The Treatment of the Discrepancy between the Systems of Input Output and National Accounting

By Zuo Li and Wang Yinchu
(Economic Information Centre of Jiangsu Province, China)

For the year there is Input Output Table, the consumption of households, the government expenditure, the fixed investment, the inventory change, the value added and the output which come from the System of Input Output (SIO) are normally not the same as the numbers reported by the System of National Accounting (SNA). The differences are called discrepancy between the two systems. For example, the discrepancy situation of the year 1987 for China is shown in Table 1.

Table 1. The difference between SIO and SNA

ITEM	SIO	SNA
Consumption of Rural Households	3498.43	3630.1
Consumption of Urban Households	2444.84	2331.1
Public Consumption	1329.90	1490.0
Fixed Investment	3802.60	3742.0
Inventory Change	570	580
Value added	11424.1	11925.5

How to deal with the discrepancy between the two systems is a problem in building up an INFORUM model because, for one same variable such as fixed investment, not only the data from input output table but also the time series from SNA are simultaneously used in that kind of model.

There are four different options to deal with the discrepancy problem:

1. Adjust the annual statistics, just for the I/O report year, according to the data from input output table.

2. Adjust the annual statistics, for every year of the study period, according to the ratio between data from annual statistics and data from input output table in the I/O report year.

3. Adjust the input output table, without adjustment for the technical coefficient matrix, according to the annual statistics for the I/O report year.

4 Adjust the input output table, with adjustment for the technical coefficient matrix, according to the annual statistics for the I/O report year.

Which option is suitable ? A comparison of them was done for MUDAN (MUltisector Development ANalysis model for China).

OPTION 1. **Adjust the annual statistics, just for the I/O report year, according to the data from input output table.** The reason to do so is the input output is the basic framework of an INFORUM model and people hope to keep various balance relationships in the input output table.

In this case, the time series of many variables, especially their growth rates, which is directly or indirectly related to the adjusted variables will probably have jumping or dropping points comparing the actual values.

Table 2 and Figure 1 is the comparison of the growth rate of GDP after the adjustment for the discrepancy in 1987 according to the option 1.

Table 2 The Growth Rate of GDP

YEAR	1986	1987	1988
Model with Option 1	13.81	12.00	30.67
Actual Value	13.81	16.89	25.20

Figure 1. The Growth Rate of GDP

It can be found that the growth rates in 1987 was changed from 16.89%, which is higher than the one in the previous year (1986, 13.81%), into 12%, which is lower than the one in the previous year. This change is not only in number, but also in concept which is higher or lower than the previous year. It is not an acceptable result for many people who know the SNA statistics very well, specially, for most of the economist who do macro economical analysis every year.

It is also not good for the estimation of the behaviour equations of the variables adjusted with option 1 because the consistency of the variable in time series will be lost after that kind of adjustment.

OPTION 2. Adjust the annual statistics, for every year of the study period, according to the ratio between data from SNA and data from SIO in the Input Output Table report year. The reason to do so is to avoid the disadvantage of the option 1 so that the growth rates can keep a consistency with the actual situation.

However, a new problem occurs: all the relative variables will have values lower (or higher) the actual ones and it looks as there is a system error in the model. A comparison between the actual values of GDP and the GDP values from the model with option 2 is shown in Table 3 and Figure 2.

Table 3. The values of GDP

YEAR	ACTUAL	MODEL
1980	4517.800	4408.748
1981	5181.100	5040.512
1982	5844.401	5672.276
1983	6507.700	6304.037
1984	7171.001	6935.802
1985	8964.399	8600.901
1986	10202.20	9785.134
1987	11925.50	11424.07
1988	14928.30	14278.66
1989	16909.20	16106.30
1990	18530.70	17665.58
1991	21617.80	20519.25
1992	26635.40	25255.56
1993	34515.10	32907.05
1994	45005.79	42962.54

It can be found that the values of GDP from the model with option 2 were about 5% lower than the ones from the actual statistics in each year. It was due to the adjustment of option 2. It is also not good for the people, especially most of the economists or the government staffs, who know the SNA statistics well,

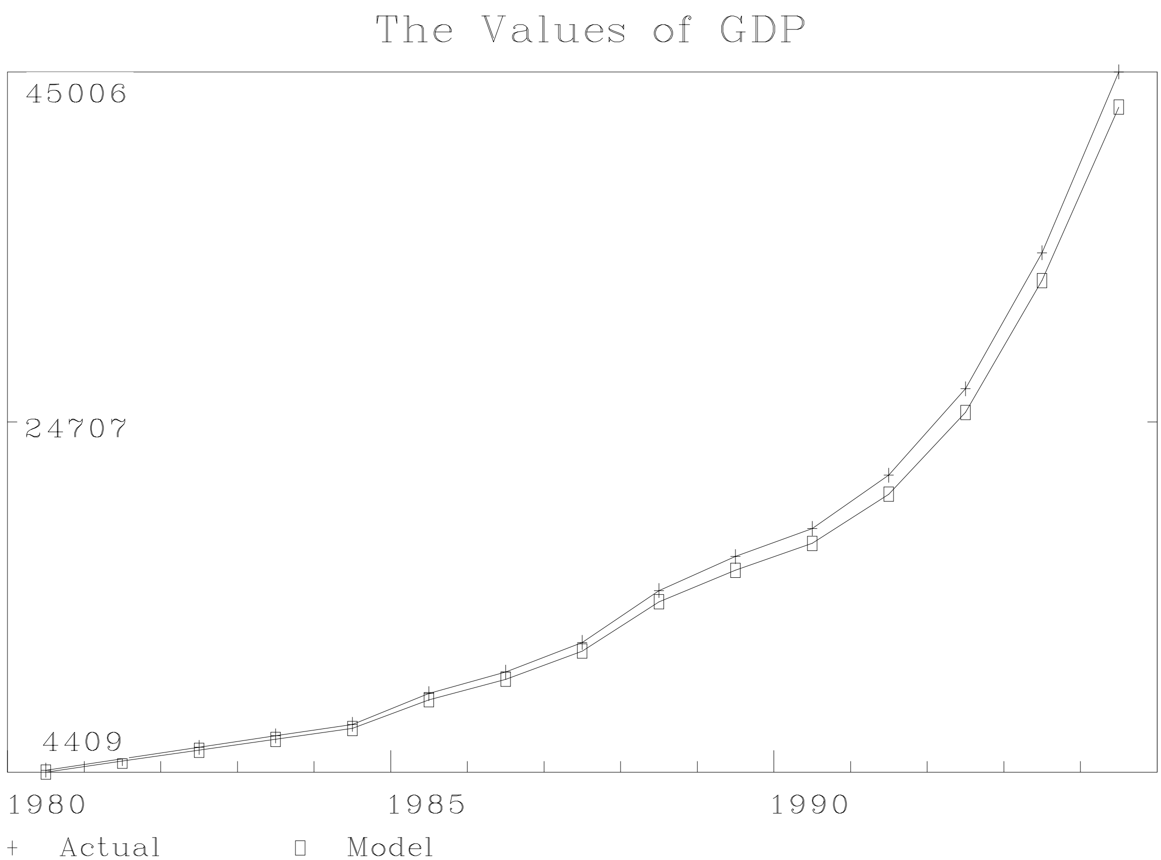

Figure 2. The Values of GDP

OPTION 3. **Adjust the input output table, without adjustment for the technical coefficient matrix, according to the annual statistics for the I/O report year.** It seems there is always problem to adjust the SNA data based on the SIO data, so we consider another direction: to adjust the SIO data on the basis of SNA data. A first scheme is to have a discrepancy column in the second quadrant and a discrepancy row in the third quadrant of the Input Output Table so that the output, the final demand, the import, the value added and their sub-items in the table can have values which are consistent with the data from SNA.

Let:

fIMSIO	is the sum of finale demand minus import from SIO
oUTSIO	is the output from SIO
vADSIO	is the value added from SIO
A	is the technical coefficient matrix from SIO
A™	is the transfer of A matrix
fIMSNA	is the sum of finale demand minus import from SNA
oUTSNA	is the output from SNA
vADSNA	is the value added from SNA
ΔoUT	is equal to oUTSIO - oUTSNA
ΔfIM	is equal to fIMSIO - fIMSNA
ΔvAD	is equal to vADSIO - vADSNA

and all of them are in current prices.

For the SIO data, there are

(1) $A*oUTSIO + fIMSIO = oUTSIO$

(2) $A^{TM}*oUTSIO + vADSIO = oUTSIO$

Replace the SIO data in (1) and (2) with SNA data, there are

(3) $A*(oUTSNA+\Delta oUT) + fIMSNA+\Delta fIM = oUTSNA + \Delta oUT$

(4) $A^{TM}*(oUTSNA+\Delta oUT) + vADSNA+\Delta vAD = oUTSNA + \Delta oUT$

and we get

(5) $A*oUTSNA + fIMSNA + dRW = oUTSNA$

(6) $$A^{TM}*oUTSNA + vADSNA + dCL = oUTSNA$$

where

(7) $$dRW = A*\Delta oUT + \Delta fIM - \Delta oUT$$

(8) $$dCL = A^{TM}*\Delta oUT + \Delta vAD - \Delta oUT$$

are the row discrepancy vector and column discrepancy vector, respectively.

The formula (5) - (8) are used for the year(s) when the input output table(s) is(are) available.

As we know, however, it is necessary to have the A matrices for the years when the input output table are not available in building an INFORUM model and there is an approach called “Across the Row” to have them. In this approach, two basic formula used as a kind of balance control are:

(9) $$A*oUTSNA + fIMSNA = oUTSNA$$

(10) $$A^{TM}*oUTSNA + vADSNA = oUTSNA$$

Is it good to use the formula (9) and (10) for the years when there are no input output tables and the formula (5) and (6) for the year when there is input output table ? It seems there is another kind of consistency problem: If the time series of fIMSNA, oUTSNA and vADSNA are consistent each other, the time series of the A matrices created by the formula (9) and (10) are probably not consist with the A matrix from SIO. From this consideration, we go further to have option 4.

OPTION 4. **Adjust the input output table, with adjustment for the technical coefficient matrix, according to the annual statistics for the I/O report year.** We think about the formula (9) and (10). They are used for creating the A matrices for the years when there are no input output tables, according to the data of oUTSNA, vADSNA and fIMSNA which are consistent with each other from the point of view of time series. For keeping consistency, can we also use the formula (9) and (10) for the year when the input output table is existing so that all the A matrices come from the same approach and should then have consistency ? In other words, the formula (9) and (10) will be used not only for creating A matrices for the years when the input output tables are not available, but also for modifying the exiting A matrix for the year when the input output table is available. The A matrix from the existing input output table is just as a kind of initial estimation of the technical coefficient matrix.

How to evaluate the results from the option 3 and. the option 4? To look at the time series of the elements of the A matrix. A reasonal way to judge them is the consistency of the values in the time series . In our experiment, there is only one possible different point in the two results which is the value in year 1987 when there is Input Output Table. We just need to look at the situation of the value in this year. Among 3036 of non zero elements of the A matrix(60 sectors), there are four different type of situations.

Situation (1): The value in year 1987 from option 4 looks obviously better than the one from option 3, as showing in the figure 3 for the element A(1,3). In figure 3 (and fig. 4, 5 and 6), the points connected by + is from option 4 and the points connected by square is from option3. The value from option 3 jumped in the year 1987 shows its inconsistency.

Among the 3036 non-zero elements, the percentage of the elements with the situation (1) is around 28.6%.

Figure 3. Option 4 is obviously better than option 3

Situation (2): It seems not very clear to say which value in year 1987 is better, as showing in the figure 4 for element A(12,3). However, we could still say that the result from option 4 is a little bit better than the one from option 3. The values from option 3 jump up in 1987 then drop down in 1988. The values from option 4 go up a little in 1987 then go up again in 1988. From the point of view of the range of the change, it seems that we can still say option 4 is better than the option 3.

Among the 3036 non-zero elements, the percentage of the elements with the situation (2) is around 10%.

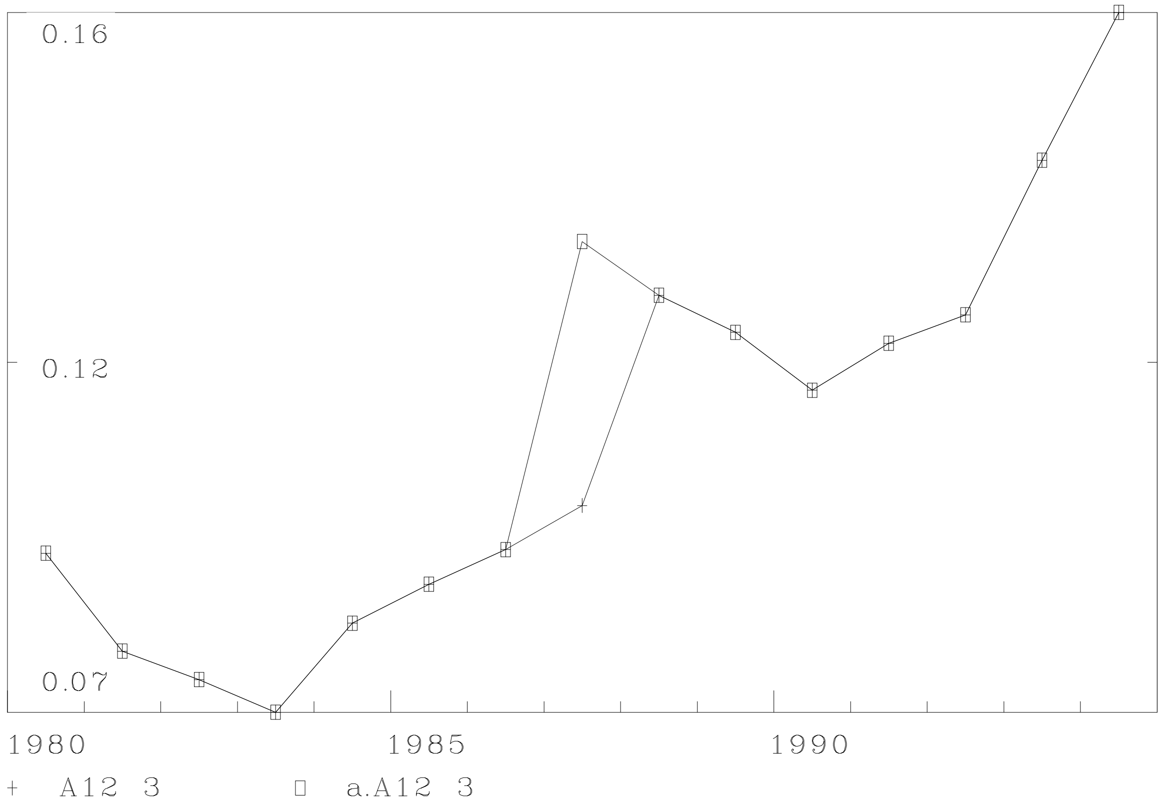

Figure 4. It is not very obvious who is better

Situation (3): The value in year 1987 from option 4 looks the same as the one from option 3, as showing in the figure 5 for the element A(1,13). We can see there is no big difference in this case.

Among the 3036 non-zero elements, the percentage of the elements with the situation (3) is around 58.8%.

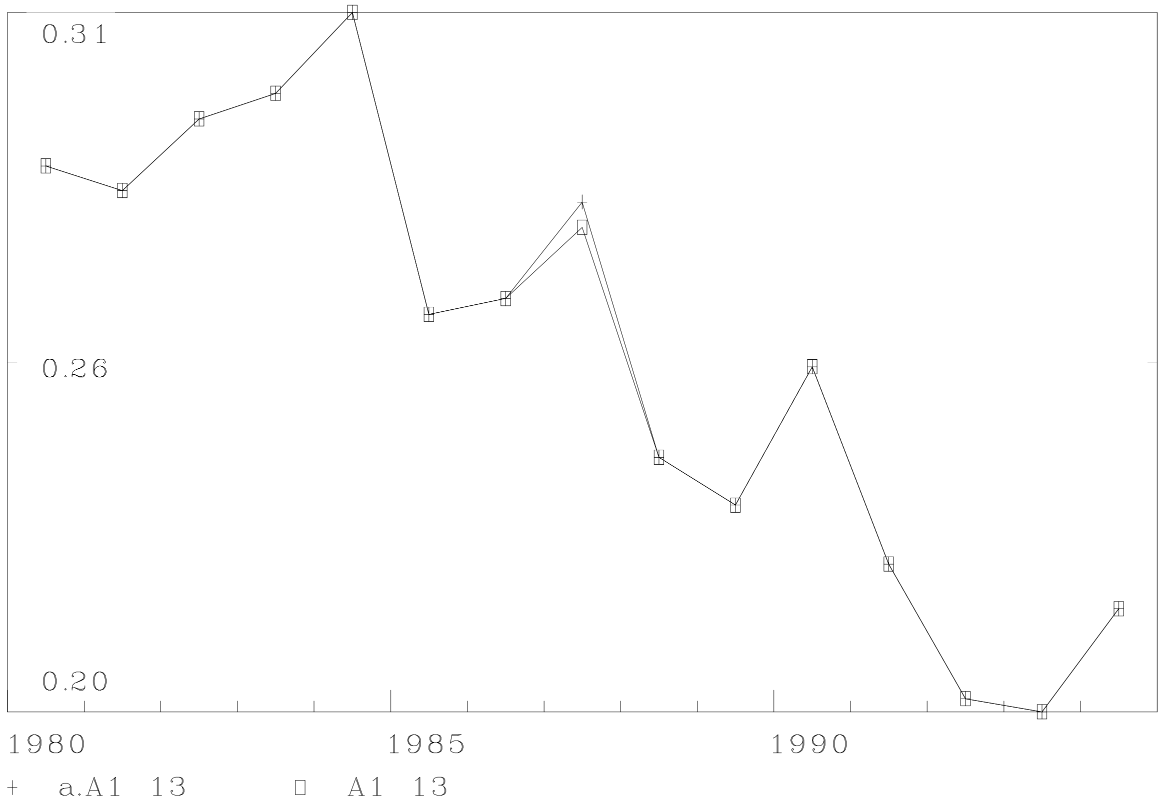

Figure 5. The results are very closed

Situation (4): The result from option 3 is a little bit better than the one from option 4 as showing in Fig. 6 in which the values of. the variable A(5,8) are multiplied by 100 because the software used for creating the graph can only display two digits after the point.

Among the 3036 non-zero elements, the percentage of the elements with the situation (4) is around 2.6%.

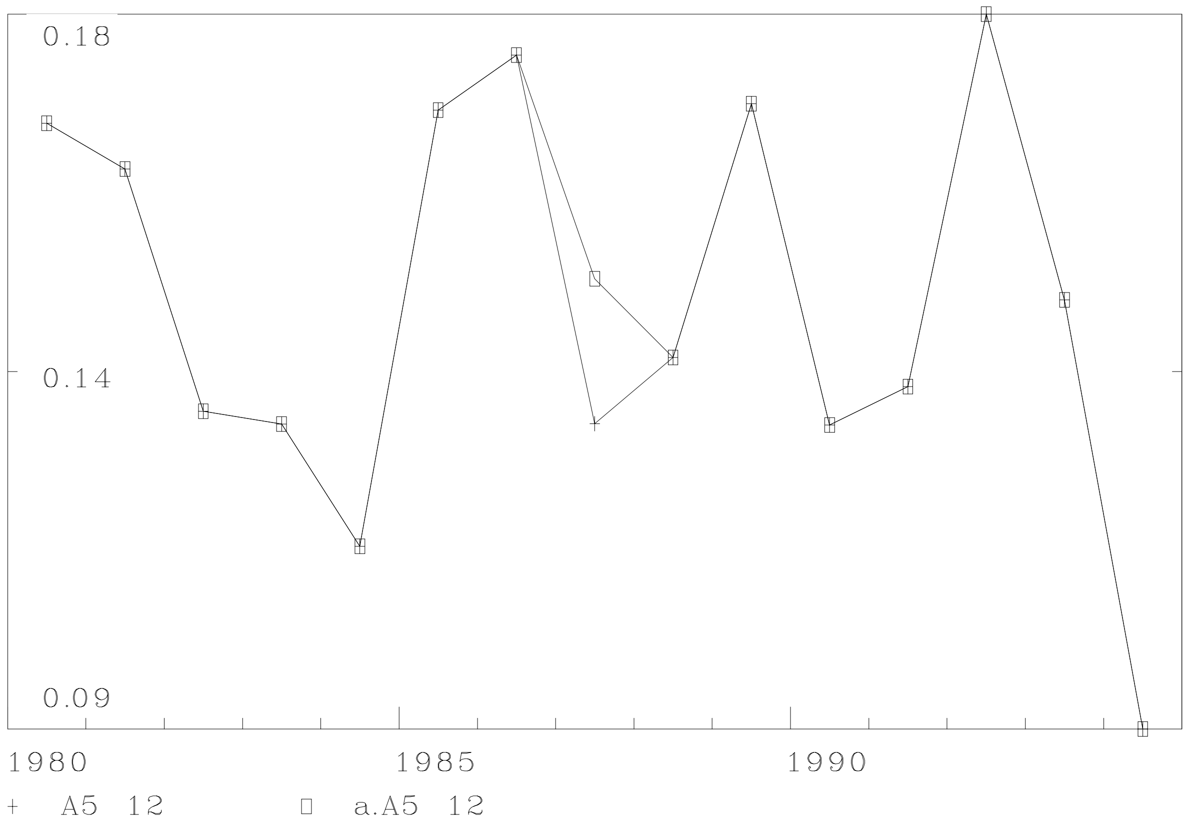

Figure 6. Option 3 is a little bit better than option 4

Conclusion:

Comparing the results from the four situation , we get conclusion that the option 4 is a better way to have the time series of the A matrix.

Reference:

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

(2). Almon Clopper, “MUDAN: A model for Multisectoral Development Analysis of China”, INFORUM working paper, Department of Economics, University of Maryland, 1994.

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