Y_t = \alpha + \beta X_t^* + u_t \tag{13}$$

donde el mecanismo de expectativas viene dado por:

$$X_t^* - X_{t-1}^* = (1-\lambda)\,(X_t - X_{t-1}^*) \tag{14}$$

donde X_t es la observación corriente de X.

La ecuación (13) se puede re-escribir como:

$$X_t^* = \lambda X_{t-1}^* + (1-\lambda)\,X_t \tag{15}$$

Es decir, el valor esperado en t de X es un promedio ponderado de la expectativa del período pasado y del valor corriente de X.

Haciendo uso del operador de rezagos, la ecuación (15) puede escribirse como:

$$X_t^* = \lambda L X_t^* + (1-\lambda)\,X_t$$

esto es, $X_t^* = \dfrac{(1-\lambda)}{1-\lambda L} X_t$.

Dado que $0 \le \lambda < 1$, $\dfrac{1}{1-\lambda L} = 1 + \lambda L + \lambda^2 L^2 + \ldots$

$X_t^* = (1-\lambda)(X_t + \lambda X_{t-1} + \lambda^2 X_{t-2} + \ldots)$

Después de insertar X_t^* en la ecuación (13), se tiene que:

$$\boxed{Y_t = \alpha + \beta(1-\lambda)\sum_{i=0}^{\infty} \lambda^i X_{t-i} + u_t} \tag{16}$$

Esta es la llamada representación del **promedio móvil (MA)** del modelo de expectativas adaptivas.

Otra representación alternativa es la del **proceso autorregresivo (AR)**:

$$Y_t = \alpha + \beta\left(\frac{1-\lambda}{1-\lambda L}\right)X_t + u_t$$

esto es, $(1-\lambda L)Y_t = \alpha + \beta(1-\lambda)X_t + (1-\lambda L)u_t$

$$\boxed{Y_t = \alpha(1-\lambda) + \lambda Y_{t-1} + \beta(1-\lambda)X_t + u_t - \lambda u_{t-1}} \qquad (17)$$

<u>b) Modelo de Ajuste Parcial</u>

$$Y_t^* = \alpha + \beta X_t \qquad (18)$$

describe el nivel deseado de Y_t. Por ejemplo, el nivel de inventario óptimo como una función de las ventas efectivas en t.

El ajuste del período actual es una proporción de la diferencia entre el nivel deseado este período y el nivel observado el período anterior, más un cierto error aleatorio:

$$Y_t - Y_{t-1} = (1-\lambda)(Y_t - Y_{t-1}) + u_t$$

Es decir, $Y_t = (1-\lambda)Y_t^* + \lambda\, Y_{t-1}$ (19)

Si substituimos (18) en (19), tenemos que:

$$\boxed{Y_t = \alpha(1-\lambda) + \lambda Y_{t-1} + \beta(1-\lambda)\, X_t + u_t} \qquad (20)$$

la cual es la **representación AR** del modelo de ajuste parcial.

El modelo de ajuste parcial también admite una representación de **promedio móvil**:

$$Y_t\,(1-\lambda L) = \alpha(1-\lambda) + \beta(1-\lambda)\, X_t + u_t$$

esto es, $Y_t = \dfrac{\alpha(1-\lambda)}{1-\lambda L} + \dfrac{\beta(1-\lambda)}{1-\lambda L}X_t + \dfrac{u_t}{1-\lambda L}$

Pero, dado que $0 \leq \lambda < 1$, $\dfrac{1}{1-\lambda L} = 1 + \lambda L + \lambda^2 L^2 + \ldots$ Entonces:

$$\frac{\alpha(1-\lambda)}{1-\lambda L} = \alpha(1-\lambda)(1+\lambda L+\lambda^2 L^2+\ldots) = \alpha(1-\lambda)(1+\lambda+\lambda^2+\ldots) = \frac{\alpha(1-\lambda)}{1-\lambda} = \alpha$$

$$\frac{\beta(1-\lambda)}{1-\lambda L} X_t = \beta(1-\lambda)(1+\lambda L+\lambda^2 L^2+\ldots) X_t = \beta(1-\lambda)\sum_{i=0}^{\infty} \lambda^i X_{t-i}$$

$$\frac{u_t}{1-\lambda L} = (1+\lambda L+\lambda^2 L^2+\ldots) u_t = \sum_{i=0}^{\infty} \lambda^i u_{t-i}$$

Con ello, $Y_t = \alpha + \beta(1-\lambda)\sum_{i=0}^{\infty} \lambda^i X_{t-i} + \sum_{i=0}^{\infty} \lambda^i u_{t-i}$

La expresión anterior se puede simplificar notando que el promedio móvil infinito correspondiente al error se puede describir como:

$\varepsilon_t = \lambda\varepsilon_{t-1} + u_t \quad \Leftrightarrow \quad (1-\lambda L)\varepsilon_t = u_t$

Con ello, $\varepsilon_t = \sum_{i=0}^{\infty} \lambda^i u_{t-i}$

De lo anterior, la representación de MA(∞) viene dada por:

$$\boxed{Y_t = \alpha + \beta(1-\lambda)\sum_{i=0}^{\infty} \lambda^i X_{t-i} + \varepsilon_t \qquad \varepsilon_t = \lambda\varepsilon_{t-1} + u_t} \qquad (21)$$

II ESTIMACION DE MODELOS CON REZAGOS GEOMETRICOS

Nos concentraremos únicamente en la estimación de la representación autorregresiva (AR) del modelo de rezagos geométricos:

$Y_t = \alpha(1-\lambda) + \lambda Y_{t-1} + \beta(1-\lambda) X_t + u_t$

Bajo la representación AR hay dos casos a ser analizados. El primer caso es aquel en el cual el error del modelo, u_t, es bien comportado. En tal circunstancia, MICO proporciona estimadores consistentes (aunque sesgados en muestras pequeñas). El segundo caso es aquel en el cual el error es autocorrelacionado. Bajo dicho escenario, MICO es inconsistente porque el error está correlacionado con uno de los regresores del modelo: Y_{t-1}. Por lo tanto, se requiere de un método alternativo de estimación, tal como variables instrumentales.

La discusión correspondiente a la estimación de la representación de promedio móvil infinito se encuentra en Greene, capítulo 19.

2.1 Errores No Correlacionados

Sea $\mathbf{z}_s'=(1 \;\; Y_{s-1} \;\; X_s)$ el vector de regresores correspondiente a la observación s-ava, $\mathbf{Z}$ la matriz que recoge los vectores $\mathbf{z}_1'$, $\mathbf{z}_2'$,..., $\mathbf{z}_s'$, ..., $\mathbf{z}_T'$, y $\mathbf{u}$ el vector columna cuyos elementos son u_1, u_2, ..., u_t, ..., u_T. MICO sería insesgado siempre y cuando $E(\mathbf{z}_s u_t)=0$, $\forall$ t, s. Sin embargo, ello no cumple porque, por ejemplo, $E(\mathbf{z}_3 u_2)=E(Y_2 u_2)\neq 0$.

En muestras grandes, sin embargo, se tiene que:

$$\operatorname{plim}\frac{1}{T}\mathbf{Z}'\mathbf{u} = \operatorname{plim}\frac{1}{T}\begin{pmatrix} \sum_{t=2}^{T} u_t \\ \sum_{t=2}^{T} Y_{t-1}u_t \\ \sum_{t=2}^{T} X_t u_t \end{pmatrix} = \begin{pmatrix} E(u_t) \\ E(Y_{t-1}u_t) \\ E(X_t u_t) \end{pmatrix} = \mathbf{0} \qquad (22)$$

Por lo tanto, siempre y cuando, $\operatorname{plim}\frac{1}{T}\sum_{s=1}^{T}\mathbf{z}_s\mathbf{z}_s'=Q$, matriz positiva definida e invertible, tendremos que MICO es consistente.

2.2 Errores Correlacionados

Supongamos que:

$Y_t = \alpha(1-\lambda) + \lambda Y_{t-1} + \beta(1-\lambda) X_t + u_t \qquad u_t = \rho u_{t-1} + \varepsilon_t$

En este caso MICO no es sólo sesgado en muestras pequeñas, sino también inconsistente:

$$\text{plim}\frac{1}{T}\mathbf{Z}'\mathbf{u} = \text{plim}\frac{1}{T}\begin{pmatrix}\sum_{t=2}^{T} u_t \\ \sum_{t=2}^{T} Y_{t-1}u_t \\ \sum_{t=2}^{T} X_t u_t\end{pmatrix} = \begin{pmatrix} E(u_t) \\ E(Y_{t-1}u_t) \\ E(X_t u_t)\end{pmatrix} = \begin{pmatrix} 0 \\ \rho E(Y_{t-1}u_{t-1}) \\ E(X_t u_t)\end{pmatrix} \neq \mathbf{0}$$

Sea **W** la matriz de instrumentos. Deseamos que $\text{plim}\frac{1}{T}\mathbf{W}'\mathbf{u} = 0$ y que $\text{plim}\frac{1}{T}\mathbf{W}'\mathbf{Z} \neq \mathbf{0}$. En este caso, la constante y la variable X_t sirven de instrumentos de sí mismos. Dado que X_{t-1} no está correlacionada con u_t pero sí lo está con Y_{t-1}, sirve de instrumento para Y_{t-1}. Sea, entonces, $\mathbf{W}=(\boldsymbol{\iota}\ \mathbf{X}\ \mathbf{X}_{-1})$. El estimador de variables instrumentales vendrá dado por:

$$\boxed{\hat{\beta}_{VI} = (\mathbf{W}'\mathbf{Z})^{-1}\mathbf{Z}'\mathbf{Y}}$$

donde $\sqrt{T}(\hat{\beta}_{VI} - \beta) \xrightarrow{d} N(\mathbf{0}, \sigma_u^2\, \text{plim}\left(\frac{\mathbf{Z}'\mathbf{W}}{T}\right)^{-1}\left(\frac{\mathbf{Z}'\mathbf{Z}}{T}\right)\left(\frac{\mathbf{W}'\mathbf{Z}}{T}\right)^{-1})$