TEMA 5: Modelos de regresión con Series de Tiempo estacionarias

Profesor: Graciela Sanroman

Año 2016

Hasta ahora explicita o implicitamente se consideraron modelos del tipo:

$$E(y_t \mid x_t) = x_t^{'}\beta \quad t = 1, 2....T$$

o el equivalente

$$\begin{aligned} y_t &= x_t^{'}\beta + u_t \\ E(u_t &\mid x_t) = 0 \end{aligned}$$

Donde la variable dependiente y los regresores refieren al mismo período de tiempo. Pero la teoría económica sugiere que en muchos casos las relaciones entre las variables económicas en distintos períodos.

Por ejemplo: la demanda de electricidad

$$\begin{aligned} Q_t &= \alpha + \beta P_t + \gamma K_t + u_t \\ K_t &= \theta_0 + \delta Y_t + \theta_1 P_{t-1} + \theta_2 P_{t-2} + ... + v_t \end{aligned}$$

Donde Q_t es la demanda doméstica de energía eléctrica, la cual es función de:

- los precios corrientes P_t,
- y de la intensidad del uso del capital K_t, que es a su vez función de la renta Y_t y de los precios en períodos anteriores.

Así, cuando cambia el precio, se espera un impacto inmediato que causará un uso menos intensivo del equipamiento eléctrico del hogar. Pero cambiar dicho equipamiento requiere tiempo (costes de ajuste) y por lo tanto el efecto completo de un cambio en los precios de la electricidad sobre la demanda no se completará hasta pasado un conjunto de períodos. Sustituyendo

$$\begin{aligned} Q_t &= \alpha + \beta P_t + \gamma \left(\theta_0 + \delta Y_t + \theta_1 P_{t-1} + \theta_2 P_{t-2} + ... + v_t \right) + u_t \\ &= \alpha_0 + \alpha_1 Y_t + \beta_0 P_t + \beta_1 P_{t-1} + \beta_2 P_{t-2} + ... + \varepsilon_t \end{aligned}$$

En resumen podemos encontrar situaciones en las cuales el impacto de una variable sobre otra se da:

- de forma instantánea (P_t)
- después de cierto tiempo $(P_{t-j}$ con $j > 1)$
- ambas cosas a la vez (P_t) y $(P_{t-j}$ con $j > 1)$
- persistente (aún luego de muchos períodos el efecto se mantiene)

Ello conduce a considerar modelos como los siguientes:

a. Modelos en los cuales no aparece la endógena rezagada como regresor

$$
\begin{aligned}
y_t &= \alpha + \beta_1 x_{t-1} + u_t \\
y_t &= \alpha + \beta_0 x_t + \beta_1 x_{t-1} + \beta_2 x_{t-2} + u_t \\
y_t &= \alpha + \beta_2 x_{t-2} + \beta_3 x_{t-3} + u_t
\end{aligned}
$$

Ejemplos: inversión, crecimiento de la oferta monetaria cambios en las tasas de interés.

b. Modelos en los cuales aparece la endógena rezagada como regresor: existen variables que presentan inercia, lo que hace que dependan de su propio pasado, por ejemplo la inflación (π_t es la tasa de inflación en el período t y m_t es la cantidad de dinero en la economía)

$$\pi_t = \alpha + \beta_1 \pi_{t-1} + \gamma m_t + u_t$$

Algunos aspectos importantes a tener en cuenta a la hora de especificar el modelo son:
- la frecuencia de los datos
- la estructura del error u_t, particularmente la correlación serial será un elemento clave a la hora de estimar modelos de series de tiempo.

Modelos dinámicos

Veremos varios modelos distinguiendo según la estructura del error:

- El error no presenta autocorrelación:
  - Modelos con regresores que no incluyen la endógena rezagada
  - Modelos con regresores que incluyen la endógena rezagada
- El error presenta autocorrelación
  - Modelos con regresores que no incluyen la endógena rezagada (el error podría estar incorrelacionado con los regresores):
  - Modelos con regresores que incluyen la endógena rezagada (aquí la autocorrelación implicará correlación entre el error y los regresores)

Como veremos más adelante estas cuestiones serán muy relevantes a la hora de determinar las propiedades de los estimadores. Previo a considerar esto veremos como operan los efectos parciales en estos modelos.

Modelos dinámicos: algunas propiedades y definiciones

Para entender como operan estos modelos supongamos la parte sistemática de un modelo

$$y_t = \alpha + \beta_0 x_t + \beta_1 x_{t-1} + \beta_2 x_{t-2}$$

Consideremos que la economía está en un equilibrio

$$\begin{aligned} & y^*, x^* \\ y^* &= \alpha + (\beta_0 + \beta_1 + \beta_2)\, x^* \end{aligned}$$

supongamos que en $t = t_0$ se produce un cambio transitorio en x

$$\begin{aligned} x_{t0} &= x^* + \Delta \\ y_{t0} &= y^* + \beta_0 \Delta \\ y_{t0+1} &= y^* + \beta_1 \Delta \\ y_{t0+2} &= y^* + \beta_2 \Delta \\ y_{t0+s} &= y^* \quad s > 2 \end{aligned}$$

Modelos dinámicos: funciones de respuesta y multiplicadores
FUNCION DE RESPUESTA AL IMPULSO: Sucesión de los efectos que tiene un cambio transitorio de 1 en x sobre la trayectoria esperada de y, en este caso

$$\beta_0, \beta_1, \beta_2, 0, 0, ..., 0$$

Notar: En cada caso los nombres y subíndices de los parámetros dependen de cómo esté escrito el modelo, por ejemplo si la parte sistemática del modelo fuera $y_t = \delta_0 + \delta_1 x_t + \delta_2 x_{t-1} + \delta_3 x_{t-2}$ la función de respuesta al impulso sería $\delta_1, \delta_2, \delta_3, 0, 0...0$.

Modelos dinámicos: funciones de respuesta y multiplicadores

FUNCIÓN DE RESPUESTA AL ESCALÓN: Es el efecto esperado en la trayectoria de y un cambio permanente en x

$$\begin{aligned}
x_s &= x^* + \Delta \text{ s>t0} \\
y_{t0} &= y^* + \beta_0 \Delta \\
y_{t0+1} &= y^* + (\beta_1 + \beta_0)\,\Delta \\
y_{t0+s} &= y^* + (\beta_2 + \beta_1 + \beta_0)\,\Delta \quad s \geqslant 2
\end{aligned}$$

Modelos dinámicos: funciones de respuesta y multiplicadores

MULTIPLICADOR DE IMPACTO/Multiplicador de corto plazo (es el efecto contemporáneo)

$$\frac{\partial}{\partial x_t} E(y_t \mid X) = \beta_0$$

MULTIPLICADOR DE EQUILIBRIO/Multiplicador de largo plazo (es el efecto acumulado en el largo plazo)

$$\sum_{j=0}^{\infty} \beta_j$$

Modelos dinámicos: funciones de respuesta y multiplicadores

NORMALIZACIÓN DE COEFICIENTES, denominamos coeficientes normalizados a:

$$\beta_j^{'} = \frac{\beta_j}{\sum_{j=0}^{\infty} \beta_j}$$

RETARDO MEDIANO, el retardo mediano es m (la cantidad de períodos que el proceso demora en alcanzar la mitad del efecto total)

$$m \text{ es tal que } \sum_{j=0}^{m} \beta_j^{'} = 0.50$$

RETARDO MEDIO (rm), el retardo medio nos brinda una idea de si el impacto total está concentrado o diluido en el tiempo

$$rm = \sum_{j=0}^{\infty} j\beta_j^{'}$$

Multiplicadores: ejemplo 1

Consideremos el caso de un modelo general con q retardos en la variable x que no incluye la endógena rezagada (la parte sistemática del modelo)

$$\begin{aligned} y_t &= \alpha + B(L)x_t \\ B(L) &= \beta_0 + \beta_1 L + \beta_2 L^2 + ...\beta_q L^q \end{aligned}$$

Nota: se considera el caso de funciones $B(L)$ donde las raíces del polínomio caen fuera del circulo unitario

Función de respuesta al impulso: β_0, β_1, $\beta_2, ...,$ β_q.

Función de respuesta al escalón: β_0, $\beta_0 + \beta_1, \beta_0 + \beta_1 + \beta_2$, $, \beta_0 + \beta_1 + \beta_2 + .. + \beta_q$

Multiplicador de impacto (corto plazo): β_0

Multiplicador de equilibrio (largo plazo): $\beta_0 + \beta_1 + \beta_2 + .. + \beta_q$

Retardo mediano: m tal que $\frac{\sum_{j=0}^{m} \beta_j}{\beta_0 + \beta_1 + \beta_2 + .. + \beta_q} = 0.5$

Retardo medio: $rm = \frac{\beta_1 + 2\beta_2 + .. + q\beta_q}{\beta_0 + \beta_1 + \beta_2 + .. + \beta_q}$

Multiplicadores: ejemplo 2

Consideremos el caso de un modelo general con 1 retardo en la variable dependiente del modelo (la parte sistemática del modelo)

$$A(L)y_t = \alpha + \beta x_t \text{ siendo } A(L) = 1 - \phi_1 L \text{ con } |\phi_1| < 1$$

Notar que

$$\begin{aligned} y_t &= \alpha + \beta x_t + \phi_1 y_{t-1} = \frac{\alpha}{1-\phi_1} + \frac{\beta}{1-\phi_1 L} x_t \\ &= \frac{\alpha}{1-\phi_1} + \beta x_t + \beta \phi_1 L x_t + \beta (\phi_1 L)^2 x_t + \ldots\ldots \end{aligned}$$

Función de respuesta al impulso: β, $\beta\phi_1, \beta\phi_1^2$

Función de respuesta al escalón: β, $\beta(1+\phi_1)$, $\beta(1+\phi_1+\phi_1^2), \ldots.$

Multiplicador de impacto (corto plazo): β

Multiplicador de equilibrio (largo plazo): $\frac{\beta}{1-\phi_1}$

Retardo mediano: $m = \frac{\ln 0.5}{\ln \phi_1} - 1$

Retardo medio: $rm = \frac{\phi_1}{1-\phi_1}$

Multiplicadores: caso general

$$
\begin{aligned}
A(L)y_t &= \alpha + B(L)x_t \\
A(L) &= 1 - \phi_1 L - \phi_2 L^2 + \dots \phi_p L^p \\
B(L) &= \beta_0 + \beta_1 L + \beta_2 L^2 + \dots \beta_q L^q
\end{aligned}
$$

donde las raíces de los polinomios $A(L)$ y $B(L)$ se encuentran todas fuera del círculo unitario

Multiplicador de impacto (corto plazo): $B(0) = \beta_0$

Multiplicador de equilibrio (largo plazo): $\frac{B(1)}{A(1)} = \frac{\beta_0+\beta_1+\beta_2+\dots\beta_q}{1-\phi_1-\phi_2-\dots\phi_p}$

Retardo medio: $rm = \frac{B'(1)}{B(1)} - \frac{A'(1)}{A(1)} = \frac{\beta_1+2\beta_2+\dots+q\beta_q}{\beta_0+\beta_1+\beta_2+\dots\beta_q} - \frac{-\phi_1-2\phi_2\dots-p\phi_q}{1-\phi_1-\phi_2-\dots-\phi_p}$

Estimación de modelos dinámicos

Consideraremos modelos del tipo

$$y_t = x_t^{'}\beta + u_t \text{ con } t = 1, 2, ...T$$

Nota importante: x_t puede incluir retardos de variables por ejemplo si p_t es un precio en el período t $x_t = \left(p_t, p_{t-1}, ...p_{t-q}\right)$ o también incluir retardos de la variable dependiente, por ejemplo $x_t = \left(y_{t-1}, p_t, p_{t-1}, ...p_{t-q}\right)$.
El tratamiento de estos modelos en la literatura distingue: si se trata de valores retardados de la variable endógena o si se trata de valores retardados de las variables explicativas, pero en definitiva la diferencia fundamental estará dada por el hecho de si se cumple o no el supuesto de ortogonalidad (es decir de independencia en media del error de la ecuación y los regresores), en la práctica es mucho más probable que este supuesto se cumpla cuando se trata de retardos de las variables explicativas que cuando se trata de la variable dependiente rezagada pero no siempre es el caso.

Estimación de modelos dinámicos

Diremos que los regresores corresponden a variables exógenas cuando

$$
\begin{aligned}
E\left(u_t \mid X\right) &= 0 \text{ con } t = 1, 2 \ldots T \\
X &= \left(x_1, x_2, \ldots . x_T\right)
\end{aligned}
$$

Notar que en el caso que y sea $I(1)$ este supuesto se deja de cumplir automáticamente. El hecho que la variable dependiente y/o las variables explicativas sean $I(d)$ con $d > 0$ hace que las reglas de análisis e inferencia en el marco del modelo de regresión lineal que se estudiaron hasta ahora no sea válido. Entonces, veremos primero como tratar con modelos en los cuales todas las variables son estacionarias y en el capítulo siguiente estudiaremos como proceder cuando alguna o varias de las variables incluidas en el modelo son n no estacionarias.

Estimación de modelos dinámicos

Si las variables y y x son estacionarias y el supuesto de independencia en media del error se cumple, las complicaciones se originan en la estimación de modelos de series temporales son básicamente:

- Potencial multicolinealidad cuando se incluyen retardos
- ¿Cómo saber el número de retardos a incluir? Si es corresponde incorporar ∞ retardos, ¿como estimar el modelo?
- Presencia de autocorrelación y/o heterocedasticidad
- Tamaños muestrales suelen ser reducidos por lo cual no será apropiado aplicar propiedades asintóticas

Propiedades en muestras finitas del estimador MCO en el contexto de series temporales

Insesgadez del estimador MCO: El estimador MCO de un modelo de series temporales

$$y_t = f(x_t, \beta) + u_t \text{ con } t = 1, 2, ... T$$

donde x_t ($x_t$puede incluir retardos de las variables) será insesgado si se cumplen los siguientes tres supuestos:

- **Supuesto 1: Linealidad** en los parámetros
$$f(x_t, \beta) = x_t^{'}\beta$$
- **Supuesto 2: Exogeneidad estricta (**Independencia en media): para cada t, el valor esperadodel termino de error, condicionando en las variables explicativas en todos los períodos, es igual a 0.
$$E(u_t \mid X) = 0 \text{ con } t = 1, 2...T \qquad X = (x_1, x_2, ...x_T)$$
- **Supuesto 3: No multicolinealidad perfecta** (rango de X es k con probabilidad 1): en la muestra (y, por lo tanto en los procesos de series temporales subyacentes) ningun regresor es constante o una combinación lineal perfecta de los restantes.

Teorema 1: Bajo los supuestos 1, 2 y 3 $E(\widehat{\beta}_{MCO} \mid X) = \beta$ y por lo tanto $E(\widehat{\beta}_{MCO}) = \beta$

Propiedades en muestras finitas del estimador MCO en el contexto de series temporales

Varianza de $\widehat{\beta}_{MCO}$ y teorema Gauss-Markov

- **Supuesto 4: Homocedasticidad** $Var(u_t \mid X) = Var(u_t) = \sigma_t^2 = \sigma^2$ con $t = 1, 2...T$ (la varianza no puede depender de X y tampoco variar a través del tiempo)
- **Supuesto 5: No correlación serial** $cov(u_t, u_j \mid X) = 0 \ \forall \ t \neq j$ (condicional en X los errores de dos períodos diferentes tienen que estar incorrelacionados)

Teorema 2: Si se cumplen los supuestos 1 a 5 entonces
$Var(\widehat{\beta}_{MCO} \mid X) = \sigma^2(X'X)^{-1}$ y $Var(\widehat{\beta}_{MCO}) = \sigma^2 E(X'X)^{-1}$
Teorema 3: Si se cumplen los supuestos 1 a 5 entonces el estimador $\widehat{\sigma}^2 = \frac{SCR}{T-k-1}$ es un estimador insesgado de σ^2, donde
$SCR = \sum_{t=1}^{T} \left(y_t - x_t' \widehat{\beta}_{MCO} \right)^2$ y k es la cantidad de regresores en el modelo.
Teorema 4: Teorema de Gauss-Markov Si se cumplen los supuestos 1 a 5 los estimadores **MCO** son los estimadores lineales insesgados óptimos, condicionando a X.

Inferencia en muestras finitas en el contexto de series temporales

Supuesto 6: Normalidad de los errores: los errores u_t son independientes de X y están idéntica e independientemente distribuidos según una distribución $N(0, \sigma^2)$

Teorema 5: Bajo los supuestos 1 a 6 los estimadores MCO siguen una distribución normal, condicionados a X. Además, el estadístico correspondiente a la hipótesis nula Ho: $\beta_j = 0$ sigue una distribución (bajo Ho) *t-student* y el estadístico F correspondiente a la hipótesis nula Ho: $R\beta - r = 0$ sigue una distribución (bajo Ho) F. La construcción de intervalos de confianza de la forma habitual es válida.

Modelo de retardos distribuidos

$$y_t = \alpha + \sum_j \beta_j x_{t-j} + u_t$$

Modelo de retardos distribuidos finitos: $j = 0, 1, 2, ...q$
Modelo de retardos distribuidos infinitos: $j = 0, 1, 2, ...\infty$

NOTACION:

L operador de retardos

$$
\begin{aligned}
Lx_t &= x_{t-1} \\
L^2 x_t &= Lx_{t-1} = x_{t-2} \\
&\;.... \\
L^q x_t &= ..... = x_{t-q}
\end{aligned}
$$

Polinomios de retardos

$$
\begin{aligned}
A(L) &= \alpha_0 - \alpha_1 L - \alpha_2 L^2 - .... - \alpha_p L^p \\
B(L) &= \beta_0 + \beta_1 L + \beta_2 L^2 + ... + \beta_q L^q
\end{aligned}
$$

Caso particular pero de mucho interés

$$
\begin{aligned}
B(L) &= 1 + bL + (bL)^2 + ... \\
&= \sum_{j=0}^{\infty} (bL)^j
\end{aligned}
$$

Si $|a| < 1$

$$A(L) = \frac{1}{1 - bL}$$

En otro orden si tengo:

$$y_t = \alpha + B(L)x_t + u_t$$

a) rezagos finitos:

$$\begin{aligned} B(L) &= \beta_0 + \beta_1 L + \beta_2 L^2 + ... + \beta_q L^q \\ B(1) &= \beta_0 + \beta_1 + \beta_2 + ... + \beta_q \text{ multiplicador de largo plazo} \end{aligned}$$

b) rezagos infinitos:

$$\begin{aligned} B(L) &= \beta_0 + \beta_1 L + \beta_2 L^2 + ... \\ B(1) &= \sum_{j=0}^{\infty} \beta_j \text{ multiplicador de largo plazo sólo si } \left| \sum_{j=0}^{\infty} \beta_j \right| < \infty \end{aligned}$$

Si tengo

$$
\begin{aligned}
y_t &= \alpha + \beta \sum_{j=0}^{\infty} \gamma^j L^j x_t + u_t \quad |\gamma| < 1 \\
y_t &= \alpha + \frac{\beta}{1-\gamma L} x_t + u_t \\
(1-\gamma L) y_t &= (1-\gamma L)\alpha + \beta x_t + (1-\gamma L) u_t \\
y_t &= \alpha^{'} + \gamma y_{t-1} + \beta x_t + v_t \\
v_t &= u_t - \gamma u_{t-1}
\end{aligned}
$$

aparece la variable endógena retardada como regresor.
El tratamiento de estos modelos en la literatura distingue si se trata de valores retardados de la variable endógena o si se trata de valores retardados de las variables explicativas, pero en definitiva la diferencia fundamental estará dada por el hecho de si se cumple o no el supuesto de ortogonalidad, en la práctica es mucho más probable que el mismo se cumpla cuando se trata de retardos de variables explicativas.

Diremos que los retardos corresponden a variables exógenas cuando

$$E\left(u_s \mid x^t\right) = 0$$
$$x^t = (x_1, x_2, ...x_t)$$

Si ese supuesto se cumple las complicaciones se originan en la estimación de modelos que incluyen retardos son básicamente:

- Potencial multicolinealidad
- ¿cómo saber el número de retardos a incluir? si es ∞ como estimarlo

Veremos:
a) Modelo de retardos distribuidos finito
b) Modelo de retardos distribuidos polinomiales (retardos de Almon)
c) El modelo de retardos infinito geométrico (variable dependiente retardada) El modelo de Koyck

Modelo de retardos distribuidos finito no restringido

$$
\begin{aligned}
y_t &= \alpha + \sum_{j=0}^{q} \beta_j x_{t-j} + u_t \\
E\left(\mathrm{u} \mid X\right) &= 0 \\
E\left(\mathrm{u}\,\mathrm{u}' \mid X\right) &= \sigma^2 I
\end{aligned}
$$

Distinguiré: q conocido, desconocido.
Si q es conocido entonces estoy en el contexto del modelo de regresión clásico.
Pero en general q no será conocida.

Para determinarla: recurro a estimar el modelo por MCO

1. $R^2 ajustado = 1 - \frac{T-1}{T-k} R^2$ (maximizar)
2. AKAIKE$=\ln\frac{\widehat{u}'\widehat{u}}{T} + \frac{2q}{T}$ (minimizar)

3. Secuencia de contrastes F
   Sea q<Q

$$y_t = \gamma + \sum_{j=0}^{Q} \beta_j x_{t-j} + u_t \text{ modelo no restringido}$$
$$y_t = \gamma^* + \sum_{j=0}^{q} \beta_j^* x_{t-j} + v_t \text{ modelo restringido}$$

$$F = \frac{\left(\frac{\widehat{u}_*' \widehat{u}_*}{q} - \frac{\widehat{u}' \widehat{u}}{Q} \right) / (Q - q)}{\frac{\widehat{u}' \widehat{u}}{Q} / (T - Q - 1)}$$

Problemas:

R2 y akaike recompensan buenos ajustes, penaliza pérdida de grados de libertad, Akaike puede llevarnos a un sobreajuste pero no tiene los problemas de los métodos secuenciales. contrastes F: significación y verdadera distribución de la secuencia de contrastes no está derivada.

Modelo de retardos distribuidos polinomiales (Almon)

Si la longitud de los retardos es larga y tenemos (relativamente) pocos datos podríamos estar interesados en imponer alguna estructura en la distribución de los retardos

Almon(1965) sugirio imponer

$$\begin{aligned} \beta_i &= \alpha_0 + \alpha_1 i + \alpha_2 i^2 + ... + \alpha_p i^p \\ i &= 0, ..., q > p \end{aligned}$$

La idea que subyace es que siempre podré aproximar una función por un polinomio del orden apropiado. El orden del polinomio generalmente se toma de forma que sea bajo

$$\begin{aligned} \beta_0 &= \alpha_0 \\ \beta_1 &= \alpha_0 + \alpha_1 + \alpha_2 + \alpha_3 + ... + \alpha_p \\ \beta_2 &= \alpha_0 + 2\alpha_1 + 4\alpha_2 + 8\alpha_3 + ... + 2^p \alpha_p \\ & \\ \beta_q &= \alpha_0 + q\alpha_1 + q^2\alpha_2 + q^3\alpha_3 + ... + q^p \alpha_p \end{aligned}$$

Entonces reduzco de q+1 a p+1 los parámetros a estimar

Después de sustituir y agrupar el modelo queda

$$
\begin{aligned}
y_t &= \alpha + \alpha_0(x_t + x_{t-1} + ... + x_{t-q}) \\
&\quad + \alpha_1(x_{t-1} + 2x_{t-2} + 3x_{t-3} ... + qx_{t-q}) \\
&\quad + \alpha_2(x_{t-1} + 4x_{t-2} + 9x_{t-3} ... + q^2 x_{t-q}) \\
&\quad + \\
&\quad + \alpha_p(x_{t-1} + 2^p x_{t-2} + 3^p x_{t-3} ... + q^p x_{t-q}) \\
&\quad + u_t
\end{aligned}
$$

$$
\begin{aligned}
y_t &= \alpha + \alpha_0 z_{0t} + \alpha_1 z_{1t} + ...\alpha_p z_{pt} \\
z_{0t} &= x_t + x_{t-1} + ... + x_{t-q} \\
z_{1t} &= x_{t-1} + 2x_{t-2} + 3x_{t-3} ... + qx_{t-q} \\
\\
z_{pt} &= x_{t-1} + 2^p x_{t-2} + 3^p x_{t-3} ... + q^p x_{t-q}
\end{aligned}
$$

En forma compacta

$$\beta = H\alpha$$

$$H = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 2 & 4 & .. & .. & 2^p \\ 1 & 3 & 9 & & & 3^p \\ & & & & & \\ & & & & & \\ 1 & q & q^2 & & & q^p \end{pmatrix}$$

$$\begin{aligned} y &= XH\alpha + u \\ y &= Z\alpha + u \end{aligned}$$

Esto lo puedo estimar por MCO o por MCGF

$$\begin{aligned} \widehat{\alpha} &= \left(Z'Z\right)^{-1} Z'y \\ \widehat{\beta} &= H\widehat{\alpha} \\ \widehat{Var}(\widehat{\beta}) &= H\widehat{Var}(\widehat{\alpha})H' \end{aligned}$$

En algunos textos se recomienda establecer $\beta_{-1} = \beta_{q+1} = 0$

Dos problemas: q?, p?

Si q conocida:

Procedimiento secuencial

a) $p_1 = q$ equivale a no imponer restricciones en los parámetros. Estimo el modelo obtengo los residuos $\widehat{u}_1$

b) procedo a reducir $p_2 < q$, estimo el modelo y obtengo $\widehat{u}_2$

$$\frac{\left[\frac{\widehat{u}_2'\widehat{u}_2}{P_2} - \frac{\widehat{u}_1'\widehat{u}_1}{P_1}\right] / (P_1 - P_2)}{\frac{\widehat{u}_1'\widehat{u}_1}{P_1} / \, (T - P_1 - 1)} \sim F(P_1 - P_2, T - P_1 - 1)$$

Un valor alto tendería a sugerir que no es aconsejable reducir el orden del polinomio.

Un aspecto que es importante tener en cuenta es que en el análisis secuencial es necesario modificar el nivel de significación: $\alpha_j = 1 - (1 - \alpha)^j$

Amemiya y Morinune (1974) p=2,3 funciona increiblemente bien.

Si q no es conocido, entonces se aconseja utilizar primero en modelo no restringido para determinar la longitud de los retardos.

Modelo de retardos distribuidos infinitos

Utilidad: es preferible porque no tengo que decidir sobre q pero tendré que imponer estructura. El modelo queda

$$y_t = \alpha + \sum_{j=0}^{\infty} \beta_j x_{t-j} + u_t$$

por lo que hay ∞ parámetros para estimar.

Una alternativa para imponer estructura es utilizar el modelo de retardos distribuidos geométricos (Modelo de Koyck). En el modelo de Koyck se supone

$$\begin{aligned} \beta_0 &= \beta \\ \beta_j &= \lambda \beta_{j-1}, \quad j \geqslant 1, \ |\lambda| < 1 \end{aligned}$$

$$y_t = \alpha + \beta \sum_{j=0}^{\infty} \lambda^j L^j x_t + u_t$$

Modelo de retardos distribuidos geométricos (Modelo de Koyck)

El modelo se convierte en

$$\begin{aligned} y_t &= \alpha + \beta x_t + \beta\lambda x_{t-1} + \beta\lambda^2 x_{t-2} + \beta\lambda^3 x_{t-3} + ... + u_t \\ &= \alpha + \beta(x_t + \lambda x_{t-1} + \lambda^2 x_{t-2} + \lambda^3 x_{t-3} + ... + \lambda^{t-1} x_1) \\ &\quad + \beta\lambda^t (x_0 + \lambda x_{-1} + ...) + u_t \end{aligned}$$

Ahora sólo tengo 3 parámetros para estimar: α, λ, β

Puedo reescribir el modelo como:

$$\begin{aligned} y_t &= \alpha + \beta x_t (1 + \lambda L + \lambda^2 L^2 + \lambda^3 L^3 + ...) + u_t \\ &= \alpha + \beta x_t \frac{1}{1 - \lambda L} + u_t \end{aligned}$$

entonces

$$\begin{aligned} y_t (1 - \lambda L) &= \alpha (1 - \lambda L) + \beta x_t + (1 - \lambda L) u_t \\ y_t - \lambda y_{t-1} &= \alpha^* + \beta x_t + u_t^* \\ y_t &= \alpha^* + \beta x_t + \lambda y_{t-1} + u_t^* \end{aligned}$$

Modelo de retardos distribuidos geométricos (Modelo de Koyck)
Multiplicador de corto plazo: β
Multiplicador de largo plazo: $\frac{\beta}{1-\lambda}$
Retardo medio $\frac{\lambda}{1-\lambda}$
Retardo mediano $\frac{\ln 0.5}{\ln \lambda} - 1$
Ejercicio: determinar la función de respuesta al impulso y la función de respuesta al escalón.

Autocorrelación en modelos de series temporales

Hasta ahora consideramos el caso en el cual se cumplían los supuestos 1 a 6. Consideraremos ahora el caso en el cual el supuesto 5 podría no cumplirse. Es decir, los errores del modelo podrían presentar correlación serial.

Notar que en el contexto de series temporales el incumplimiento del supuesto 5 es muy probable (tanto desde un punto de vista teórico como en cuanto a la evidencia que se encuentra en la literatura).

Veremos:

- Contrastes de presencia de correlación serial
- Propiedades del estimador MCO cuando hay correlación serial
- Estimadores alternativos con mejores propiedades cuando hay correlación serial

Contrastes de correlación serial

Partiremos de un modelo

$$y_t = x_{t'}\beta + u_t, \qquad t = 1, 2, ... T$$

- Supondremos:
  - que se verifican los supuestos 1 a 4, y en particular que las variables explicativas son estrictamente exógenas ($E(u_t|X) = 0$)
  - que la correlación serial sigue un proceso $AR(1)$, con $|\rho| < 1$ y $e_t \overset{iid}{\sim} (0, \sigma^2)$

$$u_t = \rho u_{t-1} + e_t \tag{1}$$

Notar que contrastar la presencia de correlación serial corresponde a someter a prueba la hipótesis $Ho : \rho = 0$ frente a $H_1 : \rho \neq 0$. Rechazar Ho implica obtener evidencia respecto al no cumplimiento del supuesto 5.

Contrastes de correlación serial

Para someter a prueba Ho :

- Si observaramos $\{u_t\}_{t=1,2,....T}$ podríamos simplemente hacer una regresión MCO de u_t sobre u_{t-1} y (bajo el supuesto 6, que aquí sería $e_t \overset{iid}{\sim} N(0, \sigma^2)$) utilizar una prueba t usando el estadístico $\frac{\hat{\rho}}{ee(\hat{\rho})}$.
- El problema es que no observamos $\{u_t\}_{t=1,2,....T}$, una alternativa es remplazar por $\{\widehat{u}_t\}_{t=1,2,....T}$ y hacer una regresión MCO de $\widehat{u}_t$ sobre $\widehat{u}_{t-1}$ calculando los errores estándar robustos a heterocedasticidad. Este contraste, no obstante, sólo tiene validez asintótica y como dijimos el tamaño de las muestras de series temporales suele ser pequeño.

Contrastes de correlación serial (Durbin-Watson)

Un contraste para la hipótesis de autocorrelación es el **Contraste de Durbin-Watson**. Es importante tener en cuenta que dicho contraste es válido únicamente bajo las siguientes condiciones:

- se trata de correlación serial de tipo AR(1) es decir $u_t = \rho u_{t-1} + e_t$, $|\rho| < 1$ y $e_t \overset{iid}{\sim} N(0, \sigma^2)$
- las X son "fijas", esto no se cumple en un conjunto de situaciones, en particular cuando la variable dependiente rezagada forma parte de x_t.

El estadístico de DW se calcula como

$$DW = \frac{\sum\limits_{t=2}^{T} (\hat{u}_t - \hat{u}_{t-1})^2}{\sum\limits_{t=1}^{T} \hat{u}_t^2}$$

donde $\widehat{u}_t = y_t - x_t' \widehat{\beta}_{MCO}$ (Puede demostrarse que está relacionado con la prueba anterior, ya que $DW \approx 2(1 - \hat{\rho})$)

Contrastes de correlación serial (Durbin-Watson)

- El estadístico DW no sigue una distribución Normal y tampoco otra distribución "conocida", Durbin y Watson derivaron la distribución del estadístico mediante procedimientos de simulación, y los valores críticos se pueden consultar en sus tablas. Estos dependen del tamaño muestral, del número de regresores y de la existencia de un término constante.
- La lectura de los valores críticos de la distribución DW es menos directa que en otros casos:
  - El estadístico varía entre 0 y 4
  - Valores cercanos a 2 llevarán a No Rechazar H_0
  - Valores cercanos a 4 o a 0 llevarán a Rechazar H_0
  - Sin embargo, existen regiones intermedias en que el valor del estadístico no permite concluir
  - En las tablas se dispone de un valor d_L y un valor d_U que dependen de la cantidad de regresores incluidos en la regresión.
  - Las regiones no concluyentes son: $d_L < DW < d_U$ y $(4 - d_U) < DW < (4 - d_L)$

Contrastes de correlación serial (h-Durbin)

Cuando la variable dependiente rezagada aparece entre los regresores el contraste DW ya no es válido. Una alternativa es utilizar el contraste propuesto por Durbin (1970) denominado h de Durbin. Consideraremos modelos del tipo

$$y_t = \alpha y_{t-1} + x_{t'}\beta + u_t, \qquad t = 1, 2, \ldots T$$

donde la correlación serial sigue un proceso $AR(1)$, $u_t = \rho u_{t-1} + e_t$ con $|\rho| < 1$ y $e_t \overset{iid}{\sim} N(0, \sigma^2)$

El contraste consiste en someter a prueba $Ho : \rho = 0$ frente $H1 : \rho \neq 0$.

El estadístico h-Durbin y su distribución asintótica (bajo Ho) es

$$\begin{aligned} h - Durbin &= \hat{\rho}\sqrt{\frac{T}{1 - T\ \widehat{Var}(\widehat{\alpha}_{MCO})}} \overset{a}{\sim} N(0,1) \\ \hat{\rho} &= 1 - \frac{DW}{2} \end{aligned}$$

Contrastes de correlación serial (h-Durbin)

Comentarios:

- El estadístico h-Durbin no está definido si $\widehat{Var}(\widehat{\alpha}_{MCO}) > 1/T$
- Su validez es asintótica
- Este contraste puede también aplicarse a modelos que incluyen otros retardos de y, $y_t = \alpha_1 y_{t-1} + \alpha_2 y_{t-2} + ... + \alpha_p y_{t-p} + x_{t'}\beta + u_t$, siempre que el modelo para los errores sea $u_t = \rho u_{t-1} + e_t$.
- IMPORTANTE: rechazar *Ho* implica encontrar evidencia que el supuesto 1 (de exogeneidad estricta) no se cumple y por lo tanto concluir que los estimadores MCO serán sesgados e inconsistentes.

Contrastes de correlación serial (Breusch-Godfrey)

Consideraremos modelos del tipo

$$y_t = x_{t'}\beta + u_t, \qquad t = 1, 2, ... T$$

donde la correlación serial sigue un proceso $AR(p)$, con $\|\rho\| < 1$ y $e_t \overset{iid}{\sim} (0, \sigma^2)$

$$u_t = \rho_1 u_{t-1} + \rho_2 u_{t-2} + .. + \rho_p u_{t-p} + e_t$$

Someteremos a prueba la hipótesis

$$\begin{aligned} Ho &: \ \rho_1 = \rho_2 = ... = \rho_p = 0 \\ H_1 &: \ \text{algún } \rho_j \neq 0 \text{ con } j = 1, 2...p \end{aligned}$$

La estrategia del contraste de Breusch-Godfrey es utilizar un contraste de los Multiplicadores de Lagrange

Contrastes de correlación serial (Breusch-Godfrey)

Procedimiento:

1. Hacer una estimación MCO de y_t sobre x_t y obtener los residuos $\widehat{u}_t = y_t - x_t^{'}\widehat{\beta}_{MCO}$
2. Hacer una estimación MCO de $\widehat{u}_t$sobre x_t , $\widehat{u}_{t-1}, \widehat{u}_{t-2}, ..., \widehat{u}_{t-p}$ para todo $t = (p+1), (p+2), ...T$ y obtener el estadístico $LM = (T-p)R^2_{\widehat{u}}$ donde $R^2_{\widehat{u}}$ es el coeficiente de determinación de la estimación realizada.
3. Bajo Ho el estadístico LM converge en distribución a una χ^2_p, por lo cual rechazaré Ho a un nivel de significación 0.05 (y por lo tanto encontraremos evidencia de correlación serial) si $LM > \chi^2_p(0.95)$

Otra alternativa es realizar el contraste a través de una prueba F para la hipótesis antes considerada.

Contraste de Breusch-Godfrey

Comentarios:

- Es un contraste más general que el DW y también que el h-Durbin
- Puede utilizarse para contrastar la presencia de autocorrelación de tipo $AR(p)$ o MA(q))
- A diferencia de DW se puede utilizar en modelos que incluyen retardos de la variable dependiente entre los regresores u otro tipo de regresores estocásticos
- Es un contraste que tiene validez asintótica, es necesario ser muy cautos cuando el tamaño de la muestra es reducido
- La elección del número de retardos p o q es importante. Un p/q pequeño puede no capturar autocorrelaciones significativas de mayor orden, pero un p/q grande puede reducir la potencia de la prueba.

Propiedades del estimador MCO con regresores estrictamente exógenos cuando hay correlación serial

En el caso que se cumplan los supuestos 1,2,3,4 y 6, pero haya correlación serial el estimador MCO seguirá siendo insesgado y consistente pero (similar a cuando existe heterocedasticidad)

- MCO ya no será óptimo (no será el estimador MELI)
- la matriz de varianzas y covarianzas y por lo tanto los errores estándar y los contrastes basados en estadísticos MCO habituales ya no serán válidos
- sin embargo, las medidas habituales de bondad de ajuste del modelo R^2 y $R^2 - \textit{ajustado}$ seguirán siendo asintóticamente válidas

Propiedades del estimador MCO cuando la variable dependiente rezagada es un regresor y hay correlación serial

Consideremos el modelo

$$y_t = \alpha y_{t-1} + u_t, \qquad t = 1, 2, \ldots T$$

con

$$\begin{aligned} u_t &= \rho u_{t-1} + e_t \\ \|\rho\| &< 1 \text{ y } \quad e_t \overset{iid}{\sim} (0, \sigma^2) \end{aligned}$$

En este caso si hay correlación serial de los residuos se viola el supuesto 1, por lo cual el estimador MCO será sesgado e inconsistente.
En la práctica es muy difícil distinguir casos de correlación serial de los errores de situaciones en las cuales existe persistencia genuina (persistencia genuina implica que la variable dependiente rezagada es una variable relevante en el modelo). Por ejemplo, en el caso anterior, la presencia de autocorrelación podría estar indicando que necesitamos un $AR(2)$ $y_t = \alpha_1 y_{t-1} + \alpha_2 y_{t-2} + a_t$.

Estimadores apropiados en modelos con regresores estrictamente exógenos y autocorrelación de tipo AR(1) en los errores

Como se planteó anteriormente si los regresores son estrictamente exógenos, MCO es insesgado y consistente pero la correlación serial de los errores implica que la estimación MCO de la matriz de varianzas y covarianzas no es válida.

Al igual que en el caso de presencia de heterocedasticidad en cortes transversales en este caso podemos proceder por MCG. Notar que si los errores son **AR(1)** con $u_t = \rho u_{t-1} + e_t$, la matriz de varianzas y covarianzas del error (suponiendo que la misma no depende de X) y que la varianza e_t es constante en el tiempo resulta en

$$
\begin{aligned}
E(uu\prime|X) &= E(uu\prime) \\
&= \frac{\sigma_e^2}{1-\rho^2}\begin{pmatrix} 1 & \rho & \rho^2 & \cdots & \rho^{T-1} \\ \rho & 1 & \rho & \cdots & \rho^{T-2} \\ \rho^2 & \rho & 1 & \cdots & \rho^{T-3} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \rho^{T-1} & \rho^{T-2} & \rho^{T-3} & \cdots & 1 \end{pmatrix}
\end{aligned}
$$

Estimadores apropiados en modelos con regresores estrictamente exógenos y autocorrelación de tipo AR(1) en los errores

Notar que una alternativa (si ρ fuese conocida) sería estimar los coeficientes a través de MCG, el estimador resultante sería MELI. El problema es que lo usual es que ρ sea desconocida. Además, para que este estimador sera MELI el modelo de autocorrelación supuesto (AR(1)) debe ser verdadero.

Veremos:

- dos procedimientos muy similares para estimar los modelos cuando los errores son AR(1) con ρ desconocido: Cochrane-Orcutt y Prais-Winsten, ambos son estrategias de estimación MCGF
- estimación MCO con cálculo de la matriz de varianzas y covarianzas robusta a autocorrelación/heterocedasticidad (siendo la forma de la autocorrelación/heterocedasticidad desconocida)

Estimador Cochrane-Orcutt

Consideraremos modelos del tipo

$$\begin{aligned} y_t &= x_t^{'}\beta + u_t, \qquad t = 1, 2, ... T \\ u_t &= \rho u_{t-1} + e_t \qquad |\rho| < 1;\ e_t \overset{iid}{\sim} (0, \sigma^2) \end{aligned}$$

El estimador Cochrane-Orcutt resulta de realizar las siguientes etapas

1. Hacer una estimación MCO de y_t sobre x_t y obtener los residuos $\widehat{u}_t = y_t - x_t^{'}\widehat{\beta}_{MCO}$
2. Hacer una estimación MCO de $\widehat{u}_t$sobre $\widehat{u}_{t-1}$ y obtener $\widehat{\rho}$
3. Obtener las cuasi-diferencias

$$\begin{aligned} \widetilde{y}_t &= y_t - \widehat{\rho} y_{t-1} \\ \widetilde{x}_{0t} &= 1 - \widehat{\rho} \\ \widetilde{x}_{jt} &= x_{jt} - \widehat{\rho} x_{jt-1}\ j = 1, 2 ... k \end{aligned}$$

4. Estimar por MCO el modelo transformado $\widetilde{y}_t$ sobre $\widetilde{x}_t (= \widetilde{x}_{0t}, \widetilde{x}_{1t}, ...\widetilde{x}_{jt})$ con $t = 2, 3 ... T$
5. Iterar en el modelo hasta obtener convergencia

Estimador Cochrane-Orcutt

Notar que el estimador Cochrane-Orcutt implica multiplicar al sistema de T ecuaciones $y_1, y_2....y_T$ por la matriz de dimensión $T-1 \times T$

$$\begin{pmatrix} -\widehat{\rho} & 1 & 0 & \cdots & 0 \\ 0 & -\widehat{\rho} & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & -\widehat{\rho} & 1 \end{pmatrix}$$

ello implica que se pierde la primera observación, y esto implica una pérdida de eficiencia (perdemos observaciones y fundamentalmente la ecuación en niveles). Esta "debilidad" del estimador Cochrane-Orcutt se resuelve utilizando el estimador de Prais-Winsten.

Estimador Prais-Winsten

Al igual que Cochrane-Orcutt se consideran modelos del tipo

$$\begin{aligned} y_t &= x_t^{'}\beta + u_t, \qquad t = 1, 2, \ldots T \\ u_t &= \rho u_{t-1} + e_t \qquad |\rho| < 1; \; e_t \overset{iid}{\sim} (0, \sigma^2) \end{aligned}$$

El estimador **Prais-Winsten** resulta de realizar las siguientes etapas

1. Hacer una estimación MCO de y_t sobre x_t y obtener los residuos $\widehat{u}_t = y_t - x_t^{'}\widehat{\beta}_{MCO}$
2. Hacer una estimación MCO de $\widehat{u}_t$sobre $\widehat{u}_{t-1}$ y obtener $\widehat{\rho}$
3. Obtener:

$$\begin{aligned} \widetilde{y}_t &= y_t - \widehat{\rho} y_{t-1}; \; \widetilde{x}_{0t} = 1 - \widehat{\rho}; \; \widetilde{x}_{jt} = x_{jt} - \widehat{\rho} x_{jt-1} \text{ para } t \geq 2, j = 1..k \\ \widetilde{y}_1 &= \sqrt{1 - \widehat{\rho}^2} y_1; \widetilde{x}_{01} = \sqrt{1 - \widehat{\rho}^2}; \widetilde{x}_{jt} = \sqrt{1 - \widehat{\rho}^2} x_{j1} \text{ para } j = 1..k \end{aligned}$$

4. Estimar por MCO el modelo transformado $\widetilde{y}_t$ sobre $\widetilde{x}_t (= \widetilde{x}_{0t}, \widetilde{x}_{1t}, \ldots. \widetilde{x}_{jt})$ con $t = 1, 2, 3 \ldots T$
5. Iterar en el modelo hasta obtener convergencia

Estimación MCO con cálculo de la matriz de varianzas y covarianzas robusta a autocorrelación

El uso de esta estrategia ha ganado terreno en los últimos años. Tiene una serie de ventajas respecto a los procedimientos de MCGF vistos previamente:

- en el caso que las variables explicativas no sean estrictamente exógenas MCGF será inconsistente
- los métodos vistos son válidos sólo si la autocorrelación sigue un modelo AR(1)
- la supuesta ganancia de eficiencia del método MCGF se ve generalmente afectada por el hecho de utilizar una estimación de ρ en lugar de su verdadero valor

Estimación MCO con cálculo de la matriz de varianzas y covarianzas robusta a autocorrelación

Hay varias alternativas para calcular errores robustos a formas desconocidas de autocorrelación/heterocedasticidad. Estos estimadores se conocen en la literatura como errores estándar HAC (Heteroskedasticity and Autocorrelation Consistent standar errors). Una alternativa es seguir el siguiente procedimiento (Newey-West, 1987; Wooldridge, 1989):

1. Estimar el modelo $y_t = x_t^{'}\beta + u_t, \quad t = 1, 2, \ldots T$ por MCO y obtener $\widehat{\beta}_{MCO}$, los residuos $\widehat{u}_{t,MCO}$, $\widehat{\sigma}^2_{MCO}$, $ee(\widehat{\beta}_j)$ $j = 0, 1, \ldots k$
2. Calcular los residuos $\widehat{r}_{1t}$, $t = 1, 2, \ldots T$ en la regresión auxiliar de x_{1t} sobre una constante y $x_{2t}, x_{3t}, \ldots . x_{kt}$, y calcular $\widehat{a}_t = \widehat{r}_{1t}\widehat{u}_{t,MCO}$
3. Obtener $\widehat{v} = \sum_{t=1}^{T} \widehat{a}_t + 2 \sum_{h=1}^{g} \left[1 - \frac{h}{g+1}\right] \sum_{t=h+1}^{T} \widehat{a}_t \widehat{a}_{t-h}$ donde g es un número entero que controla cuanta autocorrelación permitimos a la hora de calcular el error estándar.
4. Obtener $HACse\left(\widehat{\beta}_1\right) = \left(\frac{se\left(\widehat{\beta}_{1MCO}\right)}{\widehat{\sigma}^2_{MCO}}\right)^2 \sqrt{\widehat{v}}$

Este procedimiento puede realizarse para cada coeficiente $\widehat{\beta}_j$ del modelo.

Aplicaciones de modelos de retardos distribuídos: Modelo de expectativas adaptativas (Cagan, 1956)
Es un modelo para la demanda de dinero:

$$
\begin{aligned}
m_t &= \beta_1 + \beta_2 E_t \pi_{t+1} + u_t \\
& \beta_1 + \beta_2 \pi^e_{t+1} + u_t
\end{aligned}
$$

donde m_t es la demanda real de dinero agregada π^e_{t+1} es la inflación esperada. Se supone el siguiente mecanismo de formación de expectativas:

$$
\begin{aligned}
\pi^e_{t+1} &= E_t \pi_{t+1} = E_{t-1}\pi_t + \lambda\left(\pi_t - E_{t-1}\pi_t\right) \\
&= \lambda \pi_t + (1-\lambda) E_{t-1}\pi_t \\
&= \lambda \pi_t + (1-\lambda) \pi^e_t
\end{aligned}
$$

Sustituyendo recursivamente

$$
\begin{aligned}
\pi_t^e &= \lambda \pi_{t-1} + (1-\lambda)\,\pi_{t-1}^e \\
\pi_{t+1}^e &= \lambda \pi_t + \lambda(1-\lambda)\,\pi_{t-1} + (1-\lambda)^2\,\pi_{t-1}^e \\
&= \\
&= \lambda \pi_t + \lambda(1-\lambda)\,\pi_{t-1} + \lambda(1-\lambda)^2\,\pi_{t-2} + \ldots.. \\
&= \lambda \sum_{s=0}^{\infty} (1-\lambda)^s\,\pi_{t-s} = \lambda \sum_{s=0}^{\infty} \left[(1-\lambda)\,L\right]^s \pi_t \\
&= \frac{\lambda}{1-(1-\lambda)L}\pi_t
\end{aligned}
$$

Reescribiendo la ecuación del modelo resulta en

$$
\begin{aligned}
m_t &= \beta_1 + \beta_2 \pi_{t+1}^{e} + u_t \\
&= \beta_1 + \beta_2 \lambda \frac{1}{1-(1-\lambda)L} \pi_t + u_t \\
m_t \left(1-(1-\lambda)L\right) &= \beta_1 \left(1-(1-\lambda)L\right) + \beta_2 \lambda \pi_t + \left(1-(1-\lambda)L\right) u_t \\
m_t - (1-\lambda) m_{t-1} &= \lambda \beta_1 + \beta_2 \lambda \pi_t + u_t - (1-\lambda) u_{t-1} \\
m_t &= \lambda \beta_1 + \beta_2 \lambda \pi_t (1-\lambda) m_{t-1} + u_t - (1-\lambda) u_{t-1} \\
m_t &= \lambda \beta_1 + \beta_2 \lambda \pi_t + (1-\lambda) m_{t-1} + v_t
\end{aligned}
$$

Notar: por construcción $cov(m_{t-1}, v_t) \neq 0$ por lo cual la estimación MCO será sesgada e inconsistente. Una alternativa será estimar por Variables Instrumentales, π_{t-j} podría ser un instrumento válido, también se podrían buscar instrumentos "externos".

Aplicaciones de modelos de retardos distribuídos: Modelo de ajuste parcial (Nerlove, 1963).
Supongamos que el nivel deseado de capital en la economía está dado por:

$$K_t^* = \alpha_1 + \alpha_2 Y_t + u_t$$

donde K es stock de capital y Y el PBI, el problema aquí es que no se observa el capital deseado, se supone que el mecanismo por el cual el stock de capital se ajusta al nivel deseado está dado por:

$$K_t - K_{t-1} = \delta \left(K_t^* - K_{t-1} \right)$$

MODELO DE AJUSTE PARCIAL:
Si δ es 1, en cada período el stock de capital será igual al deseado (caso sin costes de ajuste) pero si existen costes de ajuste δ será en general inferior a 1. El modelo puede escribirse como:

$$K_t = \delta K_t^* + (1 - \delta) K_{t-1}$$

iterando

$$K_t = \delta K_t^* + \delta(1 - \delta) K_{t-1}^* + +\delta(1 - \delta)^2 K_{t-2}^* + ...$$

Si escribimos el modelo original en función de observables

$$K_t = \delta\alpha_1 + \delta\alpha_2 Y_t + (1 - \delta)\, K_{t-1} + \delta u_t$$

Notar que si u_t no presenta autocorrelación (lo cual es posible en el marco de este modelo) la estimación MCO será consistente para los coeficientes del modelo transformado.