

PROGRAM GAP\*- VERSION 2.2  
USER MANUAL  
May 2003

CHRISTOPHE PLANAS and ALESSANDRO ROSSI

*Emails: christophe.planas@jrc.it, alessandro.rossi@jrc.it,*

*European Commission, Joint Research Centre,  
Ispra, Italy.*

---

\*Please read disclaimer in worksheet "About" of the Excel interface.

# Contents

<b>1</b>	<b>Data input/output</b>	<b>3</b>
<b>2</b>	<b>Model specification</b>	<b>3</b>
2.1	General setting . . . . .	3
2.2	Model type . . . . .	5
2.3	Canonical trend . . . . .	5
2.4	AR orders . . . . .	5
2.5	MA order in Phillips curve equation . . . . .	6
2.6	Exogenous regressors . . . . .	6
2.7	Endogenous regressors in Phillips-curve equation . . . . .	6
<b>3</b>	<b>Estimation</b>	<b>7</b>
3.1	Simulating annealing parameters . . . . .	7
3.2	Recursive estimation . . . . .	8
<b>4</b>	<b>Parameter constraints</b>	<b>8</b>
<b>5</b>	<b>Output of the program</b>	<b>9</b>
<b>6</b>	<b>Graphs</b>	<b>9</b>

# 1 Data input/output

Before starting the analysis the user must store the data (1st, 2nd and exogenous series if any) in **worksheet "Data"** starting from the third row. Columns B and C are reserved to the endogenous variables. Ten regressors on the Phillips-curve equation can be introduced in columns D to P, while 3 other columns, Q to S, are reserved for the exogenous variable on the first variable, typically unemployment or GDP.

- Box (A): insert the location and name of the file where the 1st series will be stored, e.g.  $C : \backslash data \backslash filename1$
- Box (B): insert the location and name of the file where the 2nd series will be stored, e.g.  $C : \backslash data \backslash filename2$
- Box (C): insert the location and name of the file where the exogenous variable will be stored, e.g.  $C : \backslash data \backslash filename3$  (optional).
- Box (D): insert the location where all outputs will be stored, e.g.  $C : \backslash results$ .

*As the program does not make any type of data-transformation, users may wish to apply logarithm and/or difference operations to price series.*

## 2 Model specification

### 2.1 General setting

The most general model that can be specified is a bivariate model such that (see also Kuttner, 1994, *Journal of Business Economics and Statistics*, 12,3, pp 361-368). A short overview is developed below; full details can be read in the Technical Annex available in the worksheet Help.

**Equation 1:** The observed series is made up of a long term component or trend,  $X_{1t}^N$ , a short term component or cycle,  $X_{1t}^S$ , and a set of regressors according to:

$$X_{1t} = X_{1t}^N + X_{1t}^S + \sum_{i=1}^{Nex1} \alpha_{1i} Z_{it} \quad (1)$$

where  $Z_t$  is a vector of ( $Nex1 \times 1$ ) exogenous variables, with  $Nex1$  less or equal to 3. The trend can be specified as a second order random walk:

$$(1 - L) X_{1t}^N = \mu_{1t} + a_{Nt}. \quad (2)$$

where  $L$  is the lag operator. The slope is such that

$$(1 - L) \mu_{1t} = a_{\mu t}. \quad (3)$$

Both  $a_{Nt}$ , and  $a_{\mu t}$ , are *n iid*-innovations with variances  $V_N$  and  $V_\mu$ . Notice that if  $Var(a_{\mu t}) = 0$ , then (2) reduces to a random walk plus drift. The short-term component is such that:

$$(1 - \phi_1 L - \phi_2 L^2) X_{1t}^S = a_{St}. \quad (4)$$

The variables  $a_{St}$  are *n iid*-innovations with variance  $V_S$ .

**Equation 2:** in agreement with the Phillips curve framework, series 1 is related to a second series according to:

$$\begin{aligned} X_{2t} = & \mu_2 + \sum_{i=1}^{Nex2} \alpha_{2i} Y_{it} + \gamma (1 - L) X_{1t-1} + \sum_{i=0}^4 \beta_i X_{1t-i}^S \\ & + \phi_1^* X_{2t-1} + \phi_2^* X_{2t-2} + (1 + \theta_1 L + \theta_2 L^2 + \theta_3 L^3) a_{\pi t} \end{aligned} \quad (5)$$

where  $Y_t$  is a vector of ( $Nex2 \times 1$ ) exogenous variables, with  $Nex2$  less or equal to 10, and  $a_{\pi t}$  is a *n iid*-innovation with variance  $V_\pi$  that can be correlated with  $a_{St}$ . That correlation is denoted by  $\rho$ .

### Warnings:

1. the second series and the regressors on the second equation should be covariance stationary;
2. if  $X_{1t}^S$  is included in the Phillips curve equation (2nd equation), the innovations  $a_{\pi t}$  and  $a_{St}$  must be uncorrelated, so the program imposes  $\rho = 0$  when  $\beta_0 \neq 0$ . Without that restriction, the innovation  $a_{\pi t}$  and the regressor  $X_{1t}^S$  in the second equation would not be orthogonal and the corresponding estimator would not be consistent.

## 2.2 Model type

- The users have the possibility to choose among:
- Bivariate - AR + RW with drift (equations 1, 2, 3, 4 with restriction  $Var(a_{\mu t}) = 0$ ).
- Bivariate - AR + 2nd order RW (equations 1, 2, 3, 4).
- Univariate - AR + RW with drift (equations 1, 2, 3 with restriction  $Var(a_{\mu t}) = 0$ ).
- Univariate - AR + 2nd order RW (equations 1, 2, 3).
- Hodrick-Prescott filter. This is a facility for a univariate detrending of the first series. The box Hodrick-Prescott parameter allows users to enter the inverse signal to noise ratio.

## 2.3 Canonical trend

That specification is only possible when the trend is a random walk plus drift. The canonical trend is specified as:

$$(1 - L)X_{1t}^N = \mu_1 + (1 + L)a_{Nt}.$$

That specification removes all orthogonal noise from the long-term component and assigns it to the short-term component. Hence it can offer a solution to the case where the variance of the innovations on the trend level are estimated at zero. Notice that if the short-term component was described as a AR(2) cycle, it becomes an ARMA(2,2). If instead it was a noise, then it remains as such.

## 2.4 AR orders

Users are free to enter the autoregressive orders of equation (4) describing the first series cycle and the autoregressive order of the Phillips-curve (5). The values accepted are 0, 1 and 2.

## 2.5 MA order in Phillips curve equation

Users can enter the order of the MA term in (5). The values accepted are 0, 1, 2 and 3.

## 2.6 Exogenous regressors

The users must choice the number of exogenous regressors that will enter in equation (1) for the first series and in the Phillips-curve equation (5). These regressors must be loaded in the worksheet data - see section 1.

The maximum number of regressors accepted in equation (1) is 3 and is 10 for equation (5).

## 2.7 Endogenous regressors in Phillips-curve equation

For the endogenous terms in the right hand side of (5), users can choose among:

- Differenced. 1st series + 1st series cycle - 0 to 4 lags;
- Differenced. 1st series + 1st series cycle - 0 to 3 lags;
- Differenced. 1st series + 1st series cycle - 0 to 2 lags;
- Differenced. 1st series + 1st series cycle - 0 to 1 lag;
- Differenced. 1st series + 1st series cycle - 0 lag;
- Differenced. 1st series;
- 1st series cycle - 0 to 4 lags;
- 1st series cycle - 0 to 3 lags;
- 1st series cycle - 0 to 2 lags;
- 1st series cycle - 0 to 1 lag;
- 1st series cycle - 0 lag;
- None.

### 3 Estimation

Parameter estimation is made using the Diffuse Kalman Filter, in its collapsed version (see also Durbin and Koopman, 2001, Time Series Analysis by State Space Methods, Oxford University Press, pp.115-120). All details are given in the Technical Annex.

Likelihood maximisation can be made either via Simulating Annealing ("Estimate via simulated annealing") or via a Newton like optimization method ("Estimate via E04UCF mk. 19"). Estimation is performed when pressing **RUN**.

Because it explores the parameter space without considering the objective function gradient, the simulated annealing algorithm is less likely than the Newton-like algorithm to isolate a local maximum. The price to pay is that the simulated annealing algorithm needs much more computing time than the Newton-like technique. In most cases the gradient-based algorithm is satisfactory, and we would recommend the simulated annealing algorithm only in case of estimation problem.

#### 3.1 Simulating annealing parameters

If the simulated annealing algorithm is selected, then users can tune several parameters in the simulated annealing box. These parameters are:

- $T$  -The initial temperature.
- $RT$  - The temperature reduction factor. The suggested value .is 85.
- $EPS$  Error tolerance for termination.
- $NS$  - Number of cycles
- $NT$  - Number of iterations before temperature reduction. After  $NT \times NS \times Npar$  function evaluations, temperature ( $T$ ) is changed by the factor  $RT$ . The suggested value is  $\text{MAX}(100, 5 \times \text{Number of model parameters})$ .
- $MAXEVL$  - The maximum number of function evaluations.
- $VM$  - The initial step length vector.
- Intermediate results are produced if the user selects **Y** in the so-called box.

## 3.2 Recursive estimation

If 'Y' is selected in the cell next to "Rolling estimates", the program performs a recursive estimation of the model parameters starting at the Starting date that is entered by the user. By selecting 'Y' in the cell next to "Start at previous optimum" the maximization procedure starts from the previous maximum. Otherwise the starting values are obtained by a regression-type analysis. A file *PAR.TXT* is created in the output directory that displays on every line all model parameters estimates for a given sample size. The order of the parameters in the *PAR.TXT* file is:

$$\begin{aligned} &\phi_1, \phi_2, \mu_N, \mu_\pi, \gamma, \beta_1, \theta_1, \theta_2, \theta_3, V_N, V_S, \rho, V_\pi, \\ &\alpha_{2,1}, \dots, \alpha_{2,10}, \alpha_{1,1}, \dots, \alpha_{1,3}, \beta_2, \phi_1^*, \phi_2^*, \beta_3, \beta_4, \beta_0. \end{aligned}$$

A second file entitled *REV.TXT* is produced. It contains the smoothed estimates of the cycle computed in real-time, using the parameters in *PAR.TXT*. *REV.TXT* is a  $(t \times k)$  matrix where the rows ( $t = \text{Starting point}, \dots, \text{nobs}-11$ ), where  $t$  represents the period for which the component is estimated, and the columns are numerated as  $k = 1, 2, \dots, 12$ , with  $k$  such that  $k - 1$  is the number of additional observations available after time  $t$ . So concurrent estimates are on the first column and estimates with  $t + 11$  data are on the twelfth column.

## 4 Parameter constraints

For every parameter, the two boxes under the "**parameter constraints**" title offer the possibility to set lower and upper bounds. For the innovation variances, the lower and upper bounds are subject to several constraints.

- For the variance of the innovations in the first series unobserved components, upper bounds must be less or equal to 1.2 times the variance of the differenced first series. For the second equation residuals variance, the maximum upper bound allowed is 1.2 times the variance of the second series. In both cases, the minimum upper bound is 0, the lower bound. The program automatically resets the upper bound to the maximum authorized value if user enter a non-acceptable value.

## 5 Output of the program

When running, the program produces in the output directory (Box (D)) the following files:

- *SOL.TXT* contains the model parameters together with their standard deviations and -2 log likelihood plus the Ljung-Box statistics on the residuals. This file can be displayed by pressing EDIT RESULTS in worksheet "Graphs".
- *STATE.TXT* contains the TREND and CYCLE estimates for the first series. These estimates are those obtained via fixed-point smoother.
- *RESID.TXT* contains the residuals for the two equations.
- *RMSE.TXT* contains the Root Mean Square Errors of the smoothed unobserved component estimates.
- If "Recursive parameter estimates" is used, the program produces two additional file named *PAR.TXT* that contains the model parameters estimates for the different sample lengths, and *REV.TXT* containing smoothed estimates of the cycle computed by real-time parameters (in *PAR.TXT*).
- *PAR.TXT* contains the model parameters and -2log likelihood when the model is estimated recursively.

## 6 Graphs

When the program has run, you can see plots of the decomposition of the variable in the first equation by pushing UPDATE GRAPHS in the worksheet "Graph". Moreover users may select the magnitude of confidence bounds for the estimated short-term component by entering a number in cell **A3** in worksheet "Graphs". For instance a value of 1.96 corresponds to a 95% confidence bound.