**TMD DISCUSSION PAPER NO. 64**

# GAMS CODE FOR ESTIMATING A SOCIAL ACCOUNTING MATRIX (SAM) USING CROSS ENTROPY (CE) METHODS

**Sherman Robinson**
**and**
**Moataz El-Said**
**International Food Policy Research Institute**

**Trade and Macroeconomics Division**
**International Food Policy Research Institute**
**2033 K Street, N.W.**
**Washington, D.C. 20006, U.S.A.**

**December 2000**

*TMD Discussion Papers contain preliminary material and research results, and are circulated prior to a full peer review in order to stimulate discussion and critical comment. It is expected that most Discussion Papers will eventually be published in some other form, and that their content may also be revised. This paper is available at* ***http://www.cgiar.org/ifpri/divs/tmd/dp.htm***

# GAMS Code for Estimating A Social Accounting Matrix (SAM) Using Cross Entropy (CE) Methods

**by**

**Sherman Robinson**
**and**
**Moataz El-Said**
**International Food Policy Research Institute**

**Abstract**

This paper documents the computer code implementing the CE-SAM estimation technique in GAMS (General Algebraic Modeling System). It defines the estimation problem in a deterministic setting; extends the approach to include a stochastic treatment of errors in control totals; summarizes the equations describing CE technique for estimating a consistent SAM starting from an inconsistent data set estimated with error; and provides the GAMS code.

Key words: Entropy, cross entropy, social accounting matrices, SAM, GAMS

**Table of Contents**

## 1. Introduction[*]

The paper describes the cross entropy (CE) SAM estimation technique in situations where column sums and macro aggregates represent SAM control totals that may be measured with error. The theory of the estimation technique, comparing it to other methods, is described in S. Robinson, A. Cattaneo, and M. El Said "Updating and Estimating a Social Accounting Matrix Using Cross Entropy Methods" in *Economic Systems Research*, vol. 13, no. 1, March 2001.[1] This paper documents the computer code implementing the technique in GAMS (General Algebraic Modeling System) (Brooke, Kendrick, Meeraus, and R. Raman 1998). Section 2 defines the estimation problem in a deterministic setting. Section 3 extends the approach to include a stochastic treatment of errors in the control totals. Section 4 summarizes the equations describing the CE technique for estimating a consistent SAM starting from an inconsistent data set estimated with error. Finally, section 5 provides the GAMS code for the estimation problem both as a nonlinear programming (NLP) problem and a mixed complementarity problem (MCP).

## 2. CE-SAM Estimation: Deterministic Approach

Define $T$ as a matrix of SAM transactions, where $t_{i,j}$ is a payment from column account $j$ to row account $i$, that satisfies the condition:

$$y_i = \sum_j t_{i,j} = \sum_j t_{j,i} \qquad (1)$$

That is, for a SAM, every row sum must equal the corresponding column sum. A SAM coefficient matrix, $A$, is constructed from $T$ by dividing the cells in each column of $T$ by the column sums:

$$A_{i,j} = \frac{t_{i,j}}{y_j} \qquad (2)$$

[*] The authors would like to thank Channing Arndt, Rebecca Harris, Henning Tarp Jensen, and Finn Tarp for helpful comments on the GAMS code. Also we would like to thank Michael Ferris for providing the MCP version of the estimation problem.

[1] An earlier version of the paper can be downloaded in PDF format from the IFPRI web page "http://www.ifpri.cgiar.org/divs/tmd/tmdpubs.htm#dp" division discussion paper No. 58, August 2000.

We assume that we start with information in the form of a prior, $\bar{A}$, which may be based on data from a previous or from scattered, perhaps inconsistent, data from the current year. We also assume that we have exact information on current column sums, $y^*$. Applying the Kullback-Leibler (1951) measure of the "cross entropy" (CE) distance between two probability distributions to the CE-SAM estimation, the problem is to find a new set of $A$ coefficients which minimize the cross entropy distance between the prior $\bar{A}$ and the new estimated coefficient matrix.

$$\begin{aligned} \min_{\{A\}} I &= \left[ \sum_i \sum_j A_{i,j} \ln \frac{A_{i,j}}{\bar{A}_{i,j}} \right] \\ &= \left[ \sum_i \sum_j A_{i,j} \ln A_{i,j} - \sum_i \sum_j A_{i,j} \ln \bar{A} \right] \end{aligned} \quad (3)$$

Subject to:

$$\sum_j A_{i,j} y_j^* = y_i^* \quad (4)$$

$$\sum_j A_{j,i} = 1 \text{ and } 0 \le A_{j,i} \le 1 \quad (5)$$

Note that $x \ln x = 0$ if $x = 0$. Thus, to solve for equation (3) and allow for zero entries in the SAM in the computer code, we add an epsilon small number to the arguments of the equation. Note also that the system of constraint equations (4) is functionally dependent, since if all but one column and row sum are equal, the last one must also be equal (analogous to Walras' in general equilibrium theory). One equation can be dropped.

## 3. CE-SAM Estimation: Stochastic Approach

Specifying known column sums implies having exact information about control totals in the SAM. Most applications of economic models to real world issues must deal with the problem of extracting results from data or economic relationships with noise. One can generalize to include knowledge about any aggregates or elements of the SAM (e.g., macro aggregates from the national accounts). In this section we generalize our approach to cases where: (i) row and column sums are not fixed parameters but involve errors in measurement; and (ii) macro aggregates are not exact but are measured with error.

The general case starts from assumed prior knowledge of the standard error (perhaps due to measurement error) of the estimate of control totals—a Bayesian prior, not a maintained hypothesis. The estimated error in the $i^{th}$ control total can be represented as a weighted sum of elements in a specified error support set:

$$e_i = \sum_{jwt} w_{i,jwt} \bar{v}_{i,jwt} \tag{6}$$

where $e_i$ = error value of control total

$w_{i,jwt}$ = error weights estimated in the CE procedure ( $\sum_{jwt} w_{i,jwt} = 1$ )

$\bar{v}_{i,jwt}$ = error support set

The set $jwt$ defines the dimension of the support set for the error distribution and the number of weights that must be estimated for each error. The prior on the variance of these errors is given by:

$$s^2 = \sum_{jwt} \bar{w}_{i,jwt} \cdot \bar{v}^2_{i,jwt} \tag{7}$$

where $\bar{w}_{i,jwt}$ = prior weights on the error support set and $\sum_{jwt} \bar{w}_{i,jwt} = 1$

Starting with a prior $s$ , Golan, Judge, and Miller (1996) suggest picking the $\bar{v}$ s to define a domain for the support set of $\pm 3$ standard errors. In this case, the prior on the weights, $\bar{w}$, are then calculated to yield a consistent prior on the standard error, $s$ .[2]

*3.1. Case of three-weight error distribution*

Assume a prior mean of zero and a given standard error, $s$ . With a three-parameter error distribution that is symmetric around zero, the $\bar{v}$ s define the upper and lower bounds for the error distribution, and there are three weights, $\bar{w}$, to be estimated. That is we have:

$$\begin{aligned} \bar{v}_{i,1} &= -3\,s \\ \bar{v}_{i,2} &= \quad 0 \\ \bar{v}_{i,3} &= +3\,s \end{aligned} \tag{8}$$

and using (7):

[2] In Robinson, Cattaneo, and El-Said (2001), we specify prior weights $\bar{w}$ that are uniform and set the prior standard error by the choice of support set, $(\bar{v})$. In that paper, we use a three-weight specification (jwt = {1,2,3}).

$$s^2 = \bar{w}_{i,1} \cdot \left(+9s^2\right) + \bar{w}_{i,2} \cdot \left(0\right) + \bar{w}_{i,3} \cdot \left(9s^2\right) \tag{9}$$

Since the prior weights and support set are symmetric, $\bar{w}_{i,1} = \bar{w}_{i,3}$. Solving (9) for the weights, $\bar{w}$, we get:

$$\begin{aligned} \bar{w}_{i,1} &= \bar{w}_{i,3} = \frac{1}{18} \\ \bar{w}_{i,2} &= 1 - \bar{w}_{i,1} - \bar{w}_{i,3} = \frac{16}{18} \end{aligned} \tag{10}$$

*3.2. Case of five-weight error distribution*

For the case of a five-parameter error distribution, there are five weights, $\bar{w}$, to be estimated—the set jwt consists of five elements. We are incorporating more information about the error distribution—more moments, including variance, skewness, and kurtosis. Assuming a prior normal distribution with mean of zero and variance $s^2$, the prior on kurtosis is $3s^4$. In this case, the prior weights, $\bar{w}$, are specified so that:

$$\sum_{jwt} \bar{w}_{i,jwt} \cdot \bar{v}_{i,jwt}^4 = 3s^4 \tag{11}$$

in addition to defining the variance as above in (7). The prior weights and support set are also symmetric, so the prior on all odd moments is zero. Choose $\pm 1$ standard error for $\bar{v}_{i,2}$ and $\bar{v}_{i,4}$ (which is arbitrary). In this case we get:

$$\begin{aligned} \bar{v}_{i,1} &= -3s \\ \bar{v}_{i,2} &= -s \\ \bar{v}_{i,3} &= \ \ 0 \\ \bar{v}_{i,4} &= +s \\ \bar{v}_{i,5} &= +3s \end{aligned} \tag{12}$$

and using (7) and (12) we get:

$$\begin{aligned} s^2 &= \bar{w}_{i,1} \cdot \left(9s^2\right) + \bar{w}_{i,2} \cdot \left(s^2\right) + \bar{w}_{i,3} \cdot \left(0\right) + \bar{w}_{i,4} \cdot \left(s^2\right) + \bar{w}_{i,5} \cdot \left(9s^2\right) \\ 3s^4 &= \bar{w}_{i,1} \cdot \left(81s^4\right) + \bar{w}_{i,2} \cdot \left(s^4\right) + \bar{w}_{i,3} \cdot \left(0\right) + \bar{w}_{i,4} \cdot \left(s^4\right) + \bar{w}_{i,5} \cdot \left(81s^4\right) \end{aligned} \tag{13}$$

given that $\bar{w}_{i,1} = \bar{w}_{i,5}$ and $\bar{w}_{i,2} = \bar{w}_{i,4}$ by symmetry, we get

$$\begin{aligned} 18\bar{w}_{i,1} + 2\bar{w}_{i,2} &= 1 \\ 162\bar{w}_{i,1} + 2\bar{w}_{i,2} &= 3 \end{aligned}$$

solving (13) for the $\overline{w}$s

$$
\begin{aligned}
\overline{w}_{i,1} &= \frac{1}{72} \\
\overline{w}_{i,2} &= \frac{27}{72} \\
\overline{w}_{i,3} &= \frac{16}{72} \\
\overline{w}_{i,4} &= \frac{27}{72} \\
\overline{w}_{i,5} &= \frac{1}{72}
\end{aligned}
\tag{14}
$$

Note that all these parameters determine the prior distribution. The estimation procedure yields posterior estimates of all the moments of the error distribution (Golan, Judge, and Miller, 1996). The five parameter specification permits posterior estimation of four moments; mean, variance, skewness, and kurtosis.

## 4. Equations of the CE-SAM Estimation Problem

In this section we provide a mathematical statement of the equations involved in the CE-SAM estimation problem. The statement includes the stochastic formulation to specify errors on column sums, and errors on macro aggregates. In this case we specify two sets of errors with separate weights, $W1$'s and $W2$'s, and extend the CE minimand in equation (3) to account for the specification of the error terms as follows:

$$
\begin{aligned}
\min_{\{A, W1, W2\}} I = &\left[ \sum_i \sum_j A_{i,j} \ln A_{i,j} - \sum_i \sum_j A_{i,j} \ln \overline{A} \right] \\
&+ \left[ \sum_i \sum_{jwt} W1_{i,jwt} \ln W1_{i,jwt} - \sum_i \sum_{jwt} W1_{i,jwt} \ln \overline{W1}_{i,jwt} \right] \\
&+ \left[ \sum_k \sum_{jwt} W2_{k,jwt} \ln W2_{k,jwt} - \sum_k \sum_{jwt} W2_{k,jwt} \ln \overline{W2}_{k,jwt} \right]
\end{aligned}
\tag{15}
$$

$$
\begin{bmatrix} \text{cross} \\ \text{entropy} \end{bmatrix} = \begin{bmatrix} \text{CE for} \\ \text{coefficient} \\ \text{matrix} \end{bmatrix} + \begin{bmatrix} \text{CE for weights} \\ \text{on the column} \\ \text{sum errors} \end{bmatrix} + \begin{bmatrix} \text{CE for weights} \\ \text{on macro} \\ \text{aggregate errors} \end{bmatrix}
$$

In this case the minimization problem is to find a set of $A$'s, $W1$'s, and $W2$'s that minimize cross entropy, where the $W$'s are treated like the $A$'s, subject to the following constraints:

*SAM Equation*

$$T_{i,j} = A_{i,j} \cdot \left( \overline{X}_i + e1_i \right) \tag{16}$$

$$\begin{bmatrix} \textit{SAM cell} \\ \textit{value} \end{bmatrix} = \begin{bmatrix} \textit{SAM} \\ \textit{coefficient} \\ \textit{matrix} \end{bmatrix} \cdot \left( \begin{bmatrix} \textit{value for} \\ \textit{column sum} \end{bmatrix} + \begin{bmatrix} \textit{error term} \\ \textit{associated} \\ \textit{with column} \\ \textit{sum} \end{bmatrix} \right)$$

where $\overline{X}$ is the prior on the column sums of the SAM matrix.

*Row/column sum consistency*

$$Y_i = \overline{X}_i + e1_i \tag{17}$$

$$\begin{bmatrix} \textit{SAM row} \\ \textit{sum} \end{bmatrix} = \begin{bmatrix} \textit{value for} \\ \textit{column sum} \end{bmatrix} + \begin{bmatrix} \textit{error term} \\ \textit{associated with} \\ \textit{column sum} \end{bmatrix}$$

*Error definition (column sum)*

$$e1_i = \sum_{jwt} W1_{i,jwt} \cdot \overline{v1}_{ijwt} \tag{18}$$

$$\begin{bmatrix} \textit{error term} \\ \textit{associated} \\ \textit{with column} \\ \textit{sum} \end{bmatrix} = \begin{bmatrix} \textit{error} \\ \textit{weights} \end{bmatrix} + \begin{bmatrix} \textit{error} \\ \textit{support} \\ \textit{values} \end{bmatrix}$$

*Sum of weights on errors (column sum)*

$$\sum_{jwt} W1_{i,jwt} = 1 \tag{19}$$

$$\begin{bmatrix} \textit{sum of} \\ \textit{error} \\ \textit{weights} \end{bmatrix} = \begin{bmatrix} \textit{add up} \\ \textit{to 1} \end{bmatrix}$$

*Row Sum*

$$\sum_{j} T_{i,j} = Y_i \tag{20}$$

$$\begin{bmatrix} \text{sum of} \\ \text{row} \\ \text{elements} \end{bmatrix} = \begin{bmatrix} \text{row} \\ \text{sum} \end{bmatrix}$$

*Column sum*

$$\sum_{i} T_{i,j} = \bar{X}_i + e1_i \tag{21}$$

$$\begin{bmatrix} \textit{sum of} \\ \textit{column} \\ \textit{elements} \end{bmatrix} = \begin{bmatrix} \textit{fixed prior} \\ \textit{value for} \\ \textit{column sum} \end{bmatrix} + \begin{bmatrix} \textit{error term} \\ \textit{associated} \\ \textit{with column} \\ \textit{sum} \end{bmatrix}$$

*Additional macro control totals*

The assumption is that one has additional knowledge about the new SAM. For example, aggregate national accounts data may be available for various macro aggregates such as value added, consumption, investment, government, exports, and imports. There also may be information about some of the SAM accounts such as government receipts and expenditures. This information can be summarized as additional linear adding-up constraints on various elements of the SAM while allowing the possibility that additional information might be measured with error. We can write:

$$\sum_{i} \sum_{j} G_{i,j}^{(k)} T_{i,j} = g^{(k)} + e2_k \tag{22}$$

$$\begin{bmatrix} \text{aggregator} \\ \text{matrix} \end{bmatrix} = \begin{bmatrix} \text{k'th control} \\ \text{total} \end{bmatrix}$$

where $G$ is an n-by-n aggregator matrix, which has ones for cells in the aggregate and zeros otherwise. Assume that there are k such aggregation constraints, and $g$ is the value of the aggregate. These conditions are simply added to the constraint set in the cross entropy formulation. The error term $e2$ is associated with macro aggregates. In the example provided in the following section two macro control total equations are included: GDP at factor cost and GDP at market prices.

*Error definition (macro totals)*

$$e2_k = \sum_{jwt} W2_{k,jwt} \cdot \overline{v2}_{k,jwt} \tag{23}$$

$$\begin{bmatrix} \text{error term} \\ \text{associated} \\ \text{with macro} \\ \text{totals} \end{bmatrix} = \begin{bmatrix} \text{error} \\ \text{weights} \end{bmatrix} + \begin{bmatrix} \text{error} \\ \text{support} \\ \text{values} \end{bmatrix}$$

*Sum of weights on errors (macro totals)*

$$\sum_{jwt} W2_{k,jwt} = 1 \tag{24}$$

$$\begin{bmatrix} \text{sum of} \\ \text{error} \\ \text{weights} \end{bmatrix} = \begin{bmatrix} \text{add up} \\ \text{to 1} \end{bmatrix}$$

*Additional Cell constraints*

Defining the SAM equation (16) over non-zero elements of $A$ guarantees that the zero structure of the original SAM is maintained in the estimated SAM. Fixing all the cells with a prior value of zero to zero greatly reduces the size of the estimation problem. If it is desired to allow a zero entry to become nonzero in the estimated SAM, then the restriction that the SAM equation (16) to be defined over non-zero elements of $A$ must be replaced to include cells which are currently zero but may be nonzero.

## 5. GAMS Code

This section provides the GAMS code implementing the equations described above. It is implemented with macro data from Mozambique. The cross entropy (CE) estimation problem is inherently badly scaled, and solution algorithms often suffer. A number of approaches to improving algorithm performance are available. There are two nonlinear programming solvers that are commonly used in GAMS: MINOS and CONOPT. If the column sums are known without error, then the constraint equations are all linear. In this case, the MINOS solver works well because it is optimized for nonlinear programming problems with linear constraints. If the column sums are assumed to include errors, then the constraints are nonlinear and the CONOPT solver appears to work better.

Another approach is to convert the nonlinear programming (NLP) problem into a mixed complementarity problem (MCP). The approach is to derive all the first-order conditions and create a set of "shadow price" variables, or Lagrange multipliers, associated with each first-order condition. The resulting problem is square in that it has as many equations as variables, but there is a complementary slackness relationship between the Lagrange multipliers and corresponding first-order conditions. The MCP solver called PATH handles this kind of problem (Dirkse and Ferris, 1995). The PATH solver appears to be very efficient in solving CE problems. The MCP formulation of the problem is given in a separate GAMS "include file" below.

## GAMS code

```
$TITLE Cross Entropy SAM Estimation
$OFFSYMLIST OFFSYMXREF OFFUPPER
*#########################################################################
*
*
* CE-SAM illustrates a cross entropy technique for estimating the cells
* of a consistent SAM assuming that the initial data are inconsistent
* and measured with error. The method is applied to a stylized macro
* SAM for Mozambique. Some macro control totals are assumed known with
* error, and also all the row and column totals are assumed
* known only with error. We assume that the user can specify
* a prior estimate of the standard error of the estimates of the row
* and column sums and of the macro control totals.
*
* Programmed by Sherman Robinson and Moataz El-Said, November 2000.
* Trade and Macroeconomics Division
* International Food Policy Research Institute (IFPRI)
* 2033 K Street, N.W.
* Washington, DC 20006 USA
* Email: S.Robinson@CGIAR.ORG
*        M.El-Said@CGIAR.ORG
*
* The method is described in S. Robinson, A. Cattaneo and, M. El Said
* (2001) "Updating and Estimating a Social Accounting Matrix Using
* Cross Entropy Methods." Economic Systems Research, Vol. 13, No. 1,
* pp. 47-64.
*
* Discussion paper #58 is an earlier version of the Economic
* Systems Research paper. A copy can be downloaded from the IFPRI
* web page using the following link:
*      http://www.ifpri.cgiar.org/divs/tmd/tmdpubs.htm#dp
*
* See also A. Golan, G. Judge, and D. Miller, Maximum Entropy
* Econometrics, John Wiley & Sons, 1996.
*
* Data set is based on a SAM developed by C. Arndt, A. S. Cruz, H. T.
* Jensen, S. Robinson, and F. Tarp, "Social Accounting Matrices
* for Mozambique - 1994 and 1995." TMD Discussion Paper No. 28, IFPRI,
* July 1998.
*
*#########################################################################


SETS

 i      sam accounts   / ACT     Activities
                         COM     Commodities
                         FAC     Factors
                         ENT     Enterprises
                         HOU     Households
                         GRE     Govt recurrent expenditures
                         GIN     Govt investment
                         CAP     Capital account
                         ROW     Rest of world
                         TOTAL /

 ii(i)  all acounts in i except TOTAL
                       / ACT     Activities
                         COM     Commodities
                         FAC     Factors
                         ENT     Enterprises
                         HOU     Households
                         GRE     Govt recurrent expenditures
                         GIN     Govt investment
                         CAP     Capital account
                         ROW     Rest of world /

 macro  macro controls  /gdpfc2, gdp2 /

* The set jwt defines the dimension of the support set for the error
* distribution and the number of weights that must be estimated for each
* error. In this case, we specify a five parameter error distribution.
* For a three parameter distribution, jwt is set to /1*3/.

 jwt    set of weights for errors in variables
                      / 1*5 /
;

* ii(i)       = YES;
* ii("Total") = NO;

ALIAS (i,j), (ii,jj);


*#########################    SAM DATABASE       #########################
TABLE SAM(i,j)  social accounting matrix
                      ACT             COM             FAC             ENT
ACT                   0.0      14827.4240             0.0             0.0
COM             7917.5040             0.0             0.0             0.0
FAC             9805.4140             0.0             0.0             0.0
ENT                   0.0             0.0       3699.7060             0.0
HOU                   0.0             0.0       6031.3080       3417.5060
GRE              733.6000        357.4000         74.4000        165.2000
GIN                   0.0             0.0             0.0             0.0
CAP                   0.0             0.0             0.0        150.0000
ROW                   0.0       5573.8150             0.0             0.0
Total          18456.5180       20758.639        9805.414        3732.706

  +                   HOU             GRE             GIN             CAP
ACT             2101.0490         -0.3270             0.0             0.0
*COM            6753.3320       1764.5000       2118.5000       2197.7980
COM             6953.3320       1564.5000       2518.5000       2597.7980
FAC                   0.0             0.0             0.0             0.0
ENT                   0.0         33.0000             0.0             0.0
HOU                   0.0         29.6000             0.0             0.0
GRE              139.5000             0.0             0.0             0.0
GIN                   0.0             0.0             0.0             0.0
CAP              649.1560       -356.6730       -406.2000             0.0
ROW                   0.0             0.0             0.0             0.0
Total            9643.037          1470.1          1712.3        2197.798

  +                   ROW       Total
ACT             1488.1570       18416.303
COM                   0.0       20751.634
FAC                   0.0        9805.414
ENT                   0.0        3732.706
HOU              209.5010        9687.915
GRE                   0.0          1470.1
GIN             1712.3000          1712.3
```

```
CAP              2163.8570        2200.14
ROW                    0.0       5573.815
Total            5573.815
     ;

*#########################  Parameters and Scalars #####################
PARAMETER

 SAM0(i,j)        Base SAM transactions matrix
 T0(i,j)          Matrix of SAM transactions (flow matrix)
 T1(i,j)          SAM transactions Adjusted to eliminate negative entries
 Abar0(i,j)       Prior SAM coefficient matrix
 Abar1(i,j)       Prior SAM adjusted to eliminate negative coefficients
 Target0(i)       Targets for macro SAM column totals
 vbar1(i,jwt)     Error support set 1
 vbar2(macro,jwt) Error support set 2
 wbar1(i,jwt)     Weights on error support set 1
 wbar2(macro,jwt) Weights on error support set 2
 sigmay1(i)       Prior standard error of column sums
 sigmay2(macro)   Prior standard error of macro aggregates
 epsilon          Tolerance to allow zero entries in SAM
;

SCALARS

 gdp0          base GDP
 gdp00         GDP from final SAM
 gdpfc0        GDP at factor cost
;

*################# Initializing Parameters
 SAM("TOTAL",jj)      = sum(ii, SAM(ii,jj));
 SAM(ii,"TOTAL")      = sum(jj, SAM(ii,jj));
 sam0(i,j)            = sam(i,j);

*#################
* Divide SAM entries by 1000 for better scaling.
* The SAM is scaled to enhance solver efficiency. Nonlinear solvers are
* more efficient if variables are scaled similarly. In this case,
* coefficients to be estimated range between 0 and 1, so SAM values
* are also scaled.

 Scalar scalesam Scaling value  /1000/ ;

 sam(i,j)                 = sam(i,j)/scalesam ;
 Abar0(ii,jj)$SAM(ii,jj)  = SAM(ii,jj)/SAM("TOTAL",jj) ;

 T0(ii,jj)                = SAM(ii,jj);
 T0("TOTAL",jj)           = sum(ii, SAM(ii,jj));
 T0(ii,"TOTAL")           = sum(jj, SAM(ii,jj));

 epsilon                  = .00001;

 Display T0, Abar0 ;
*########################  CROSS ENTROPY  ###############################


*########################  RED ALERT!!!   ##############################

* The ENTROPY DIFFERENCE procedure uses LOGARITHMS: negative flows in
* the SAM are NOT GOOD!!!
*
* The option used here is to detect any negative flows and net them out
* of their respective symmetric cells, e.g.
*     negative flow column to row is set to zero
*     and added to corresponding row to column as a positive number.
* The entropy difference method can then be implemented.
* After balancing, the negative SAM values are returned to their
* original cells for printing.

SET
 red(i,j)           Set of negative SAM flows
;

Parameter
 redsam(i,j)        Negative SAM values only
 rtot(i)            Row total
 ctot(i)            Column total
;

 rtot(ii)                            = sum(jj, T0(ii,jj));
 ctot(jj)                            = sum(ii, T0(ii,jj));

 red(ii,jj)$(T0(ii,jj) LT 0)         = yes ;
 redsam(ii,jj)                       = 0;
 redsam(ii,jj)$red(ii,jj)            = T0(ii,jj);
 redsam(jj,ii)$red(ii,jj)            = T0(ii,jj);

*Note that redsam includes each entry twice, in corresponding row
*and column. So, redsam need only be subtracted from T0.
 T1(ii,jj)                           = T0(ii,jj) - redsam(ii,jj);
 T1("Total",jj)                      = sum(ii, T1(ii,jj));
 T1(ii,"Total")                      = sum(jj, T1(ii,jj));

 redsam("total",jj)                  = sum(ii, redsam(ii,jj));
 redsam(ii,"total")                  = sum(jj, redsam(ii,jj));

 sam(ii,"total")                     = sum(jj, T1(ii,jj));
 sam("total",jj)                     = sum(ii, T1(ii,jj));

 rtot(ii)                            = sum(jj, T1(ii,jj));
 ctot(jj)                            = sum(ii, T1(ii,jj));

 Abar1(ii,jj)                        = T1(ii,jj)/sam("total",jj);

 display "NON-NEGATIVE SAM" ;
 display redsam, T1, Abar0, Abar1, rtot, ctot ;

* Define set of elements of SAM that can be nonzero. In this case, only
* elements which are nonzero in initial SAM.
 SET NONZERO(i,j)  SAM elements that can be nonzero ;

 NONZERO(ii,jj)$(Abar1(ii,jj)) = yes ;

*#### Initializing Parameters after accounting for negative values #####
*   Note that target column sums are being set to average of initial
*   row and column sums. Initial column sums or other values
*   could have been used instead, depending on knowledge of data quality
*   and any other prior information.

 target0(ii)        = (sam(ii,"total") +  sam("total",ii))/2 ;
 gdpfc0             = T1("fac","act");
```

```
 gdp0                = T1("fac","act") + T1("gre","act")
                        - T1("act","gre") + T1("gre","com") ;
Display gdpfc0, gdp0;

*############### Define variable bounds on errors #####################
* Start from assumed prior knowledge of the standard error (perhaps due
* to measurement error) of the column sums. Below, we assume that all
* column sums have a standard error of 5%. This is a Bayesian prior,
* not a maintained hypothesis.
* The estimated error is weighted sum of elements in an error support
* set:
*   ERR(ii) = SUM(jwt, W(ii,jwt)*VBAR(ii,jwt))
* where the W's are estimated in the CE procedure.
* The prior variance of these errors is given by:
*   (sigmay(ii))**2 = SUM(jwt, WBAR(ii,jwt)*(VBAR(ii,jwt))**2 )
* where the WBAR's are the prior on the weights.
* The VBARs are chosen to define a domain for the support set of +/- 3
* standard errors. The prior on the weights, WBAR, are then calculated
* to yield the specified prior on the standard error, sigmay.
* In Robinson, Cattaneo, and El-Said (2001), we specify prior weights
* (WBAR) that are uniform and set the prior standard error by the
* choice of support set, VBAR. In that paper, we use a three-weight
* specification (jwt /1*3/);
*
* We define two sets of errors with separate weights, W1 and W2. The
* first is for specifying errors on column sums, the second for errors
* on macro aggregates (defined in the set macro).

* First, define standard error for errors on column sums.

  sigmay1(ii)    = 0.05 * target0(ii) ;

* This code assumes a prior mean of zero and a two-parameter
* distribution with specified prior standard error. There are three
* weights, W(ii,jwt), to be estimated. The actual moments are estimated
* as part of the estimation procedure.
$ontext
* Set constants for 3-weight error distribution
  vbar1(ii,"1")  = -3 * sigmay1(ii);
  vbar1(ii,"2")  =  0              ;
  vbar1(ii,"3")  = -3 * sigmay1(ii);

  wbar1(ii,"1")  =  1/18  ;
  wbar1(ii,"2")  = 16/18  ;
  wbar1(ii,"3")  =  1/18  ;
$offtext

* This code assumes a prior mean of zero and a prior value of kurtosis
* consistent with a prior normal distribution with mean zero, variance
* sigmay**2, and kurtosis equal to 3*sigmay**4. The addition of a prior
* on kurtosis requires estimation of 5 weights (jwt = 5);
* The prior weights wbar are specified so that:
* SUM(jwt, wbar(ii,jwt)*vbar(ii,jwt)**4) = 3*sigmay(ii,jwt)**4
* as well as defining the variance as above.
* The prior weights and support set are also symmetric, so the prior
* on all odd moments is zero. The choice of +/- 1 standard error
* for vbar(ii,"2") and vbar(ii,"4") is arbitrary.
* The actual moments are estimated as part of the estimation procedure.

* Set constants for 5-weight error distribution
  vbar1(ii,"1")  = -3 * sigmay1(ii) ;
  vbar1(ii,"2")  = -1 * sigmay1(ii) ;
  vbar1(ii,"3")  =  0               ;
  vbar1(ii,"4")  = +1 * sigmay1(ii) ;
  vbar1(ii,"5")  = +3 * sigmay1(ii) ;

  wbar1(ii,"1")  =   1/72       ;
  wbar1(ii,"2")  =  27/72       ;
  wbar1(ii,"3")  =  16/72       ;
  wbar1(ii,"4")  =  27/72       ;
  wbar1(ii,"5")  =   1/72       ;

* Second, define standard errors for errors on macro aggregates

  sigmay2("gdpfc2")  = 0.05*gdpfc0  ;
  sigmay2("gdp2")    = 0.05*gdp0 ;

$ontext
* Set constants for 3-weight error distribution
  vbar2(ii,"1")  = -3 * sigmay2(ii);
  vbar2(ii,"2")  =  0             ;
  vbar2(ii,"3")  = -3 * sigmay2(ii);

  wbar2(ii,"1")  =  1/18  ;
  wbar2(ii,"2")  = 16/18  ;
  wbar2(ii,"3")  =  1/18  ;
$offtext

* Set constants for 5-weight error distribution
  vbar2(macro,"1")  = -3 * sigmay2(macro) ;
  vbar2(macro,"2")  = -1 * sigmay2(macro) ;
  vbar2(macro,"3")  =  0                  ;
  vbar2(macro,"4")  = +1 * sigmay2(macro) ;
  vbar2(macro,"5")  = +3 * sigmay2(macro) ;

  wbar2(macro,"1")  =   1/72       ;
  wbar2(macro,"2")  =  27/72       ;
  wbar2(macro,"3")  =  16/72       ;
  wbar2(macro,"4")  =  27/72       ;
  wbar2(macro,"5")  =   1/72       ;

 Display vbar1, vbar2, sigmay1, sigmay2 ;

*###################### VARIABLES #################################
 VARIABLES
 A(ii,jj)       Post SAM coefficient matrix
 TSAM(ii,jj)    Post matrix of SAM transactions
 Y(ii)          row sum of SAM
 X(ii)          column sum of SAM
 ERR1(ii)       Error value on column sums
 ERR2(macro)    Error value for macro aggregates
 W1(ii,jwt)     Error weights
 W2(macro,jwt)  Error weights
 DENTROPY       Entropy difference (objective)
 GDPFC          GDP at factor cost
 GDP            GDP at market prices
  ;

*######################### INITIALIZE VARIABLES #################

 A.L(ii,jj)         = Abar1(ii,jj) ;
 TSAM.L(ii,jj)      = T1(ii,jj)    ;
```

```
 Y.L(ii)             = target0(ii)  ;
 X.L(ii)             = target0(ii)  ;
 ERR1.L(ii)          = 0.0          ;
 ERR2.L(macro)       = 0.0          ;
 W1.L(ii,jwt)        = wbar1(ii,jwt) ;
 W2.L(macro,jwt)     = wbar2(macro,jwt) ;
 DENTROPY.L          = 0            ;
 GDPFC.L             = gdpfc0       ;
 GDP.L               = gdp0         ;

*############ CORE EQUATIONS
 EQUATIONS

 SAMEQ(i)        row and column sum constraint
 SAMMAKE(i,j)    make SAM flows
 ERROR1EQ(i)     definition of error term 1
 ERROR2EQ(macro) definition of error term 2
 SUMW1(i)        Sum of weights 1
 SUMW2(macro)    Sum of weights 2
 ENTROPY         Entropy difference definition
 ROWSUM(i)       row target
 COLSUM(j)       column target
 GDPFCDEF        define GDP at factor cost
 GDPDEF          define GDP
 ;

*CORE EQUATIONS====================================================

 SAMEQ(ii)..   Y(ii)       =E= X(ii) + ERR1(ii) ;

 SAMMAKE(ii,jj)$nonzero(ii,jj)..
                  TSAM(ii,jj) =E= A(ii,jj) * (X(jj) + ERR1(jj)) ;

 ERROR1EQ(ii).. ERR1(ii)    =E= SUM(jwt, W1(ii,jwt)*vbar1(ii,jwt)) ;

 SUMW1(ii)..   SUM(jwt, W1(ii,jwt)) =E= 1 ;

 ENTROPY..     DENTROPY    =E= SUM((ii,jj)$nonzero(ii,jj),
                             A(ii,jj)*(LOG(A(ii,jj) + epsilon)
                           - LOG(Abar1(ii,jj) + epsilon)))
                           +
                             SUM((ii,jwt), W1(ii,jwt)
                           * (LOG(W1(ii,jwt) + epsilon)
                           - LOG(wbar1(ii,jwt) + epsilon)))
                           +
                             SUM((macro,jwt), W2(macro,jwt)
                           * (LOG(W2(macro,jwt) + epsilon)
                           - LOG(wbar2(macro,jwt) + epsilon))) ;

* Note that we exclude one rowsum equation since if all but one column
* and rowsum are equal, the last one must also be equal. Walras' Law
* at work.

 ROWSUM(ii)$(NOT SAMEAS(ii,"ROW"))..   SUM(jj, TSAM(ii,jj)) =E= Y(ii) ;

 COLSUM(jj)..     SUM(ii, TSAM(ii,jj)) =E= (X(jj) + ERR1(jj)) ;

*ADDITIONAL MACRO CONTROL-TOTAL EQUATIONS==========================

 GDPFCDEF..         GDPFC =E= TSAM("fac","act") + ERR2("gdpfc2") ;
```

```
 GDPDEF..           GDP   =E= TSAM("fac","act") + TSAM("gre","act")
                            - TSAM("act","gre") + TSAM("gre","com")
                            + ERR2("gdp2") ;

 ERROR2EQ(macro)..  ERR2(macro)
                        =E= SUM(jwt, W2(macro,jwt)*vbar2(macro,jwt)) ;

 SUMW2(macro)..     SUM(jwt, W2(macro,jwt)) =E= 1 ;

*###############  Define bounds for cell values  ####################

* Defining equation SAMMAKE over non-zero elements of A ($Abar1(ii,jj))
* guarantees that the zero structure of the original SAM is maintained
* in the estimated SAM. Fixing all the zero entries to zero greatly
* reduces the size of the estimation problem. If it is desired to
* allow a zero entry to become nonzero in the estimated SAM, then
* the condition $ABAR1(ii,jj) must be replaced with a new set that
* does not include cells which are currently zero but may be nonzero.

 A.LO(ii,jj)$nonzero(ii,jj)          = 0 ;
 A.UP(ii,jj)$nonzero(ii,jj)          = 1 ;
 A.FX(ii,jj)$(NOT nonzero(ii,jj))    = 0;

 TSAM.lo(ii,jj)                      = 0.0 ;
 TSAM.up(ii,jj)                      = +inf ;
 TSAM.FX(ii,jj)$(NOT nonzero(ii,jj)) = 0 ;

* Upper and lower bounds on the error weights
 W1.LO(ii,jwt)                       = 0   ;
 W1.UP(ii,jwt)                       = 1   ;
 W2.LO(macro,jwt)                    = 0   ;
 W2.UP(macro,jwt)                    = 1   ;

* Set target column sums, X. If these are not fixed, then the column sum
* constraints will not be binding and the solution values or ERR1 will
* be 0.

  X.FX(ii)                            = TARGET0(ii) ;

* Fix Macro aggregates.
* If these are not fixed, then the macro constraints will not be binding
* and the solution values of ERR2 will be zero.
  GDP.FX              = GDP0 ;
  GDPFC.FX            = GDPFC0 ;

*######################## DEFINE MODEL ############################

 MODEL SAMENTROP / ALL /

*######################## SOLVE MODEL #############################

 OPTION ITERLIM       = 5000;
 OPTION LIMROW        = 0,   LIMCOL    = 0;
 OPTION SOLPRINT      = ON;


* SAMENTROP.optfile   = 1 ;
  SAMENTROP.HOLDFIXED = 1 ;
  option NLP          = MINOS5 ;
* OPTION NLP          = CONOPT;
* SAMENTROP.WORKSPACE = 25.0;
```

```
*########################### Solve statement #####################

  SOLVE SAMENTROP using nlp minimizing dentropy ;

*##################################################################


*##(alternative formulation)#### MCP Formulation ################

* Add code restating the nonlinear-programming (NLP) minimization
* problem as an MCP problem solved using the MCP solver.
*$include CE-MCP.INC

*##################################################################

*---------------- Parameters for reporting results
Parameters
 Macsam1(i,j)       Assigned new balanced SAM flows from CE
 Macsam2(i,j)       Balanced SAM flows from entropy diff x scalesam
 SEM                Squared Error Measure
 percent1(i,j)      percent change of new SAM from original SAM
 PosUnbal(i,j)      Positive unbalanced SAM
 PosBalan(i,j)      Positive balanced SAM
 Diffrnce(i,j)      Differnce btw original SAM and Final SAM in values
 NormEntrop         Normalized Entropy a measure of total uncertainty
 ;

 macsam1(ii,jj)         = TSAM.l(ii,jj);
 macsam1("total",jj)    = SUM(ii, macsam1(ii,jj)) ;
 macsam1(ii,"total")    = SUM(jj, macsam1(ii,jj)) ;
 macsam2(i,j)           = macsam1(i,j) * scalesam ;
 SEM                    = Sum((ii,jj), SQR(A.L(ii,jj)
                                 - Abar1(ii,jj)))/SQR(card(ii));
 percent1(i,j)$(T1(i,j))= 100*(macsam1(i,j)-T1(i,j))/T1(i,j);
 PosUnbal(i,j)          = T1(i,j) * scalesam;
 PosBalan(i,j)          = macsam2(i,j);
 Diffrnce(i,j)          = PosBalan(i,j) - PosUnbal(i,j);
 NormEntrop             = SUM((ii,jj)$(Abar1(ii,jj)), A.L(ii,jj)*
                          LOG (A.L(ii,jj))) /
SUM((ii,jj)$(Abar1(ii,jj)),
                          Abar1(ii,jj)* LOG (Abar1(ii,jj)))
 ;
 display macsam1, macsam2, percent1, sem, dentropy.l, PosUnbal,
        PosBalan, NormEntrop, Diffrnce ;

*########### Return negative flows to initial cell position ############

 macsam1(ii,jj)         = macsam1(ii,jj) + redsam(ii,jj) ;
 macsam1("total",jj)    = SUM(ii, macsam1(ii,jj)) ;
 macsam1(ii,"total")    = SUM(jj, macsam1(ii,jj)) ;
 macsam2(i,j)           = macsam1(i,j) * scalesam ;

 gdp00                  = macsam1("fac","act") + macsam1("gre","act")
                          - macsam1("act","gre") + macsam1("gre","com")
;

 display macsam1, macsam2 ;
 display gdp0, gdp00, gdp.l, gdpfc0, gdpfc.l ;

*#######

 Parameter ANEW(i,j) ;
* print some stuff
 ANEW("total",jj)   = SUM(ii, A.L(ii,jj)) ;
 ANEW(ii,"total")   = SUM(jj, A.L(ii,jj)) ;

 ABAR1("total",jj) = SUM(ii, ABAR1(ii,jj)) ;
 ABAR1(ii,"total") = SUM(jj, ABAR1(ii,jj)) ;

 Display ANEW, ABAR1 ;

 scalar meanerr1, meanerr2 ;
 meanerr1 = SUM(ii, abs(err1.l(ii)))/card(ii) ;
 meanerr2 = SUM(macro, abs(err2.l(macro)))/card(macro) ;
 display meanerr1, meanerr2 ;

* Use the following code to specify that the column sums are known
* exactly. The errors are thus fixed to zero and two equations are
* dropped from the estimation procedure. The computational gains are
* that the constraints are all linear and the estimation problem is
* considerably smaller.

*######
$ontext
*######

 ERR1.FX(ii)       = 0.0 ;
 W1.FX(ii,jwt)     = WBAR1(ii,jwt) ;

 MODEL SAMENTROP2 /SAMEQ
                 SAMMAKE
*                ERROR1EQ
*                SUMW1
                 ERROR2EQ
                 SUMW2
                 ENTROPY
                 ROWSUM
                 COLSUM
                 GDPFCDEF
                 GDPDEF
                          /
;
 SAMENTROP2.holdfixed = 1 ;

*######################## Solve statenment ##################

 SOLVE SAMENTROP2 using nlp minimizing dentropy ;

*##############################################################

*#######
$offtext
*#######

*#*#*#*#*#*#*#*#*#*#*#*#* THE END *#*#*#*#*#*#*#*#*#*#*#*#*#*#*
```

## GAMS code for the MCP Formulation

File name: CE-MCP.INC

```
* The code below is a translation of the NLP problem into a
* mixed complementarity problem (MCP), which can be solved
* using an MCP solver in GAMS. The translation was done using
* a preliminary version of a program called NLP2MCP written
* by Michael Ferris and Jeffrey Horn (1998) at the University
* of Wisconsin. The translation adds "shadow price" or
* complementarity variables for all constraint equations and
* also provides equations for all the first-order conditions
* for minimizing the objective function. The resulting model
* is square with as many variables as equations.

alias(macro,a_macro);
alias(ii,a_ii);
alias(jj,a_jj);
alias(jwt,a_jwt);

*#####  "SHADOW PRICE" OR COMPLEMENTARITY VARIABLES OR   ##########
*##### LAGRANGE MULTIPLIERS FOR ALL CONSTRAINT EQUATIONS ##########

variables
  m_SAMEQ(i)         Multiplier for row and column sum constraint
  m_SAMMAKE(i,j)     Multiplier for make SAM flows constraint
  m_ERROR1EQ(i)      Multiplier for definition of error term 1
  m_ERROR2EQ(macro)  Multiplier for definition of error term 2
  m_SUMW1(i)         Multiplier for Sum of weights 1 constraint
  m_SUMW2(macro)     Multiplier for Sum of weights 1 constraint
  m_ROWSUM(i)        Multiplier for row target constraint
  m_COLSUM(j)        Multiplier for column target constraint
  m_GDPFCDEF         Multiplier for GDP at factor cost constraint
  m_GDPDEF           Multiplier for GDP at market pricesconstraint
;

*############# EQUATIONS FOR THE FIRST ORDER CONDITIONS ############
*############# FOR MINIMIZING THE OBJECTIVE FUNCTION    ############

EQUATIONS
  d_A(a_ii,a_jj)        FOC wrt the choice variable A
  d_TSAM(a_ii,a_jj)     FOC wrt the variable TSAM
  d_Y(a_ii)             FOC wrt the variable Y
  d_X(a_ii)             FOC wrt the variable X
  d_ERR1(a_ii)          FOC wrt the variable ERR1
  d_ERR2(a_macro)       FOC wrt the variable ERR2
  d_W1(a_ii,a_jwt)      FOC wrt the choice variable W1
  d_W2(a_macro,a_jwt)   FOC wrt the choice variable W2
  d_GDPFC               FOC wrt the macro control variable GDPFC
  d_GDP                 FOC wrt the macro control variable GDP
;


*##### EQUATION: FOC wrt the choice variable A  #################
d_A(a_ii,a_jj) ..

 ((log(a(a_ii,a_jj)+epsilon)-log(abar1(a_ii,a_jj)+epsilon))
 $(nonzero(a_ii,a_jj)) + a(a_ii,a_jj)*(1/(a(a_ii,a_jj)+epsilon))
 $(nonzero(a_ii,a_jj)))-

 SUM((ii,jj)$(nonzero(ii,jj)),m_sammake(ii,jj)*(-(1$((sameas
 (a_ii,ii) and sameas(a_jj,jj)))))*(x(jj)+err1(jj)))) =e= 0;


*##### EQUATION: FOC wrt the variable TSAM #####################
d_TSAM(a_ii,a_jj) .. -

 m_gdpfcdef*(-(1$((sameas(a_ii,"fac") and sameas(a_jj,"act"
 )))))-

 m_colsum(a_jj)-

 m_gdpdef*(-(1$((sameas(a_ii,"fac") and sameas(a_jj,"act"))
 )+1$((sameas(a_ii,"gre") and sameas(a_jj,"act")))-(1$((sameas
 (a_ii,"act") and sameas(a_jj,"gre"))))+1$((sameas(a_ii,"gre"
 ) and sameas(a_jj,"com")))))-

 m_rowsum(a_ii)$((not sameas(a_ii,"ROW")))-

 m_sammake(a_ii,a_jj)$(nonzero(a_ii,a_jj)) =e= 0;


*##### EQUATION: FOC wrt the variable Y ########################
d_Y(a_ii) .. -

 m_sameq(a_ii)-

 SUM((ii)$((not sameas(ii,"ROW"))),m_rowsum(ii)*(-(1$((sameas
 (a_ii,ii))))))

 =e= 0;

*##### EQUATION FOC wrt the variable X #########################
d_X(a_ii) .. -

 SUM(jj,m_colsum(jj)*(-(1$((sameas(a_ii,jj))))))-

 SUM(ii,m_sameq(ii)*(-(1$((sameas(a_ii,ii))))))-

 SUM((ii,jj)$(nonzero(ii,jj)),m_sammake(ii,jj)*(-a(ii,jj)*(1
 $((sameas(a_ii,jj)))))) =e= 0;

*##### EQUATION: FOC wrt the variable ERR1 #####################
d_ERR1(a_ii) .. -

 m_error1eq(a_ii)-

 SUM(jj,m_colsum(jj)*(-(1$((sameas(a_ii,jj))))))-

 SUM(ii,m_sameq(ii)*(-(1$((sameas(a_ii,ii))))))-
```

```
 SUM((ii,jj)$(nonzero(ii,jj)),m_sammake(ii,jj)*(-a(ii,jj)*(1
 $((sameas(a_ii,jj))))))

 =e= 0;

*##### EQUATION: FOC wrt the variable ERR2 ######################
d_ERR2(a_macro) .. -

 m_error2eq(a_macro)-

 m_gdpfcdef*(-(1$((sameas(a_macro,"gdpfc2")))))-

 m_gdpdef*(-(1$((sameas(a_macro,"gdp2")))))

 =e= 0;

*##### EQUATION: FOC wrt the choice variable W1 #################
d_W1(a_ii,a_jwt) .. ((log(w1(a_ii,a_jwt)+epsilon)-log(wbar1(a_ii
 ,a_jwt)+epsilon))+w1(a_ii,a_jwt)*(1/(w1(a_ii,a_jwt)+epsilon)))-

 SUM(ii,m_error1eq(ii)*(-(vbar1(ii,a_jwt)$(sameas(a_ii,ii))
 )))-

 m_sumw1(a_ii)

 =e= 0;

*##### EQUATION: FOC wrt the choice variable W2 #################
d_W2(a_macro,a_jwt) .. ((log(w2(a_macro,a_jwt)+epsilon)-log(wbar2
 (a_macro,a_jwt)+epsilon))+w2(a_macro,a_jwt)*(1/(w2(a_macro,a_jwt
 )+epsilon)))-

 SUM(macro,m_error2eq(macro)*(-(vbar2(macro,a_jwt)$(sameas(a_macro
 ,macro)))))-

 m_sumw2(a_macro)

 =e= 0;

*##### EQUATION: FOC wrt the macro control variable GDPFC #######
d_GDPFC .. -

 m_gdpfcdef =e= 0;

*##### EQUATION: FOC wrt the macro control variable GDP #########
d_GDP .. -

 m_gdpdef =e= 0;

*####################### DEFINE MODEL ###########################
* In GAMS the "." is used for pairing the complementarity variables
* and equations for the MCP solver. For example the equation
* defined by d_A is complementary to the variable A and must be
* defined over the same sets.

 MODEL m_SAMENTROP /
    d_A.A
    d_TSAM.TSAM
    d_Y.Y
    d_X.X
    d_ERR1.ERR1
    d_ERR2.ERR2
    d_W1.W1
    d_W2.W2
    d_GDPFC.GDPFC
    d_GDP.GDP
    ERROR1EQ.m_ERROR1EQ
    ERROR2EQ.m_ERROR2EQ
    GDPFCDEF.m_GDPFCDEF
    COLSUM.m_COLSUM
    SAMEQ.m_SAMEQ
    GDPDEF.m_GDPDEF
    SUMW1.m_SUMW1
    ROWSUM.m_ROWSUM
    SUMW2.m_SUMW2
    SAMMAKE.m_SAMMAKE
                    /
 ;

*############################ SOLVE MODEL ########################

*Shock the NLP solution
 A.L(ii,jj) = 0.9*A.l(ii,jj) ;

*########################### Solve statenment ####################

  SOLVE m_SAMENTROP using mcp;

**###############################################################

*Compare NLP and MCP results.
 Scalar savedent ;
 savedent = dentropy.l ;

DENTROPY.l   = SUM((ii,jj)$nonzero(ii,jj),
                              A.l(ii,jj)*(LOG(A.l(ii,jj) + epsilon)
                             - LOG(Abar1(ii,jj) + epsilon)))
                             +
                               SUM((ii,jwt), W1.l(ii,jwt)
                             * (LOG(W1.l(ii,jwt) + epsilon)
                             - LOG(wbar1(ii,jwt) + epsilon)))
                             +
                               SUM((macro,jwt), W2.l(macro,jwt)
                             * (LOG(W2.l(macro,jwt) + epsilon)
                             - LOG(wbar2(macro,jwt) + epsilon))) ;

 option decimals=8 ;
 display dentropy.l, savedent ;
 option decimals=3 ;

*###################### END MCP INCLUDE FILE #####################



*###### NOTE ON THE USE OF "SAMEAS" GAMS COMMAND ##################
*######   Undocumented Feature IN GAMS Manual    ##################
$ontext
Matching Set Elements

New features in GAMS allow one to introduce conditional statements
controlling execution in cases where certain items match up . The
```

```
syntax involves using the commands

  SAMEAS(setelement1,setelement2)
             or
  DIAG(setelement,setelement)

the SAMEAS command returns a true false indicator which is true
if the text string defining the name of set element 1 equals that
for setelement 2 and false otherwise. DIAG returns a 1 under
equality and a zero otherwise.

For example
  x=sum((i,j)$(not SAMEAS(i,j)),z(i)*z(j));
                  or
  x=sum((i,j)$(DIAG(i,j) eq 0),z(i)*z(j));

would exclude the cases where i=j from the sum

while
  x=sum((i,j)$(SAME AS(i,"case1") or SAME AS(j,"case2")),z(i)+z(j));

would only include cases where the text for i equaled the string
"case1" or the text for j corresponded to "case2."

If interested check the following web address Undocumented Features
and Usage Tips

 http://agrinet.tamu.edu/mccarl/gamstip.htm

$offtext

*############### END Note on "SAMEAS" GAMS command ################
```

**References**

Brooke, A., D. Kendrick, A. Meeraus, and R. Raman (1998) *GAMS a User's Guide* (GAMS Development Corporation, Washington D.C.).

Dirkse, S. and M. Ferris (1995) “The PATH solver: A non-monotone stabilization scheme for mixed complementarity problems.” *Optimization Methods and Software,* 5:123-156.

Ferris, Michael, and Jeffrey Horn (1998) “NLP2MCP: Automatic conversion of nonlinear programs into mixed complementarity problems.” Technical report, Computer Sciences Department, University of Wisconsin.

Golan, Amos, George Judge, and Douglas Miller (1996) *Maximum Entropy Econometrics, Robust Estimation with Limited Data* (John Wiley & Sons).

Kullback, S. and R. A. Leibler (1951) “On Information and Sufficiency.” *Ann. Math. Stat.* 4, 99-111.

S. Robinson, A. Cattaneo and, M. El Said (2001) “Updating and Estimating a Social Accounting Matrix Using Cross Entropy Methods.” *Economic Systems Research, vol. 13, No.1, pp.47-64.*

S. Robinson, A. Cattaneo and, M. El Said (2000) “Updating and Estimating a Social Accounting Matrix Using Cross Entropy Methods.” TMD Discussion Paper No. 58, IFPRI, August 2000.

**List of Discussion Papers**

No. 1 - "Land, Water, and Agriculture in Egypt: The Economywide Impact of Policy Reform" by Sherman Robinson and Clemen Gehlhar (January 1995)

No. 2 - "Price Competitiveness and Variability in Egyptian Cotton: Effects of Sectoral and Economywide Policies" by Romeo M. Bautista and Clemen Gehlhar (January 1995)

No. 3 - "International Trade, Regional Integration and Food Security in the Middle East" by Dean A. DeRosa (January 1995)

No. 4 - "The Green Revolution in a Macroeconomic Perspective: The Philippine Case" by Romeo M. Bautista (May 1995)

No. 5 - "Macro and Micro Effects of Subsidy Cuts: A Short-Run CGE Analysis for Egypt" by Hans Löfgren (May 1995)

No. 6 - "On the Production Economics of Cattle" by Yair Mundlak, He Huang and Edgardo Favaro (May 1995)

No. 7 - "The Cost of Managing with Less: Cutting Water Subsidies and Supplies in Egypt's Agriculture" by Hans Löfgren (July 1995, Revised April 1996)

No. 8 - "The Impact of the Mexican Crisis on Trade, Agriculture and Migration" by Sherman Robinson, Mary Burfisher and Karen Thierfelder (September 1995)

No. 9 - "The Trade-Wage Debate in a Model with Nontraded Goods: Making Room for Labor Economists in Trade Theory" by Sherman Robinson and Karen Thierfelder (Revised March 1996)

No. 10 - "Macroeconomic Adjustment and Agricultural Performance in Southern Africa: A Quantitative Overview" by Romeo M. Bautista (February 1996)

No. 11 - "Tiger or Turtle? Exploring Alternative Futures for Egypt to 2020" by Hans Löfgren, Sherman Robinson and David Nygaard (August 1996)

No. 12 - "Water and Land in South Africa: Economywide Impacts of Reform - A Case Study for the Olifants River" by Natasha Mukherjee (July 1996)

No. 13 - "Agriculture and the New Industrial Revolution in Asia" by Romeo M. Bautista and Dean A. DeRosa (September 1996)

No. 14 - "Income and Equity Effects of Crop Productivity Growth Under Alternative Foreign Trade Regimes: A CGE Analysis for the Philippines" by Romeo M. Bautista and Sherman Robinson (September 1996)

No. 15 - "Southern Africa: Economic Structure, Trade, and Regional Integration" by Natasha Mukherjee and Sherman Robinson (October 1996)

No. 16 - "The 1990's Global Grain Situation and its Impact on the Food Security of Selected Developing Countries" by Mark Friedberg and Marcelle Thomas (February 1997)

No. 17 - "Rural Development in Morocco: Alternative Scenarios to the Year 2000" by Hans Löfgren, Rachid Doukkali, Hassan Serghini and Sherman Robinson (February 1997)

No. 18 - "Evaluating the Effects of Domestic Policies and External Factors on the Price Competitiveness of Indonesian Crops: Cassava, Soybean, Corn, and Sugarcane" by Romeo M. Bautista, Nu Nu San, Dewa Swastika, Sjaiful Bachri and Hermanto (June 1997)

No. 19 - "Rice Price Policies in Indonesia: A Computable General Equilibrium (CGE) Analysis" by Sherman Robinson, Moataz El-Said, Nu Nu San, Achmad Suryana, Hermanto, Dewa Swastika and Sjaiful Bahri (June 1997)

No. 20 - "The Mixed-Complementarity Approach to Specifying Agricultural Supply in Computable General Equilibrium Models" by Hans Löfgren and Sherman Robinson (August 1997)

No. 21 - "Estimating a Social Accounting Matrix Using Entropy Difference Methods" by Sherman Robinson and Moataz-El-Said (September 1997)

No. 22 - "Income Effects of Alternative Trade Policy Adjustments on Philippine Rural Households: A General Equilibrium Analysis" by Romeo M. Bautista and Marcelle Thomas (October 1997)

No. 23 - "South American Wheat Markets and MERCOSUR" by Eugenio Díaz-Bonilla (November 1997)

No. 24 - "Changes in Latin American Agricultural Markets" by Lucio Reca and Eugenio Díaz-Bonilla (November 1997)

No. 25* - "Policy Bias and Agriculture: Partial and General Equilibrium Measures" by Romeo M. Bautista, Sherman Robinson, Finn Tarp and Peter Wobst (May 1998)

No. 26 - "Estimating Income Mobility in Colombia Using Maximum Entropy Econometrics" by Samuel Morley, Sherman Robinson and Rebecca Harris (Revised February 1999)

No. 27 - "Rice Policy, Trade, and Exchange Rate Changes in Indonesia: A General Equilibrium Analysis" by Sherman Robinson, Moataz El-Said and Nu Nu San (June 1998)

No. 28* - "Social Accounting Matrices for Mozambique - 1994 and 1995" by Channing Arndt, Antonio Cruz, Henning Tarp Jensen, Sherman Robinson and Finn Tarp (July 1998)

No. 29* - "Agriculture and Macroeconomic Reforms in Zimbabwe: A Political-Economy Perspective" by Kay Muir-Leresche (August 1998)

No. 30* - "A 1992 Social Accounting Matrix (SAM) for Tanzania" by Peter Wobst (August 1998)

No. 31* - "Agricultural Growth Linkages in Zimbabwe: Income and Equity Effects" by Romeo M. Bautista and Marcelle Thomas (September 1998)

No. 32* - "Does Trade Liberalization Enhance Income Growth and Equity in Zimbabwe? The Role of Complementary Polices" by Romeo M.Bautista, Hans Lofgren and Marcelle Thomas (September 1998)

No. 33 - "Estimating a Social Accounting Matrix Using Cross Entropy Methods" by Sherman Robinson, Andrea Cattaneo and Moataz El-Said (October 1998)

No. 34 - "Trade Liberalization and Regional Integration: The Search for Large Numbers" by Sherman Robinson and Karen Thierfelder (January 1999)

No. 35 - "Spatial Networks in Multi-Region Computable General Equilibrium Models" by Hans Löfgren and Sherman Robinson (January 1999)

No. 36* - "A 1991 Social Accounting Matrix (SAM) for Zimbabwe" by Marcelle Thomas, and Romeo M. Bautista (January 1999)

No. 37 - "To Trade or not to Trade: Non-Separable Farm Household Models in Partial and General Equilibrium" by Hans Löfgren and Sherman Robinson (January 1999)

No. 38 - "Trade Reform and the Poor in Morocco: A Rural-Urban General Equilibrium Analysis of Reduced Protection" by Hans Löfgren (January 1999)

No. 39 - " A Note on Taxes, Prices, Wages, and Welfare in General Equilibrium Models" by Sherman Robinson and Karen Thierfelder (January 1999)

No. 40 - "Parameter Estimation for a Computable General Equilibrium Model: A Maximum Entropy Approach" by Channing Arndt, Sherman Robinson and Finn Tarp (February 1999)

No. 41 - "Trade Liberalization and Complementary Domestic Policies: A Rural-Urban General Equilibrium Analysis of Morocco" by Hans Löfgren, Moataz El-Said and Sherman Robinson (April 1999)

No. 42 - "Alternative Industrial Development Paths for Indonesia: SAM and CGE Analysis" by Romeo M. Bautista, Sherman Robinson and Moataz El-Said (May 1999)

No. 43* - "Marketing Margins and Agricultural Technology in Mozambique" by Channing Arndt, Henning Tarp Jensen, Sherman Robinson and Finn Tarp (July 1999)

No. 44 - "The Distributional Impact of Macroeconomic Shocks in Mexico: Threshold Effects in a Multi-Region CGE Model" by Rebecca Lee Harris (July 1999)

No. 45 - "Economic Growth and Poverty Reduction in Indochina: Lessons From East Asia" by Romeo M. Bautista (September 1999)

No. 46* - "After the Negotiations: Assessing the Impact of Free Trade Agreements in Southern Africa" by Jeffrey D. Lewis, Sherman Robinson and Karen Thierfelder (September 1999)

No. 47* - "Impediments to Agricultural Growth in Zambia" by Rainer Wichern, Ulrich Hausner and Dennis K. Chiwele (September 1999)

No. 48 - "A General Equilibrium Analysis of Alternative Scenarios for Food Subsidy Reform in Egypt" by Hans Lofgren and Moataz El-Said (September 1999)

No. 49*- " A 1995 Social Accounting Matrix for Zambia" by Ulrich Hausner (September 1999)

No. 50 - "Reconciling Household Surveys and National Accounts Data Using a Cross Entropy Estimation Method" by Anne-Sophie Robilliard and Sherman Robinson (November 1999)

No. 51 - "Agriculture-Based Development: A SAM Perspective on Central Viet Nam" by Romeo M. Bautista (January 2000)

No. 52 - "Structural Adjustment, Agriculture, and Deforestation in the Sumatera Regional Economy" by Nu Nu San, Hans Löfgren and Sherman Robinson (March 2000)

No. 53 - "Empirical Models, Rules, and Optimization: Turning Positive Economics on its Head" by Andrea Cattaneo and Sherman Robinson (April 2000)

No. 54 - "Small Countries and the Case for Regionalism vs. Multilateralism" by Mary E. Burfisher, Sherman Robinson and Karen Thierfelder (May 2000)

No. 55 - "Genetic Engineering and Trade: Panacea or Dilemma for Developing Countries" by Chantal Pohl Nielsen, Sherman Robinson and Karen Thierfelder (May 2000)

No. 56 - "An International, Multi-region General Equilibrium Model of Agricultural Trade Liberalization in the South Mediterranean NIC's, Turkey, and the European Union" by Ali Bayar, Xinshen Diao, A. Erinc Yeldan (May 2000)

No. 57* - "Macroeconomic and Agricultural Reforms in Zimbabwe: Policy Complementarities Toward Equitable Growth" by Romeo M. Bautista and Marcelle Thomas (June 2000)

No. 58 - "Updating and Estimating a Social Accounting Matrix Using Cross Entropy Methods " by Sherman Robinson, Andrea Cattaneo and Moataz El-Said (August 2000)

No. 59 - “Food Security and Trade Negotiations in the World Trade Organization : A Cluster Analysis of Country Groups ” by Eugenio Diaz-Bonilla, Marcelle Thomas, Andrea Cattaneo and Sherman Robinson (November 2000)

No. 60* - “Why the Poor Care About Partial Versus General Equilibrium Effects Part 1: Methodology and Country Case’’ by Peter Wobst (November 2000)

No. 61 - “Growth, Distribution and Poverty in Madagascar : Learning from a Microsimulation Model in a General Equilibrium Framework ” by Denis Cogneau and Anne-Sophie Robilliard (November 2000)

No. 62 - “Farmland Holdings, Crop Planting Structure and Input Usage: An Analysis of China’s Agricultural Census” by Xinshen Diao, Yi Zhang and Agapi Somwaru (November 2000)

No. 63 - “Rural Labor Migration, Characteristics, and Employment Patterns: A Study Based on China’s Agricultural Census” by Francis Tuan, Agapi Somwaru and Xinshen Diao (November 2000)

No. 64 - “GAMS Code for Estimating a Social Accounting Matrix (SAM) Using Cross Entropy (CE) Methods” by Sherman Robinson and Moataz El-Said (December 2000)

***TMD Discussion Papers marked with an ‘*’ are MERRISA-related. Copies can be obtained by calling Maria Cohan at 202-862-5627 or e-mail: m.cohan@cgiar.org***