Boletín Trimestral de Coyuntura, **n. 88**
Junio, 2003

Desagregación temporal mediante modelos dinámicos: el método de Santos Silva y Cardoso

Enrique M. Quilis[1]

S.G. Cuentas Nacionales
Instituto Nacional de Estadística

[1] Agradezco a Ana Abad sus comentarios así como la colaboración de Fátima Cardoso. Las opiniones formuladas en esta nota corresponden al autor, sin que el INE coincida de forma necesaria con ellas.

1. Introducción

En esta nota se presenta una nueva función Matlab que amplía la librería de desagregación temporal descrita en Quilis (2003). Esta ampliación está orientada a la extensión de la metodología básica de Chow y Lin (1971) para tener en cuenta especificaciones dinámicas más complejas, de forma que se incrementa la precisión de las estimaciones y se extiende el marco de aplicación de estas técnicas.

El procedimiento aplicado ha sido expuesto por Santos Silva y Cardoso (2001) y se inscribe a su vez en una línea de desarrollo iniciada por Salazar et al. (1994) y Gregoir (1994), en la que se utilizan modelos dinámicos explícitos para abordar el problema de la desagregación temporal. Una revisión panorámica de estos trabajos se encuentra en di Fonzo (2002).

El texto se organiza de la siguiente forma: en la segunda sección se expone el método de Santos Silva y Cardoso, detallándose su relación con el procedimiento de Chow y Lin. La codificación en Matlab, los resultados y gráficos principales se presentan en la tercera sección.

2. El procedimiento de Santos Silva y Cardoso: Teoría

A continuación se presenta el método de Santos Silva y Cardoso (2001) tal y como es expuesto en di Fonzo (2002). Dichos autores proponen el siguiente modelo de alta frecuencia como representación de la relación existente entre una variable inobservable y_t y un vector x_t de p indicadores:

[2.1] $$y_t = \phi y_{t-1} + x'_t \beta + \varepsilon_t \qquad -1 < \phi < 1 \qquad t = 1 \cdots n$$

siendo $\varepsilon_t \sim iid\ N(0,\sigma_a)$.

Desarrollando [2.1] se obtiene el siguiente modelo de retardos distribuidos:

[2.2] $$y_t = \sum_{j=1}^{p} \sum_{h=0}^{\infty} \theta_{j,h} x_{t-h} + \sum_{h=0}^{\infty} \psi_h \varepsilon_{t-h}$$

donde:

[2.3] $$\theta_{j,h} = \beta_j \phi^h \qquad \psi_h = \phi^h$$

En consecuencia, el modelo [2.1] es una parametrización parsimoniosa de un modelo dinámico bastante general, gracias a las restricciones expresadas en [2.3]. Expresando [2.1] en forma matricial se obtiene:

[2.4] $$D_\phi y = x\beta + q\eta + \varepsilon$$

siendo:

$$D_\phi : nxn = \begin{bmatrix} 1 & 0 & 0 & \cdots & 0 & 0 \\ -\phi & 1 & 0 & \cdots & 0 & 0 \\ 0 & -\phi & 1 & \cdots & 0 & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & 0 & \cdots & -\phi & 1 \end{bmatrix}$$

Las condiciones iniciales del sistema se encuentran en los términos *q* y η (*truncation remainder*):

$$q = \begin{bmatrix} \phi \\ 0 \\ \cdots \\ 0 \end{bmatrix} \qquad \eta = E(y_0)$$

Agrupando términos en [2.4]:

[2.5] $$D_\phi y = z\gamma + \varepsilon$$

siendo $z=[x \;\; q]$ y $\gamma=[\beta \;\; \eta]'$.

De esta manera, la condición inicial se incluye como parámetro adicional que ha de ser estimado junto con los restantes parámetros del modelo. Despejando *y* en [2.5] se obtiene:

[2.6] $$y = D_\phi^{-1} z\gamma + D_\phi^{-1}\varepsilon = z(\phi)\gamma + u$$

El modelo de baja frecuencia implicado por [2.6] se obtiene premultiplicando esta expresión por la matriz de agregación temporal *C* de dimensión Nxn:

[2.7a] $$C = I_N \otimes c$$

donde *N* es el número de observaciones de baja frecuencia, $\otimes$ denota el producto tensorial, *c* es un vector fila de tamaño *s* que define el tipo de agregación temporal que se desea realizar y *s* es el número de períodos de alta frecuencia contenidos en uno de baja. Si *c=[1,1,...,1]* se trata de un caso de agregación de un flujo, si *c=[1/s,1/s,...,1/s]* se trata del promedio de un índice y, si *c=[0,0,...,1]*, se obtiene un problema de interpolación.

Cuando el tamaño muestral de los indicadores no es conformable con el de la serie de baja frecuencia (*n>sN*), se presenta una situación de extrapolación. Este caso se puede tratar fácilmente mediante una ampliación de la matriz de agregación temporal *C*, de forma que se consideran nuevas columnas de ceros que no distorsionan la relación de agregación temporal ni dificultan la incorporación de los últimos *n-sN* datos de los indicadores en el proceso de estimación de y_t. En consecuencia, la matriz *C* expandida es:

[2.7b] $$C = (I_N \otimes c \mid O_{N,n-sN})$$

Tanto si se aplica [2.7a] (distribución o interpolación) como [2.7b] (extrapolación), el modelo de baja frecuencia es:

[2.8] $$Y = Z(\phi)\gamma + U$$

Este modelo, a diferencia de su generador de alta frecuencia, sólo involucra variables observables: *Y* y $Z(\phi)$, si bien esta última es resultado de la transformación de *Z* mediante CD_ϕ.

En virtud de la hipótesis de gaussianidad, la función de verosimilitud, logarítmicamente transformada, del modelo anterior es:

[2.9] $$\ell = -\frac{N}{2}\ln(2\pi) - \frac{1}{2}\ln(|C v(\phi) C'|) - \frac{1}{2}(Y - Z(\phi)\gamma)'(C v(\phi) C')^{-1}(Y - Z(\phi)\gamma)$$

La correspondiente matriz de varianzas y covarianzas es:

$$V(\phi) = C\, v(\phi)\, C'$$

con:

$$v(\phi) = \sigma_a (D_\phi\, D'_\phi)^{-1}$$

La función de verosimilitud es altamente no lineal en ϕ por lo que su estimación se realiza mediante una búsqueda en su dominio estacionario $-1<\phi<1$, junto con la estimación por mínimos cuadrados generalizados de los parámetros contenidos en γ:

[2.10] $$\hat{\gamma} = (Z(\phi)'(C\,v(\phi)\,C')^{-1}Z(\phi))^{-1}(Z(\phi)'(C\,v(\phi)\,C')^{-1}Y)$$

La estimación de la serie de alta frecuencia procede siguiendo la propuesta ELIO[1] de Chow-Lin:

[2.11] $$\hat{y} = z(\phi)\hat{\gamma} + v(\phi)C'\,V(\phi)^{-1}\hat{U} = z(\phi)\hat{\gamma} + L_{\phi}\hat{U}$$

Los intervalos de confianza asociados a [2.11] se obtienen a través de su matriz de varianzas y covarianzas:

[2.12] $$\Sigma_{\hat{y}} = (I_n - L_{\phi}C)v(\phi) + (z(\phi) - L_{\phi}Z(\phi))\,\Sigma_{\hat{\gamma}}\,(z(\phi) - L_{\phi}Z(\phi))'$$

Un análisis detallado del procedimiento de Chow-Lin se encuentra en di Fonzo (1987) y en Quilis (2001).

[1] Estimación lineal, insesgada y óptima. Optimalidad definida en términos de varianza mínima.

3. El procedimiento de Santos Silva y Cardoso: Función Matlab

A continuación se expone una función Matlab que permite la aplicación del procedimiento de desagregación temporal expuesto en la sección anterior. Asimismo, esta función también permite realizar extrapolaciones. Para ello basta con considerar como *input* un vector de indicadores de dimensión *4N+e*, siendo *N* el número de observaciones de baja frecuencia y *e* el de extrapolaciones.

La función posee la siguiente estructura:

```
function res=ssc(Y,x,ta,s,type)
% PURPOSE: Temporal disaggregation using the dynamic Chow-Lin method
%                   proposed by Santos Silva-Cardoso (2001).
% ------------------------------------------------------------
% SYNTAX: res=ssc(Y,x,ta,s,type);
% ------------------------------------------------------------
% OUTPUT: res: a structure
%        res.meth               ='Santos Silva-Cardoso';
%        res.ta                 = type of disaggregation
%        res.type               = method of estimation
%        res.N                  = nobs. of low frequency data
%        res.n                  = nobs. of high-frequency data
%        res.pred               = number of extrapolations
%        res.s                  = frequency conversion between low and high freq.
%        res.p                  = number of regressors (+ intercept)
%        res.Y                  = low frequency data
%        res.x                  = high frequency indicators
%        res.y                  = high frequency estimate
%        res.y_dt               = high frequency estimate: standard deviation
%        res.y_lo               = high frequency estimate: sd - sigma
%        res.y_up               = high frequency estimate: sd + sigma
%        res.u                  = high frequency residuals
%        res.U                  = low frequency residuals
%        res.gamma              = estimated model parameters (including y(0))
%        res.gamma_sd           = estimated model parameters: standard deviation
%        res.gamma_t            = estimated model parameters: t ratios
%        res.rho                = dynamic parameter phi
%        res.beta               = estimated model parameters (excluding y(0))
%        res.beta_sd            = estimated model parameters: standard deviation
%        res.beta_t             = estimated model parameters: t ratios
%        res.aic                = Information criterion: AIC
%        res.bic                = Information criterion: BIC
%        res.val                = Objective function used by the estimation method
%        res.r                  = grid of dynamic parameters used by the estimation method
%        res.et                 = elapsed time
% ------------------------------------------------------------
% INPUT: Y: Nx1 ---> vector of low frequency data
%       x: nxp ---> matrix of high frequency indicators (without intercept)
%       ta: type of disaggregation
%           ta=1 ---> sum (flow)
%           ta=2 ---> average (index)
%           ta=3 ---> last element (stock) ---> interpolation
%       s: number of high frequency data points for each low frequency data points
%           s= 4 ---> annual to quarterly
%           s=12 ---> annual to monthly
%           s= 3 ---> quarterly to monthly
%       type: estimation method:
%           type=0 ---> weighted least squares
%           type=1 ---> maximum likelihood
% ------------------------------------------------------------
% LIBRARY: aggreg
% ------------------------------------------------------------
% SEE ALSO: chowlin, litterman, fernandez, td_plot, td_print
% ------------------------------------------------------------
% REFERENCE: Santos, J.M.C. y Cardoso, F.(2001) "The Chow-Lin method
% using dynamic models",Economic Modelling, vol. 18, p. 269-280.
```

Se invoca mediante un guión (*script*) de Matlab de la forma genérica:

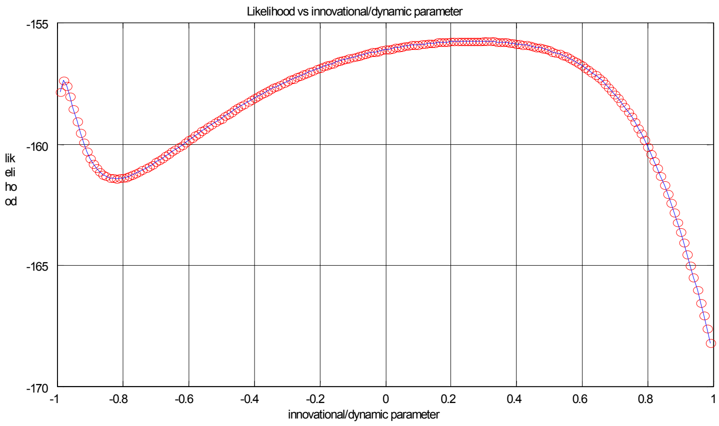
```
Y=load('Y.anu');
x=load('x.tri');
% Type of aggregation
ta=1;
% Frequency conversion
s=4;
% Method of estimation
type=1;
% Calling the function: output is loaded in a structure called res
res=ssc(Y,x,ta,s,type);
% Calling printing function
% Name of ASCII file for output
file_sal='td.sal';
output=0; % Exclude series from ASCII file
td_print(res,file_sal,output);
edit td.sal;
% Calling graph function
td_plot(res);
```

La salida impresa es la siguiente[1]:

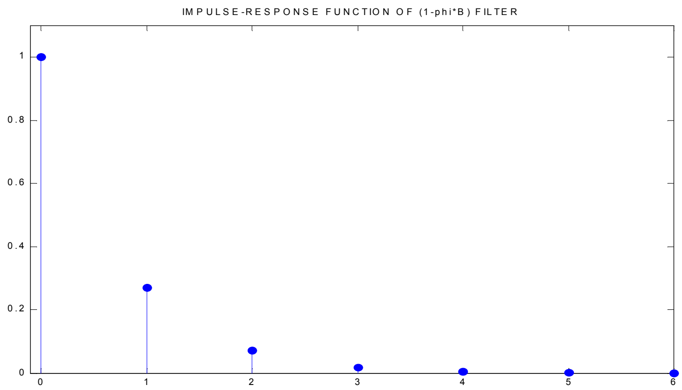
```
*************************************************
 TEMPORAL DISAGGREGATION METHOD: Santos Silva-Cardoso
 *************************************************
 -------------------------------------------------
 Number of low-frequency observations :   32
 Frequency conversion                 :    4
 Number of high-frequency observations:  128
 Number of extrapolations             :    0
 Number of indicators (+ constant)    :    2
 -------------------------------------------------
 Type of disaggregation: sum (flow).
 -------------------------------------------------
 Estimation method: Maximum likelihood.
 -------------------------------------------------
 Beta parameters (columnwise):
   * Estimate
   * Std. deviation
   * t-ratios
 -------------------------------------------------
    1.0946          3.7817          0.2895
    0.6718          0.0049        136.9983
 -------------------------------------------------
 Dynamic parameter:   0.2600
 -------------------------------------------------
 Long-run beta parameters (columnwise):
     1.4792
     0.9078
 -------------------------------------------------
 Truncation remainder: expected y(0):
   * Estimate
   * Std. deviation
   * t-ratios
 -------------------------------------------------
  310.3328         90.5351          3.4278
 -------------------------------------------------
 AIC:   5.2524
 BIC:   5.3898
 -------------------------------------------------
 Low-frequency correlation
  - levels     : 0.9994
  - yoy rates  : 0.8561
 -------------------------------------------------
 High-frequency correlation
  - levels     : 0.9993
  - yoy rates  : 0.8881
 -------------------------------------------------
 High-frequency volatility of yoy rates
  - estimate   : 2.0592
  - indicator  : 2.3430
  - ratio      : 0.8789
 -------------------------------------------------
```

[1] Los datos son los utilizados por Santos Silva y Cardoso, recopilados en Greene (1997), tabla 17.5.

Algunos de los gráficos generados se muestran a continuación:



Especialmente relevante para conocer la dinámica del modelo es la función de respuesta al impulso implicada por el filtro *(1-ϕB)*:



En determinadas ocasiones el analista puede estar interesado en aplicar esta función con un parámetro ϕ determinado a priori. En este caso la función se invoca de la siguiente forma:

```
res=ssc_fix(Y,x,ta,s,type,phi);
```

siendo *phi* un escalar que recoge el valor que el analista asigna al parámetro dinámico del modelo.

Referencias

Chow, G. y Lin, A.L. (1971) "Best linear unbiased distribution and extrapolation of economic time series by related series", *Review of Economic and Statistics*, vol. 53, n. 4, p. 372-375.

di Fonzo, T. (1987) "La stima indiretta di serie economiche trimestrali", Cleup Editore, Padua, Italia.

di Fonzo, T. (2002) "Temporal disaggregation of economic time series: towards a dynamic extension", Dipartimento di Scienze Statistiche, Università di Padova, Working Paper n. 2002-17 (disponible en http://www.stat.unipd.it/~difonzo).

Greene, W.H. (1997) *Econometric Analysis*, 3ª edición, Macmillan, New York, Estados Unidos de América.

Gregoir, S. (1994) "Propositions pour une désagrégation temporelle basée sur des modèles dynamiques simples", en Eurostat (Ed.) *Workshop on Quarterly National Accounts*, Eurostat, Luxembourg.

Quilis, E.M. (2001a) "Notas sobre desagregación temporal de series económicas", Instituto de Estudios Fiscales, Papeles de Trabajo n. 1/01 (disponible en http://www.minhac.es/ief).

Quilis, E.M. (2003) "A Matlab library of temporal disaggregation methods", Documento Interno, Instituto Nacional de Estadística (disponible en http://www.ine.es).

Salazar, E., Smith, R., Wright, S. y Weale, M. (1994) "Indicators of monthly national accounts", en Eurostat (Ed.) *Workshop on Quarterly National Accounts*, Eurostat, Luxembourg.

Santos Silva, J.M.C. y Cardoso, F. (2001) "The Chow-Lin method using dynamic models", *Economic Modelling*, vol. 18, p. 269-280.