From Data Preparation to Construction of Simulation Models: the QUICKDYME Approach in COMPASS

Wang Yinchu , Senior Economist, Vice Director of the Economic Information Centre of Jiangsu Province , P.R. of China

Abstract:

Econometric input-output model is a kind of very useful economic analysis tool which describes the behaviour of economy not only at macro level, but also at sector level and production branch level. Time series from the System of National Account and input-output tables are the basis in building an econometric input-output model. In reality, to build a such kind of model is not an easy task. For various reasons, the tables probably lack internal consistency among A matrix, final demand and value added components. Tables are available for selected years only, whereas annual series are required for analytical purposes. They are often compiled in current prices, whereas constant price tables are needed for model building. Corresponding figures from the input-output framework and the related data such as the SNA do not necessarily match. There are lots of calculations and operations on matrices, vectors and their components.

The paper proposes to automate the entire process of consistency check of the data to construction of the model and to policy simulation. In particular, the following steps are implemented for various countries.

Step 1, prepare necessary data
Step 2, check the consistency of the data
Step 3, fill the A matrices for the years there are no input-output tables.
Step 4, convert some data from current prices into constant prices.
Step 5, simulate import and get output by solving the equation (1).
Step 6, simulate wages, depreciation and tax minus subsides.
Step 7, simulate the formation of gross output prices and get operating surplus or simulate operating surplus and get gross output prices by solving the equation (2).
Step 8, economic forecasting or policy analysis with different scenarios.

Experience in the construction of country models (China, Japan, other Asian countries) are briefly discussed. It is envisaged to expend the system to cover international trade flows and international linkages of the country models. Energy and environment-related spheres will also be covered. The suggested system incorporates a simulation engine developed by the University of Maryland, but the users are not required to engage in C++ programming as has been the case previously.

From Data Preparation to Construction of Simulation Models:

the QUICKDYME Approach in COMPASS

Wang Yinchu *

Economic Information Centre
Jiangsu Province of China

1. Introduction

Econometric input-output model is a kind of very useful economic analysis tool which describes the behaviour of economy not only at aggregated level, but also at sector and production branch level. Time series from the System of National Account and input-output tables are the basis in building an econometric input-output model. The theoretical framework of this type of model has steps for year t as following:

(1) Determine final demand by production branches in constant price which includes

households' consumption, marked as vector **ctt**,
government expenditure, marked as vector **gtt**,
fixed investment, marked as vector **itt**,
inventory change, marked as vector **ntt**,
export, marked as vector **xtt**,
import, marked as vector **mtt**,

(2) Determine gross output by production branches in constant price which is marked as vector **out** and the calculation formula is

$$A*\mathbf{out} + \mathbf{ctt} + \mathbf{gtt} + \mathbf{itt} + \mathbf{ntt} + \mathbf{xtt} - \mathbf{mtt} = \mathbf{out}, \qquad (1)$$

where A is the technical coefficient matrix in constant price.

---

* Sincere thanks go to Professor, Dean Kimio Uno, Keio University of Japan, who supported and is supporting this research and related international trips. Thanks also go to Dr. Dirk Vanwynsberghe, Economic Research Unit of Belgium, Professor Clopper Almon, University of Maryland of US and Professor Bernd Meyer, University of Osnabrueck of Germany for friend helps from them.

(3) Determine the components of value added in current price by production branches:

depreciation, marked as vector **dET**,
wages, marked as vector **wAT**,
indirect taxes minus subsides, marked as vector **tST**
operating surplus, marked as vector **uRT**,

Therefore, for the value added marked as **vAT** , there is

$$\mathbf{vAT} = \mathbf{dET} + \mathbf{wAT} + \mathbf{tST} + \mathbf{uRT},$$

(4) Determine gross output price indexes by production branches, marked as vector **p_out**, by following formula

$$(A - M)^{T} * \mathbf{p\_out} + M * \mathbf{p\_mtt} + \mathbf{C\_voo} = \mathbf{p\_out}, \qquad (2)$$

where M is the import share matrix of the technical coefficient A, **p_mtt** is the price index vector of import, **C_voo** is the unit value added vector defined as value added in current prices per unit of output in constant prices, i.e., there is

$$\mathbf{C\_voo} = \mathbf{vAT}/\mathbf{out}, \qquad (3)$$

where the operator "/" between two vectors means divided element by element.

(5) Determine other relative variables such as productivity and employment at production branch level. Determine the values of some macro economic variables such as GDP in both current and constant prices, GDP deflator, households' income, consumer's price index, national income, government revenue and so on at aggregation level.

(6) According to some criteria, go back to step 1 to repeat the whole process or go to year t+1.

The basic structure of the econometric input-output model in COMPASS (COmprehensive Model for Policy ASSessment, Uno and Meyer, 1997) is shown in Fig. 1. The model is linked with SNA model and trade model. The components of the final demand, excluding import and export, come from SNA model at aggregation level in constant prices. The vector export and the vector of import prices at production branch level come from the trade model. The IO model dis-aggregates the values which come from SNA model into production branches and calculate the output and import at production branch level by a Seidel iteration procedure on the basis of having estimated equations of the import vector. The import depends upon relative price (defined as domestic gross output price divided by import price) and domestic demand. Next step is to estimate the components of value added at production branch level in current prices and to get unit value added defined as value added in current prices per unit of output in constant prices. Then the IO model determine the domestic gross output price at production branch level according to the formula (2). These results are aggregated into depreciation in total, wages in total, tax minus subsides in total, operating surplus in total, consumer price indexes, GDP deflator. They are the input of the SNA model. The vectors of import demand and the price vector of gross output are the input of the trade model.

It can be seen that to build an econometric input-output model is not an easy task. For various reasons, the tables probably lack internal consistency among A matrix, final demand and value added components. Tables are available for selected years only, whereas annual series are required for analytical purposes. They are often compiled in current prices, whereas constant price tables are needed for model building. Corresponding figures from the input-output framework and the related data such as the SNA do not necessarily match. There are lots of calculations and operations on matrices, vectors and their components. All of these mean that we need a helpful model building software.

There is a software called INTERDYME (A package of Programs for Building Interindustry Dynamic Macroeconomic Models) developed by INFORUM (INterindustry FORecasting at University of Maryland). There are more than 15 countries(Austria, Belgium, China, France, Germany, Italy, Japan, Korea, Mexico, Poland, Russia, Swiss, USA and so on) which are using INTERDYME for building econometric input-output models.

QUICKDYME is a tool which incorporates the INTERDYME and helps users who hope to build multi-sectoral and multi-branch model with at least one input output table and related time series from SNA. The users are not required to engage in C++ programming and to master many INTERDYME commands and concepts as has been the case previously.

Trade model **SNA model**

by branches:
export, **xtt**
import prices, **p_mtt**

aggregated values:
households' consumption,
government expenditure,
investment

---

by production branches, constant prices: **I**
households' consumption, **ctt** **N**
government expenditure, **gtt** **P**
fixed investment, **itt** **U**
inventory change, **ntt** **T**
|
O
by production branches, constant prices: **U**
mtt **out** = A***out + ctt + gtt + itt + ntt + xtt - mtt** **T**
mtt = **f(out + mtt)** or **f(out + mtt - xtt)** **P**
U
T
by production branches, current prices:
depreciation, **dET** **M**
wages, **wAT** **O**
tax minus subsides, **tST** **D**
operating surplus, **uRT** **E**
value added, **vAT = dET + wAT + tST + uRT** **L**

Determine the prices by production branches:
C_voo(j) = **vAT**(j)/**out**(j), j=1,2,…,n.
$(A - M)^T$ ***p_out** + M***p_mtt** + **C_voo** = **p_out**,

Aggregated values:
depreciation in total
wages in total
tax minus subsides in total
operating surplus in total
consumer price index

Fig. 1 The basic structure of the IO model in COMPASS

2. The Characteristics of QUICKDYME

QUICKDYME has following characteristics:

(1) The econometric input-output model built by using QUICKDYME has standard framework in logical structure. However, necessary flexibility, especially in selecting or modifying the form of regression equations, is provided. User can have specific forms for various regression equations when running the model.

(2) The whole process of the model building is divided into several steps from the data preparation to the construction of the model and to policy simulation. At each step, only one main thing (or check data, or extend the A matrix, or convert into constant prices, or one vector variable becomes endogenous) will be done. User can get into one step only if he is successful in previous steps.

(3) Each step of the QUICKDYME corresponds to a sub-directory (or sub-folder) in the directory created by user when installing the software. The sub-directories' name starts from number 1, 2, 3 and so on, according to their order in the whole approach, so that user can easily know what should be done after one successful step.

(4) Various variable's name, including matrix, vector or aggregate variable, will be provided by QUICKDYME according some rules. From the name of a variable, user can easily judge it is a representative of a matrix, a vector or an aggregate variable, which is in current price or in constant price or is an index or a share.

(5) At each step, only several simple text files need to be prepared or modified by the user. After that, one command "ALL" with or without very simple parameters will automatically do whole things of the model building and simulation, including data preparation, regression, C++ programming, compiling and linkage of the model and running the model. To look at the results from simulations, in graphics or tables, needs also only very simple command.

(6) The users are required to master only very few concepts about the modelling engine INTERDYME. Two of them are G bank and VAM bank.

G bank is the data bank in which all the time series of aggregated variables are stored. A program called G.EXE which belongs to the INTERDYME can manage (add, delete, update, type, graph and calculate) the data in G bank and can do regression on these time series.

VAM bank is the data bank in which all the time series of vectors and matrices are stored. A program called VAM.EXE which belongs to the INTERDYME can manage (add, delete, update, type, graph and calculate) the data in VAM bank.

3. The steps included in current version of QUICKDYME.

There are 8 steps in the current version of QUICKDYME. They sit in 8 sub-directories and each sub-directory corresponds one step, respectively. These sub-directories are:

\1data—prepare necessary data
\2check—check the consistency of the data
\3fillam—fill the A matrices for the years there are no input-output tables.
\4conpri—convert some data from current prices into constant prices.
\5mttsim—simulate import and get output by solving the equation (1).
\6wagesim—simulate wages, depreciation and tax minus subsides.
\7pricsim—simulate the formation of gross output prices and get operating surplus or simulate operating surplus and get gross output prices by solving the equation (2).
\8forcast –economic forecasting or policy analysis with different scenarios.

The contents of each step are described as following:

(1) step 1 (sub-directory \1data)

Required data files for matrices, vectors and aggregated variables which are necessary in building QUICKDYME model are prepared according to required format in this step. All of these data files are text files and directly come from reported input-output table(s), SNA reports(statistics year books). It means most of the data, especially the input output data, are in current prices. Table 1 listed all the necessary data (matrices, vectors and aggregate variables) and their names for the current version of QUICKDYME.

Table 1. The Necessary Data for Current QUICKDYME

	Content of the Variable	Real or Nominal	Name in Quickdyme
matrix	Technical coefficients	current price	A
matrix	Import share of matrix A		M
vector	Output	current price	oUT
vector	Consumption of households	current price	cTT
vector	Government expenditure	current price	gTT
vector	Fixed investment	current price	iTT
vector	Inventory changes	current price	nTT
vector	Exports	current price	xTT
vector	Imports	current price	mTT
vector	Depreciation	current price	dET
vector	Wage rate	current price	wAR
vector	Wages	current price	wAT
vector	Taxes	current price	tAT
vector	Operating surplus	current price	uRT
vector	Subsidies	current price	sUT
vector	Index of import prices		p_mtt
vector	Index of gross output prices		p_out

Note: (1) The matrices A and M are in current prices firstly and they will be in constant prices from step 4. To do so is for saving space in the approach. **(2)** In the case there are no data of import share matrix M, QUICKDYME will create it according to the ratio between import and domestic demand by branches.

(2) step 2 (sub-directory \2check)

The objectives of this step are:

I. To create a G bank in which the historic data of aggregated variables are stored.

II. To create a VAM bank in which the historic data of matrices and vectors are stored.

III. To check the consistency among the historic data introduced into G bank and VAM bank.

The consistency consists of four aspects:

A. Whether the equation

$$A*\mathbf{oUT} + \mathbf{fIM} = \mathbf{oUT}, \qquad (3)$$

is held for the I-O table report years or not, where matrix A is in current prices and

$$\mathbf{fIM = cTT} + \mathbf{gTT} + \mathbf{iTT} + \mathbf{nTT} + \mathbf{xTT} - \mathbf{mTT},$$

B. Whether the equation

$$A^{T}*\mathbf{oUT} + \mathbf{vAT} = \mathbf{oUT}, \qquad (4)$$

is held for the I-O table report years or not, where matrix A is in current prices.

C. Whether the relationship

$$\sum_{i=1}^{n} \mathbf{fIM}(i) = \sum_{\mathrm{j}=1}^{\mathrm{n}} \mathbf{vAT}(j)$$

is held for every historical year or not.

D. Whether the relationship

$$\mathbf{vAT}(j) < \mathbf{oUT}(j), \qquad j=1,2,\ldots,n$$

is held for every historical year or not.

(3) step 3 (sub-directory \3fillam)

The A matrices for the years when the input-output tables are not available are created firstly in this step. Then, these created matrices are modified by using a procedure called "ACROSS" in order that they can have consistency with other data such as output, final demand and value added.

The basis of the ACROSS procedure is the equation (3) and (4) and it includes two directions' modifications, rows and columns, on the A matrix.
First is the row adjustment which is based on equation (3). Calculate the row adjustment factors (a vector, called **amitemp**)

$$\mathbf{amitemp} = (\mathbf{oUT} - \mathbf{fIM}) / A^{*}\mathbf{oUT}, \qquad (5)$$

and then let

$$A(i,j) = A(i,j)^{*}\,\mathbf{amitemp}(i), \qquad j=1,2,...,n, \quad i=1,2,...,n. \qquad (6)$$

The reason to do so is both the (**oUT - fIM**) and (A***oUT**) in (5) are the total intermediate demand vectors and they should be the same if the A matrix is a right one. The formula (6) is to reach this consistency by modifying the A matrix row by row.

Another consistency, the equation (4), also needs to be checked and reached. From (4), we should have

$$\sum_{i=1}^{n} A(i,j) = 1 - \mathbf{vAT}(j)/\mathbf{oUT}(j), \qquad j = 1,2,...,n. \qquad (7)$$

So, after the row adjustment, a column scale operation should be done. However, this will probably destroy the balance reached through the row adjustments and a new row adjustment operation is needed, and so on. This iteration approach, from row adjustment to column scale, will converge if the data are qualified.

(4) step 4 (sub-directory \4conpri)

The technical coefficient matrices A, the vectors of output, import and all the vectors which are belong to final demand side are converted from current prices into constant prices in this step so that the equation (1) and (2) can be used in further steps of modelling. There are only simple calculations by using the vector of gross output prices, **p_out**. For example, the conversion of A matrix from current prices into constant prices is

$$A(i,j) = A(i,j)*\mathbf{p\_out}(j)/\mathbf{p\_out}(i), \quad i, j = 1,2,\ldots, n.$$

(5) step 5 (sub-directory \5mttsim)

In this step, the import vector, **mtt**, become endogenous variable. There are 2 different basic regression forms of import equation which can be selected by user. They are:

$$\mathbf{mtt} = \mathbf{a}*(\mathbf{p\_out}/\mathbf{p\_mtt}) + \mathbf{b} * (\mathbf{out} + \mathbf{mtt}) \quad (8)$$
$$\mathbf{mtt} = \mathbf{a} *(\mathbf{p\_out}/\mathbf{p\_mtt}) + \mathbf{b} * (\mathbf{out} + \mathbf{mtt} - \mathbf{xtt}) \quad (9)$$

where **out** is gross output, **xtt** is export, **a** and **b** are the parameters to be estimated, the operator * is multiple element by element. Equations above say that the import depends upon relative prices and domestic demand. The difference between Equation (8) and (9) is the definition of domestic demand. In equation (9), the domestic demand excludes export. Which kind of import equation is proper? The user can compare the results and make choice from running the model. User also has flexibility to further modify the import equation forms when running the model by using equation fixer function of the software.

The export, vector **xtt,** comes from trade model and the other components of the final demand by branches, **ctt, gtt, itt** and **ntt** come from SNA model at aggregated level and their shares which are determined in exogenous. After that, the output, vector **out**, is determined from solving the equation (1) by using a Seidel approach.

(6) step 6 (sub-directory \6wagesim)

In this step, the depreciation by branches (vector **dET**), the tax minus subsides by branches(vector **tST**) and the wages by branches(vector **wAT**) become endogenous variables. The regression equations have following basic forms and they can be modified flexibly by user when running the model.

1. dET = α ***out,**

2. **tST =** β ***out,**
3. **wAR =** wAR[gdp(t-1)/totwhr(t-1)**,**cTTSIO(t-1)/cttsio(t-1)],
   where gdp is gross domestic product,
   cTTSIO is total households' consumption in current prices,
   cttsio is total households' consumption in constant prices.
   totwhr is total working hours which comes from vector **whr**, working hours by branches. The vector whr comes from

4. **whr** = whr[**wAR/p_out**, **out,** t]
   where t is a time trend.

(7) step 7 (sub-directory \7pricsim)

In this step, prices of gross output by branches (vector **p_out**) and operating surplus by branches (vector **uRT**) become endogenous variables. There are two optional approaches which user can chose. One is to determine the gross output price vector firstly by regression on unit cost and then get operating surplus as a kind of residual. Another one is to determine the operating surplus firstly by regression on unit cost, therefore the value added and the unit value added, and then get gross output from solving the equation (2) by a Seidel procedure. Of course, all the regression equations can be modified flexibly by user when running the model.

The unit cost is defined as

$$\mathbf{uCT} = (\text{A-M})^T\,\mathbf{p\_out} + \text{M}^T\,\mathbf{p\_mtt} + \mathbf{wAT/out} + \alpha + \beta$$

If the first approach is selected, the prices of the gross output will be calculated according to regression equation

$$\mathbf{p\_out} = \mathbf{a} + \mathbf{b*uCT}$$

and the operating surplus is calculated as a residual:

$$\mathbf{uRT} = \mathbf{p\_out*out - uCT*out}$$

If the second option is selected, the unit operation surplus will be firstly determined by regression equation

uPT = **a** + **b*uCT**

where **uPT** = **uRT/out**

and the prices of gross output will be determined by solving the equation (2).

User can decide which approach is proper through comparing the results from running the model with different options.

(8) step 8 (sub-directory \model)

After selecting a proper form of import equation and proper approach of price determination, economic forecasting or policy simulation under different scenario can be done in this step.

The values of the exogenous variables in different scenario are given in text file **scenario.%1,** where **%1** is a parameter in the file's name and it can be **a** or **b** or **c** or **d** and so on. Therefore, file **scenario.a** is corresponding to scenario **a,** file **scenario.b** is corresponding to scenario **b,** and so on.

The command to run the model has one simple parameter which points out the scenario to be used in the model. For example, to run the model with scenario **a,** just enter a command

```
ALL  a
```

To run the model with scenario **b**, just enter a command

```
ALL  b
```

and so on.

4. The Applications of QUICKDYME

QUICKDYME has been used to build country model for China, Japan, US, Korea, Thailand, Philippine, Singapore, Indonesia, Malaysia and other countries. After the preparation of required data (step 1), the step 2 helped a lot to find out consistency problem exited in the data. After correcting the errors in data, user can easily go through step 3 to step 8 in one day if he (or she) pays full attention to the model construction in that day.

Notice that the country models built up by QUICKDYME

A. describe the behaviour of the economy at sector level and production branch level
B. have same variable names,
C. have import prices in the pricing approaches,

therefore, these country models can easily be used for the analysis of energy production and consumption because there is detail related information from the technical coefficients and the amount of production and consumption. It can also be used for the analysis of environment issue such an the CO2 emission connected to the energy consumption and the technical conditions. It is also easy to use these country models in an international linkage system. In fact, some of the country models developed by QUICKDYME have been used in a pilot linkage research which belongs to the APEC Economic, Energy and Environment-related linkage project.

QUICKDYME is still in developing. On one hand, there are many other steps should be added into the whole approach. On the other hand, it should run under windows so that it will be more flexible and more easy to use.

References:

1. Bernd Meyer and Kimio Uno, “The Theoretical Structure of the Internationally Linked Dynamic Input-Output-Model COMPASS”, Paper presented at the Fifth INFORUM World Conference, September 1997, Italy.

2. Dirk Vanwynsberghe and Wang Yinchu, “Naming Convention, the QUICKDYME Models”, Keio University at Shonandai Fujisawa Campus, February, 1997.

3. INFORUM (Interindustry Forecasting at University of Maryland), “INTERDYME, A Package of Programs for Building Interindustry Dynamic Macroeconomic Models”, University of Maryland, December, 1995.

4. Kimio Uno, “Policy application of Global Economic-Energy_Environment(3E) accounting and modelling”, Keio University, January, 1998.

5. Wang Yinchu and Dirk Vanwynsberghe, “QUICKDYME, Country Model Software, User’s Manual”, Keio University at Shonandai Fujisawa Campus, February, 1997.

6. Wang Yinchu, “Price Equation for the National Model Linked with Trade Model”, Economic Information Centre of Jiangsu Privince, August, 1997.

7. Wang Yinchu, “An Approach to Determine Gross Output Prices from Value Added Prices”, Keio University, December, 1997.

8. Zuo Li, “The APEC Trade Linkage Model, Concepts and Preparations”, Paper presented at the fifth INFORUM World Conference, September 1997, Italy.

9. Zuo Li and Wang Yinchu, “The Treatment of the Discrepancy between the Systems of Input Output and National Accounting”, Economic Information Centre of Jiangsu Province, November, 1997