

# 1 A user's guide to the subroutine

The syntax of the subroutine call is

```
@y2qrev(options) yseries yqseries qseries qresids start end1 end2  
# <list of indicators in regression format>
```

Where

**yseries** is the series to interpolate. This should be set up as a quarterly series. **y2q** takes the last quarter of each year as the annual value to be interpolated. Reading an annual series from a RATS format file, when the CALENDAR statement sets up a quarterly periodicity, produces this set-up

**yqseries** contains the known quarterly values of the **yseries** series.

**qseries** is the interpolated quarterly series. For the period from **end1** to **end2** the values in this series should be the same as the corresponding values of **yqseries**

**start end1** is the range to use in interpolation. These should refer to quarters 1 and 4 respectively.

**end2** Specifies the end of available quarterly data in **yqseries** and in the indicators.

The list of indicators may include constant and trend terms. Constant and trend terms may also be included as options in the procedure call. If they are included in the list of indicators, they will be used in the calculation of the Principal Component which is used to make the interpolating regression 'homoscedastic'. If they are specified in the option list they are added to the regression after the 'heteroscedasticity' adjustment has been completed. If there is a single indicator do not use the PC option in conjunction the CONST option. In this case specify a constant in your indicators and use the NOCONST option (if this is what you require).

The annual series should contain no missing data (#NA) over the interpolation range (from **start** to **end1**. All indicators must contain no #N/A from the first period of the start year (**start**) to the final period of the data set (**end2**). If an #N/A is found in either series the procedure halts and returns an error message. If **start** is not the first quarter of a year or **end1** is not the fourth quarter the procedure also halts and returns an error message.

The options available are as follows:

[SC] NOSC Adjusts for autocorrelation in the residuals of the indicator model. The default is a first order AR process.

PC [NOPC] Heteroscedasticity adjustment required using sole indicator or principal component if more than one indicator. The default is not to adjust.

CONST/[NOCONST] Add or do not add a constant to the regression after the 'homoscedastic' transformation. Default is not to add constant

TREND/[NOTREND] Add or do not add a trend to the regression after the 'homoscedastic' transformation. Default is not to add the trend

STARTRHO Initial value for  $\rho$  in maximum likelihood estimation. Default is 0.5. If the estimate does not converge for this value of  $\rho$  you may need to try a larger value. Some initial values may cause the program to exit with a message that some matrix is not invertible

The procedure includes an amended version of the principal components procedure distributed with RATS

## Example

```
OPEN DATA xq.rat
CALENDAR 1981 1 4
ALL 2004:04
DATA(FORMAT=RATS) / INTERCEPT X1 X2
OPEN DATA ya.rat
data(format=rats) 1981:1 1997:4 ya
OPEN DATA yq.rat
data(format=rats) 1998:1 2004:04 yq
*print 1981:1 2004:4 intercept x1 x2 ya yq
* quarterly indicators are intercept x1 and x2
*
* annual series to be interpolated is ya and goes from 1981 to 1997
* annual value is repeated in each quarter
*
```

```

* Known quarterly series is yq
*
source y2qrev.src
@y2qrev ya yq yinterp yresid 1981:1 1997:4 2004:4
# intercept x1 x2
*
* yinterp is the interpolated series
* yresid are the residue in the interpolation model corresponding
* to the period for which quarterly data are used in the estimation
* of that model.

```

The first eight statements load the data and indicators. The source statement will load the subroutine into memory for further use. The statement

```

@y2qrev ya yq yinterp yresid 1981:1 1997:4 2004:4
# intercept x1 x2

```

runs the procedure. `ya` contains the annual data to be distributed to quarters. `yq` contains known quarterly data for the series. `yresid` may be used to examine the suitability of the model

Examples of the use of the other options are given in Frain(2004). The output produced by the program consists of

1. Progress of the maximization process used in estimating the autocorrelation coefficient. If the system fails to converge this may indicate a possible new starting value for  $\rho$  (The auto-correlation coefficient in the quarterly model). It may also imply that the indicator model is very bad and that the indicators are not indicating.
2. Estimates of the parameters of the indicator model

## 2 A description of the distribution/interpolation method

Let  $\mathbf{y}$  be a  $((4m + k) \times 1)$  vector of the quarterly series. Only annual observations are available for the first  $m$  years of the series. Let  $\mathbf{X}$  be a  $((4m + k) \times p)$  matrix of related series. Each column of this matrix represents a quarterly data series which will serve as an indicator for the distribution/interpolation of the annual series. We assume that there is a multivariate regression type of relationship between  $\mathbf{y}$  and  $\mathbf{X}$  of the form.

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon} \quad (1)$$

where

$$E[\boldsymbol{\varepsilon}] = 0 \text{ and } E[\boldsymbol{\varepsilon}\boldsymbol{\varepsilon}'] = \boldsymbol{\Omega}$$

Let  $\mathbf{C}_m$  be the  $(m \times 4m)$  matrix

$$\begin{pmatrix} 1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 & \dots & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 & \dots & 0 & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \dots & 1 & 1 & 1 & 1 \end{pmatrix} \quad (2)$$

Let  $\mathbf{I}_k$  be the  $k$ -dimension Identity Matrix.

$\mathbf{C}$  be the block-diagonal matrix.

$$\begin{pmatrix} \mathbf{C}_m, \mathbf{0} \\ \mathbf{0}, \mathbf{I}_k \end{pmatrix} \quad (3)$$

Let

$$\mathbf{y}_d = \mathbf{C}\mathbf{y} \text{ and } \mathbf{X}_d = \mathbf{C}\mathbf{X}$$

Then  $\mathbf{y}_d$  is a  $((m + k) \times 1)$  vector which contains the  $m$  known annual observations on the national accounts variable stacked on the  $k$  known quarterly observations.  $\mathbf{X}_d$  is a corresponding stacked version of the indicator variable matrix.

Let  $\mathbf{A}$  be matrix such that  $\hat{\mathbf{y}} = \mathbf{A}\mathbf{y}_d$  is a linear unbiased estimator of the quarterly series. Premultiplying equation (1) by  $\mathbf{C}$  we get

$$\begin{aligned}
\mathbf{y}_d &= \mathbf{C}\mathbf{y} \\
&= \mathbf{C}\mathbf{X}\boldsymbol{\beta} + \mathbf{C}\boldsymbol{\varepsilon} \\
&= \mathbf{X}_d\boldsymbol{\beta} + \boldsymbol{\varepsilon}_d
\end{aligned} \tag{4}$$

where

$$E[\boldsymbol{\varepsilon}_d] = 0 \text{ and } E[\boldsymbol{\varepsilon}_d\boldsymbol{\varepsilon}_d'] = \mathbf{C}\boldsymbol{\Omega}\mathbf{C}'$$

Thus

$$\begin{aligned}
\hat{\mathbf{y}} - \mathbf{y} &= \mathbf{A}\mathbf{y}_d - \mathbf{y} \\
&= \mathbf{A}(\mathbf{X}_d\boldsymbol{\beta} + \boldsymbol{\varepsilon}_d) - (\mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon}) \\
&= (\mathbf{A}\mathbf{X}_d - \mathbf{X})\boldsymbol{\beta} + (\mathbf{A}\boldsymbol{\varepsilon}_d - \boldsymbol{\varepsilon})
\end{aligned} \tag{5}$$

If  $\hat{\mathbf{y}}$  is an unbiased estimator of  $\mathbf{y}$  then  $E[\hat{\mathbf{y}} - \mathbf{y}] = 0$ . Taking expectations of equation (5) gives

$$E[\hat{\mathbf{y}} - \mathbf{y}] = (\mathbf{A}\mathbf{X}_d - \mathbf{X})\boldsymbol{\beta} = 0 \tag{6}$$

This represents  $(4n + k)$  equations in the  $p$  unknowns. Thus we must have

$$\mathbf{A}\mathbf{X}_d - \mathbf{X} = 0 \tag{7}$$

The covariance matrix of  $(\hat{\mathbf{y}} - \mathbf{y})$  is given by

$$\begin{aligned}
\text{cov}(\hat{\mathbf{y}} - \mathbf{y}) &= E[(\mathbf{A}\boldsymbol{\varepsilon}_d - \boldsymbol{\varepsilon})(\mathbf{A}\boldsymbol{\varepsilon}_d - \boldsymbol{\varepsilon})'] \\
&= E[\mathbf{A}\boldsymbol{\varepsilon}_d\boldsymbol{\varepsilon}_d'\mathbf{A}' - \boldsymbol{\varepsilon}\boldsymbol{\varepsilon}_d'\mathbf{A}' - \mathbf{A}\boldsymbol{\varepsilon}_d\boldsymbol{\varepsilon}' + \boldsymbol{\varepsilon}\boldsymbol{\varepsilon}'] \\
&= \mathbf{A}\mathbf{C}\boldsymbol{\Omega}\mathbf{C}'\mathbf{A}' - \boldsymbol{\Omega}\mathbf{C}'\mathbf{A}' - \mathbf{A}\mathbf{C}\boldsymbol{\Omega} + \boldsymbol{\Omega}
\end{aligned} \tag{8}$$

The diagonal elements of this covariance matrix represent the variance of the corresponding element of  $(\hat{\mathbf{y}} - \mathbf{y})$ . The sum of these elements or the trace of this covariance matrix is a multiple of the variance of the estimator. Thus the minimum variance unbiased estimator is that choice of  $\mathbf{A}$  which minimises the trace of this covariance matrix subject to the constraints given by equation (7). We may write the Lagrangian for this constrained optimisation problem as—

$$\mathcal{L} = \frac{1}{2}\text{trace}(\mathbf{A}\mathbf{C}\boldsymbol{\Omega}\mathbf{C}'\mathbf{A}' - \boldsymbol{\Omega}\mathbf{C}'\mathbf{A}' - \mathbf{A}\mathbf{C}\boldsymbol{\Omega} + \boldsymbol{\Omega}) - \text{trace}(\mathbf{M}'(\mathbf{A}\mathbf{X}_d - \mathbf{X})) \tag{9}$$

where  $\mathbf{M}$  is a  $((4m + k) \times p)$  matrix of Lagrange multipliers.

Differentiating (9) with respect to the elements of  $\mathbf{A}$  and equating the derivatives to zero for a minimum we get—

$$\mathbf{A}\mathbf{C}\boldsymbol{\Omega}\mathbf{C}' - \boldsymbol{\Omega}\mathbf{C} - \mathbf{M}\mathbf{X}'_d = 0 \quad (10)$$

Thus

$$\mathbf{A} = \boldsymbol{\Omega}\mathbf{C}'(\mathbf{C}\boldsymbol{\Omega}\mathbf{C}')^{-1} + \mathbf{M}\mathbf{X}'_d(\mathbf{C}\boldsymbol{\Omega}\mathbf{C}')^{-1} \quad (11)$$

Substitute (11) into (7) to get

$$\boldsymbol{\Omega}\mathbf{C}'(\mathbf{C}\boldsymbol{\Omega}\mathbf{C}')^{-1}\mathbf{X}_d + \mathbf{M}[\mathbf{X}'_d(\mathbf{C}\boldsymbol{\Omega}\mathbf{C}')^{-1}\mathbf{X}_d] - \mathbf{X} = 0 \quad (12)$$

Thus

$$\mathbf{M} = \mathbf{X}[\mathbf{X}'_d(\mathbf{C}\boldsymbol{\Omega}\mathbf{C}')^{-1}\mathbf{X}_d]^{-1} - \boldsymbol{\Omega}\mathbf{C}'(\mathbf{C}\boldsymbol{\Omega}\mathbf{C}')^{-1}\mathbf{X}_d[\mathbf{X}'_d(\mathbf{C}\boldsymbol{\Omega}\mathbf{C}')^{-1}\mathbf{X}_d]^{-1} \quad (13)$$

Substitute (13) into (11) to get —

$$\begin{aligned} \mathbf{A} &= \boldsymbol{\Omega}\mathbf{C}'(\mathbf{C}\boldsymbol{\Omega}\mathbf{C}')^{-1} + \mathbf{X}[\mathbf{X}'_d(\mathbf{C}\boldsymbol{\Omega}\mathbf{C}')^{-1}\mathbf{X}_d]^{-1}\mathbf{X}'_d(\mathbf{C}\boldsymbol{\Omega}\mathbf{C}')^{-1} \\ &\quad - \boldsymbol{\Omega}\mathbf{C}'(\mathbf{C}\boldsymbol{\Omega}\mathbf{C}')^{-1}\mathbf{X}_d[\mathbf{X}'_d(\mathbf{C}\boldsymbol{\Omega}\mathbf{C}')^{-1}\mathbf{X}_d]^{-1}\mathbf{X}'_d(\mathbf{C}\boldsymbol{\Omega}\mathbf{C}')^{-1} \\ &= \mathbf{X}[\mathbf{X}'_d(\mathbf{C}\boldsymbol{\Omega}\mathbf{C}')^{-1}\mathbf{X}_d]^{-1}\mathbf{X}'_d(\mathbf{C}\boldsymbol{\Omega}\mathbf{C}')^{-1} \\ &\quad + \boldsymbol{\Omega}\mathbf{C}'(\mathbf{C}\boldsymbol{\Omega}\mathbf{C}')^{-1}[I - [\mathbf{X}'_d(\mathbf{C}\boldsymbol{\Omega}\mathbf{C}')^{-1}\mathbf{X}_d]^{-1}\mathbf{X}'_d(\mathbf{C}\boldsymbol{\Omega}\mathbf{C}')^{-1}] \end{aligned} \quad (14)$$

The resulting estimator is now

$$\begin{aligned} \hat{\mathbf{y}} &= \mathbf{A}\mathbf{y}_d \\ &= \mathbf{X}\hat{\boldsymbol{\beta}} + \boldsymbol{\Omega}\mathbf{C}'(\mathbf{C}\boldsymbol{\Omega}\mathbf{C}')^{-1}\hat{\mathbf{u}}_d \end{aligned} \quad (15)$$

where

$$\hat{\boldsymbol{\beta}} = [\mathbf{X}'_d(\mathbf{C}\boldsymbol{\Omega}\mathbf{C}')^{-1}\mathbf{X}_d]^{-1}\mathbf{X}'_d(\mathbf{C}\boldsymbol{\Omega}\mathbf{C}')^{-1}\mathbf{y}_d \quad (16)$$

is the *GLS* estimator in the regression of  $\mathbf{y}_d$  on  $\mathbf{X}_d$  and

$$\hat{\mathbf{u}}_d = \mathbf{y}_d - \mathbf{X}[\mathbf{X}'_d(\mathbf{C}\boldsymbol{\Omega}\mathbf{C}')^{-1}\mathbf{X}_d]^{-1}\mathbf{X}'_d(\mathbf{C}\boldsymbol{\Omega}\mathbf{C}')^{-1}\mathbf{y}_d \quad (17)$$

are the residuals in this regression.

We may summarise the result as follows

1. Estimate the *GLS* estimates of the coefficients in the annual regression of  $\mathbf{y}_a$  on  $\mathbf{X}_a$ . The distributed series is the sum of two parts

2. The first part is found by applying the estimated regression coefficients to the quarterly indicators
3. The second part is found by distributing the residual in the annual regression over the quarters. This ensures that the distributed series sum to the known annual series. Pre-multiplying equation 15 by  $\mathbf{C}$  shows that  $\mathbf{y}_d = \mathbf{y}$  for values of  $\mathbf{y}$  for which the quarterly data were used.

#### **Extracted from RT/2/2004**

The implementation of this procedure involves two practical problems. The first is the choice of indicators. This must be dealt with on a pragmatic basis for each series. The second problem is that the covariance matrix  $\mathbf{\Omega}$  is, in general, unknown. Obviously we can not estimate all  $((4m + k) \times (4m + k))$  elements in  $\mathbf{\Omega}$ . We must assume some structure for  $\mathbf{\Omega}$ . Four cases are provided in the program. Cases (1) to (3) below are set out in Chow and Lin (1971). In previous distribution/interpolation exercises either a test was carried out for the significance of autocorrelation in regression of the annual variables. If this was not significant case (1) was used. If significant autocorrelation was found then case (3) was used.

In our practical applications of the methodology the estimated equations produced very poor out of sample forecasts of the distributed/interpolated variables. The autocorrelation found in the estimations is likely to be due to miss-specification of the variable-indicator model. We are also concerned that the the assumption of homoscedasticity may cause a loading of errors at the start of the sample in the type of regressions estimated for the model.

Case 4 implements an extension to cater for both autocorrelation and heteroscedasticity. The program estimates the unknown parameters in the model using pseudo maximum likelihood rather than the iterative procedure suggested by Chow and Lin (1971). The procedure can be extended to account for various other autocorrelation structures. A maximum likelihood estimate of the other cases can also be completed by an appropriate choice of options in running the procedure.

**Case 1 :  $\varepsilon$  are uncorrelated variables with constant variance**

In this case

$$\Omega C'(C\Omega C')^{-1} = \begin{pmatrix} \frac{1}{4} & 0 & \dots & 0 \\ \frac{1}{4} & 0 & \dots & 0 \\ \frac{1}{4} & 0 & \dots & 0 \\ \frac{1}{4} & 0 & \dots & 0 \\ 0 & \frac{1}{4} & \dots & 0 \\ 0 & \frac{1}{4} & \dots & 0 \\ 0 & \frac{1}{4} & \dots & 0 \\ 0 & \frac{1}{4} & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \frac{1}{4} \\ 0 & 0 & \dots & \frac{1}{4} \\ 0 & 0 & \dots & \frac{1}{4} \\ 0 & 0 & \dots & \frac{1}{4} \end{pmatrix}$$

$\beta$  is estimated by *OLS* and the residual is divided equally between the four quarters.



**Case 2 :  $\varepsilon$  are uncorrelated variables but are heteroscedastic**

In this case

$$\Omega = \begin{pmatrix} \omega_1 & 0 & \dots & 0 \\ 0 & \omega_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \omega_{4n} \end{pmatrix}$$

and

$$\Omega C'(C\Omega C')^{-1} = \begin{pmatrix} \frac{\omega_1}{\omega_1 + \omega_2 + \omega_3 \omega_4} & 0 & \dots & 0 \\ \frac{\omega_2}{\omega_1 + \omega_2 + \omega_3 \omega_4} & 0 & \dots & 0 \\ \frac{\omega_3}{\omega_1 + \omega_2 + \omega_3 \omega_4} & 0 & \dots & 0 \\ \frac{\omega_4}{\omega_1 + \omega_2 + \omega_3 \omega_4} & 0 & \dots & 0 \\ 0 & \frac{\omega_5}{\omega_5 + \omega_6 + \omega_7 \omega_8} & \dots & 0 \\ 0 & \frac{\omega_6}{\omega_5 + \omega_6 + \omega_7 \omega_8} & \dots & 0 \\ 0 & \frac{\omega_7}{\omega_5 + \omega_6 + \omega_7 \omega_8} & \dots & 0 \\ 0 & \frac{\omega_8}{\omega_5 + \omega_6 + \omega_7 \omega_8} & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \frac{\omega_{4n-3}}{\omega_{4n-3} + \omega_{4n-2} + \omega_{4n-1} \omega_{n4}} \\ 0 & 0 & \dots & \frac{\omega_{4n-2}}{\omega_{4n-3} + \omega_{4n-2} + \omega_{4n-1} \omega_{n4}} \\ 0 & 0 & \dots & \frac{\omega_{4n-1}}{\omega_{4n-3} + \omega_{4n-2} + \omega_{4n-1} \omega_{n4}} \\ 0 & 0 & \dots & \frac{\omega_{4n}}{\omega_{4n-3} + \omega_{4n-2} + \omega_{4n-1} \omega_{n4}} \end{pmatrix}$$

In this case  $\beta$  is estimated by *GLS* and the residuals are distributed in proportion to the variance in the quarter. The variance may be taken as proportional to a function of the indicator variable(s) or perhaps of the principal component of the indicator variables.

**Case 3 : Innovations in  $\varepsilon$  are homoscedastic but  $\varepsilon$  follows an  $AR(1)$  process**

In this case we assume that the quarterly residuals follow an  $AR(1)$  process. In this case

$$\mathbf{\Omega} = \frac{\sigma^2}{1 - \rho^2} \begin{pmatrix} 1 & \rho & \rho^2 & \dots & \rho^{4n-1} \\ \rho & 1 & \rho & \dots & \rho^{4n-2} \\ \rho^2 & \rho & 1 & \dots & \rho^{4n-3} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \rho^{4n-1} & \rho^{4n-2} & \rho^{4n-3} & \dots & 1 \end{pmatrix}$$

The covariance of the annual residuals is then given by  $\mathbf{C}\mathbf{\Omega}\mathbf{C}'$ . It can be shown that the ratio of element (1,2) to element (1,1) in this covariance matrix is given by —

$$\alpha = \frac{\rho + 2\rho^2 + 3\rho^3 + 4\rho^4 + 3\rho^5 + 2\rho^6 + \rho^7}{4 + 6\rho + 4\rho^2 + 2\rho^3}$$

Taking  $\alpha$  as a function of  $\rho$

$$\alpha(0) = 0$$

$$\alpha(1) = 1 \text{ and}$$

$$\alpha(\rho) > 0 \text{ for } 0 < \rho \leq 1.$$

Thus for a particular value of  $\alpha$  in the range (0, 1) there corresponds a unique value of  $\rho$ .

When  $-0.130544 < \alpha] < 0$  there are two values of  $\rho$  in the range  $(-0.680297, 0)$  for each value of  $\alpha$ . For values of  $\alpha$  in the range  $(-1, -0.130544)$  there is no real solution to the equation. Thus this methodology is not valid for the distribution of quarterly data when  $\alpha$  is negative. In previous implementations of this procedure we found no negative values of  $\alpha$

**Case 4 : Innovations in  $\varepsilon$  are heteroscedastic and “pseudo AR(1)” process**

This Special Case is a combination of 3.2 and 3.3. We may write

$$\begin{aligned} \Omega &= \mathbf{WVW} \\ &= \frac{1}{1-\rho^2} \begin{pmatrix} \omega_1 & 0 & \dots & 0 \\ 0 & \omega_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \omega_{4n} \end{pmatrix} \times \\ &\quad \begin{pmatrix} 1 & \rho & \rho^2 & \dots & \rho^{4n-1} \\ \rho & 1 & \rho & \dots & \rho^{4n-2} \\ \rho^2 & \rho & 1 & \dots & \rho^{4n-3} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \rho^{4n-1} & \rho^{4n-2} & \rho^{4n-3} & \dots & 1 \end{pmatrix} \begin{pmatrix} \omega_1 & 0 & \dots & 0 \\ 0 & \omega_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \omega_{4n} \end{pmatrix} \end{aligned}$$

where  $\omega_1, \omega_2, \dots, \omega_{4n}$  are the square roots of the variances of the quarterly regressions and  $\rho$  is the first order auto-correlation of the quarterly residuals. This methodology may be implemented using the following steps:

1. Estimate the  $\omega$ 's using an appropriate function of the indicator variable(s) (e.g. proportional to a principal component of the indicator variable – this is the assumption made in the subroutine). This determines  $\mathbf{W}$  up to a factor  $\sigma$ .
2. If the annual variable  $\mathbf{y}_a$  is multivariate normal with mean  $\mathbf{CX}\beta$  and variance  $\mathbf{C}'\mathbf{W}'\mathbf{V}\mathbf{W}\mathbf{C}$  the log of the likelihood is given by

$$\begin{aligned} & -\frac{N}{2}\text{Log}(2\pi) - \frac{1}{2}\text{Log}(|\mathbf{C}'\mathbf{W}'\mathbf{V}\mathbf{W}\mathbf{C}|) \\ & - \frac{1}{2}(\mathbf{y}_a - \mathbf{CX}\beta)'(\mathbf{X}'_a(\mathbf{C}'\mathbf{W}'\mathbf{V}\mathbf{W}\mathbf{C})^{-1}\mathbf{X}_a)^{-1}(\mathbf{y}_a - \mathbf{CX}\beta) \end{aligned}$$

3. For a given value of  $\rho$  the maximum of this expression occurs when

$$\hat{\beta} = (\mathbf{X}'_a(\mathbf{C}'\mathbf{W}'\mathbf{V}\mathbf{W}\mathbf{C})^{-1}\mathbf{X}_a)^{-1}\mathbf{X}'_a(\mathbf{C}'\mathbf{W}'\mathbf{V}\mathbf{W}\mathbf{C})^{-1}\mathbf{y}_a$$

and

$$\hat{\sigma}^2 = \frac{1}{N}(\mathbf{y}_a - \mathbf{CX}\hat{\beta})'(\mathbf{X}'_a(\mathbf{C}'\mathbf{W}'\mathbf{V}\mathbf{W}\mathbf{C})^{-1}\mathbf{X}_a)^{-1}(\mathbf{y}_a - \mathbf{CX}\hat{\beta})$$

and the maximum value is

$$-\frac{N}{2}(\log(2\pi) + 1) - \frac{N}{2}\text{Log}(\hat{\sigma}^2) - \frac{1}{2}\text{Log}(|\mathbf{C}'\mathbf{C}\mathbf{C}|)$$

The FIND command in RATS is then used to maximise this expression with respect to  $\rho$ .

4. Use the results at (3) to distribute  $\mathbf{y}_a$  between quarters using equations 15 to 17 above with  $\mathbf{\Omega}$  replaced by  $\mathbf{W}'\mathbf{V}\mathbf{W}$

One can verify that this solution satisfies the constraints by premultiplying (9) by  $\mathbf{C}$