



Social Accounting Matrix Exercises: Module M1

cgemod

These exercises focus on developing an understanding of the principles of social accounting and Social Accounting Matrices (SAMs). The first exercise is basic exercise in interpreting the information in a SAM. The second and third exercises are applications of the SAM-Leontief model to derive income and price multipliers that further expose the information found in SAMs. The final, fourth, exercise is a first data collection exercise in preparation for subsequent modules.

Table of Contents

1.	Introduction	2
2.	The Course Materials	4
2.	Ex M1.1: Constructing SAMs from T-Account Data	5
4.	Ex M1.2 SAM Interpretation Exercise.....	9
5.	Ex M1.3 SAM Income Multiplier Exercises	12
6.	Ex M1.4 SAM Price Multiplier Exercises.....	15
7.	Ex M1.5 Collecting National Accounts Data	17

1. Introduction

The course design is based on the principle of learning-by-doing. Accordingly, during the course the participants will develop a SAM for their chosen country. Ideally, the country for which the SAM will be developed will be the participant's own country or a country where the participant is studying. Participants will need access data for the chosen country, typically aggregate national accounts data, aggregate tax instrument data, Supply and Use tables, household income and expenditure survey data and labour force survey data.

This module, M1, is concerned with the theory and understanding of Social Accounting and the role of SAMs as accounting systems used to calibrate whole economy models. The exercises for this module emphasise the interpretation of the information content in SAMs using coefficients and simple SAM-Leontief models. The exercises conclude with the collection of basic aggregate national accounts data, aggregate tax instrument data, Supply and Use tables, household income and expenditure survey data and labour force survey data. You will find that the databases you first collect are incomplete and almost invariably unreconciled. You will also likely find that the databases are for different years; national accounts agencies rarely have the resources to conduct all surveys for each year.

The SAM exercises are formulated to help you develop your understanding of SAMs; the exercises are not exhaustive, and they are not a substitution for more extensive study of SAMs.

The first set of exercises (M1.1) involves the generation of two simple SAMs from a series of T-Accounts. The first SAM is a so-called 'macro' SAM that uses simple macroeconomic data to derive a basic SAM; this is an example of a SAM that might be derived at the start of a process that is undertaken when deriving a SAM for a country using a top-down process. The second SAM is a so-called 'micro' SAM that uses disaggregated data to derive a basic SAM; this is an example of disaggregating a 'macro' SAM. The data in the 'micro' SAM aggregate up to the data in the 'macro' SAM; all the data used are fictional.

The second set of exercises (M1.2) involves interpreting the information in a SAM: if you understand and can interpret the information contained in a SAM you can usually understand why even complex models produce the insights the models provide. This is complemented by two further exercises (M1.3 and M1.4) that involve deriving fixed price

multipliers and then interpreting the information they provide. Although fixed price models are now dated, they remain a very valuable, and underused, tool for evaluating the information content of SAMs.

The final exercise is an introductory guide to the data you will need to collect to develop a SAM.

There is no defined end to the course, implicitly the course ends when you have compiled a preliminary SAM, although it is highly unlikely to be complete. The reason is simple. The process of compiling a SAM takes time and depends on the immediate accessibility of certain necessary data. However, participants should have a good understanding of what is necessary to compile a completed SAM along with an understanding of estimation techniques that can be used to ease the task.

The demonstration materials for the course use data from various sources to illustrate the processes. The demonstration materials for basic estimation and reconciliation techniques will use data from several sources; all these constructed data are supplied.

2. The Course Materials

This course does not use a GAMS model library to organise the correct files for each exercise. The computer/code/data files will be downloaded from the cgemod site (www.cgemod.org.uk/sames.html) as they are required: however, it is recommended that course directory is formed, e.g., C:\cgemod\samest_lib, to store the downloaded materials.

The data for this module are in two workbooks 'SAM T-Account exercise.xlsx' and 'SAM exercises M1.xlsx' on the course website. You should download these workbooks to C:\cgemod\samest_lib\M1\. Files with worked examples are available from the website.

RESIST THE TEMPTATION TO EXPLORE THE CONTENTS OF THE SUBSEQUENT EXERCISE FILES PRIOR TO COMPLETING THE PREVIOUS EXERCISES; THIS WAY LIES CONFUSION. YOU GET TO SEE AND USE ALL THE FILES.

All the computer/code/data files used in this course are in GAMS and/or Excel.

FOR THIS COURSE YOU NEED A GAMS LICENCE THAT INCLUDES THE CONOPT 4 SOLVER. IT IS ALSO HELPFUL TO HAVE THE PATH/PATHNLP AND KNITRO SOLVERS.

No GAMS licenses are supplied. For GAMS licenses contact GAMS (https://www.gams.com/buy_gams/) or email sales@gams.com.

2. Ex M1.1: Constructing SAMs from T-Account Data

This set of exercises are about the process of constructing a SAM, i.e., a single-entry accounting system, from data presented as T-accounts, i.e., a double-entry accounting system. The workbook 'SAM T-Account exercise.xlsx' has two worksheets: 'Macro SAM Exercise' and 'Micro SAM Exercise'. The first exercise uses the worksheet 'Macro SAM Exercise' and the second uses the worksheet 'Micro SAM Exercise'.

Both worksheets contain a template/framework for the SAM to be constructed and a series of T-Accounts that will be used to construct the SAM. The worksheets use a 'freeze panes' view: in the top frame is the template and in the bottom frame are the T-accounts. This arrangement allows you to scroll through the T-accounts while still viewing the SAM; you may need to adjust the zoom setting (bottom right-hand corner of the screen) on your PC to keep everything in view. The screen shot in Figure 1 shows what you should see – line 16 and below is the frame with the T-accounts, line 15 and above is the frame with the SAM.

The exercise involves deciding which values in the T-accounts should be entered in which cells in the SAM. The simplest way to do this is to use formulas that link the cells in the T-accounts with the cells in the SAM. This can be done in three different ways:

1. Incomes: use the income side of each T-account to determine the row entries in the SAM.
2. Expenditure: use the expenditure side of each T-account to determine the column entries in the SAM.
3. Hybrid: use a mix of the income and expenditure sides of selected T-accounts to determine the entries in the SAM.

Implicitly options 1 and 2 work from the T-accounts to the SAM while option 3 works from the SAM to the T-accounts. Note that the row and column total must equate; if they do not the SAM will not be complete¹, although they may (rarely) be consistent².

¹ A SAM is said to be complete if it includes all transactions in an economy.

² A SAM is said to be 'consistent' if all the transactions by agents are reconciled, i.e., an income to one account has the same value as an expenditure by another account. If it is complete and consistent the row and column totals will equate. The row and column totals may equate for a SAM that is not 'complete', i.e., the transactions are 'consistent', but this usually requires that one or more transactions do not have correct values. Thus, the fact that the row and column totals equate does not necessarily mean that a SAM

Figure 1 Exercise Worksheet Layout

Macro SAM Template

	Commodities	Activities	Factors	Households	Enterprises	Government	Savings	RoW	Total
Commodities									
Activities									
Factors									
Households									
Enterprises									
Government									
Savings									
RoW									
Total									

T- Accounts

Commodities

Incomes		Expenditures	
Agent	Value	Agent	Value
Activities	265	Activities	715
Households	305		
Government	100	Government	80
Savings	65	Savings	0
RoW	165	RoW	105
Total	900	Total	900

Activities

Incomes		Expenditures	
Agent	Value	Agent	Value

A Macro SAM

For this exercise select one of the ways to derive the Macro SAM and map the values for each cell in the SAM from the respective value in the T-account. Use exactly the same option for ALL transactions; if you chop and change between options you are likely to take longer and risk getting confused. HINT: for option 1 work by rows starting at the top, for option 2 work by columns starting on the left, and for option 3 use rows or columns or work down the SAM from top left to bottom right.

The row and column totals should be:

is ‘complete’ and ‘consistent’, whereas if a SAM is ‘complete’ and ‘consistent’ the row and column totals will equate.

Table 1 **Macro SAM Account Totals**

Commodities	900
Activities	715
Factors	385
Households	392
Enterprises	80
Government	197
Savings	65
RoW	145

Now repeat the exercise using a different option. Note how you can determine the transaction value for any cells of the SAM from either an income or an expenditure value in a T-account. This is useful since all the data available to generate SAMs have been estimated with some degree of error, so comparisons of estimates of data points from two different perspectives – income v expenditure side – assist when making judgements about data reliability. This can be especially useful when working with certain data points, e.g., tobacco and alcohol consumption is regularly under recorded in household expenditure surveys, so supply side estimates are particularly important.

In this example, the T-accounts for the macro-SAM transactions were fully reconciled. However, when trying to construct a macro-SAM from specific (macro) national accounts data it is often the case that the T-accounts are not fully reconciled and/or that the macroeconomic T-accounts are incomplete.

Compare your macro-SAM with those reported in the workbook ‘SAM T-account solution.xlsx’. These were constructed using option 1 and option 2 and the cells in the SAM are linked by formulae to the T-accounts.

A Micro SAM

For this exercise select one of the ways to derive the Micro SAM, ideally not the ways your first chose for the macro-SAM, and map the values for each cell in the SAM from the respective value in the T-account. Use the same option for ALL transactions; if you chop and change between options you are likely to take longer and risk getting confused. **HINT:** for option 1 work by rows starting at the top, for option 2 work by columns starting on the left, and for option 3 use rows or columns or work down the SAM from top left to bottom right.

The row and column totals should be:

Table 2 **Micro SAM Account Totals**

Primary	310
Secondary	590
Agriculture	235
Industry	480
Labour	250
Capital	135
Urban	215
Rural	177
Enterprises	80
Government	197
Savings	65
RoW	145

Now repeat the exercise using a different option. Note how the sum of related transaction values, e.g., intermediate input demand, payment by activities to factors, etc., in the micros SAM equal the aggregate values in the macro-SAM. This is useful since all the data available to generate SAMs have been estimated with some degree of error. If the sum of these groups of transaction values in the micro-SAM do not equal the aggregates in macro-SAM, the compiler of a SAM will need to determine the extent to which the transactions in the micro-SAM need ‘adjusting’ to be consistent with the macro-SAM aggregates or some mix of both. Inevitably this is a judgement call that must be based on some opinion about the reliability of the different components of the underlying data. Since there is strong evidence that certain data points are very unreliable, e.g., tobacco and alcohol consumption estimates in household expenditure surveys, the decision may be straightforward. In other cases, the decision is more difficult.

Compare your macro-SAM with those reported in the workbook ‘SAM T-account solution.xlsx’. These were constructed using option 1 and option 2 and the cells in the SAM are linked by formulae to the T-accounts.

4. Ex M1.2 SAM Interpretation Exercise

Botswana SAM Interpretation Exercise

The interpretation exercise using the macro-SAM for Botswana requires you to calculate the column and row coefficients and then to use those coefficients and the SAM to track through the interpretation in Section 7 (Interpreting the Information in a Social Accounting Matrix) of the document 'Intro to SAMs cgemod.pdf'.

The macro-SAM for Botswana is reproduced in the workbook 'SAM exercises M1.xlsx' in the worksheet 'Botswana SAM'. You should spend a few minutes making sure you can interpret the meaning of the transactions recorded in each cell.

The worksheet 'Botswana Coeffs' is a template. In this workbook calculate

1. The row coefficients: these should appear in area B24:N36. **(NB: We use the standard way of defining the bounds of an array in Excel. This is defined by two cells separated by a colon (:); the first cell is the top left-hand corner and the second is the bottom right-hand corner.)**
2. The column coefficients: these should appear in the area B43:N55.
3. Include check calculations to ensure that your calculations are correct.
4. Explain why the row sums for the column coefficients do not equal one and the column sums for the row coefficients do not equal one.

Work through the description in Section 7 (Interpreting the Information in a Social Accounting Matrix) of the document 'Intro to SAMs cgemod.pdf', making sure you can identify the row or column coefficients that are used in each stage of the description.

South Africa SAM Interpretation Exercise

The interpretation exercise using the SAM for the Republic of South Africa (RSA) requires you to calculate the column and row coefficients and then to use those coefficients to answer some questions.

The SAM for South Africa is reproduced in the workbook 'SAM exercises M1.xlsx' in the workbook 'RSA SAM'. You should spend a few minutes making sure you

can interpret the meaning of the transactions recorded in each cell, descriptions of the accounts are recorded in column A, and make brief notes on:

1. The tax instruments and the transactions to which they relate.
2. The different factor types.
3. The different household types.
4. The commodity and activity accounts.

Column and Row Coefficients

The worksheet 'RSA sam coefficients' is a template. In this workbook calculate

1. The column coefficients: these should appear in the area C71:BG128.
2. The row coefficients: these should appear in the area C137:BG193.
3. Include check calculations to ensure that your calculations are correct.

Interpretation of the South Africa SAM

Spend some time interpreting the information provided by the row and column coefficients.

Then provide answers to the following questions

1. What are the shares of output of 'cpetchem' supplied by different domestic activities and imports at basic prices? (Hint: you may need to adjust for taxes and margins.)
2. What is the total share of taxes in the purchaser prices of 'cagric' and 'cveh'?
3. What are the factor input cost shares for the activity 'aaagric'?
4. What are the internal and external balances? What is the share of Government income accounted for by government borrowings?
5. What are the shares of government incomes from different taxes?
6. Compare the shares of income from different factors for the households 'hafflow' and 'hwhhigh'.
 - a. What do the incomes from each factor and these shares imply about the endowments owned by each of the two households?
 - b. How important are transfers from the government to each of these households?
7. Compare the shares of consumption expenditures by the households 'hafflow' and 'hwhhigh'.
 - a. What do the consumption expenditures and these shares imply about the preferences of the two households?
 - b. Are these patterns of expenditure consistent with your expectations (given inter alia Engels Law).

8. Compare the shares of expenditures by the households 'hafflow' and 'hwhigh' on direct taxes and savings.
 - a. What do the incomes taxes paid and these shares imply about the implications of changes in income tax rates for two households?
 - b. What do the savings and these shares imply about the preferences of the two households?
 - c. Are these patterns of savings consistent with your expectations (given the implied ownership of factors).
9. Compare the transaction values and column and row coefficients for exports. What do these imply about:
 - a. The relative importance of different commodities as sources of foreign exchange?
 - b. The shares of domestic commodity production exported? (You will need to do a side calculation for this one.)
10. Compare the transaction values and column and row coefficients for imports. What do these imply about:
 - a. The relative importance of different commodities as sources of imports?
 - b. The shares of imports in the total supply of commodities, valued at basic prices, to South Africa? (You will need to do a side calculation for this one.)

5. Ex M1:3 SAM Income Multiplier Exercises

The income multiplier exercise using the South African SAM involves three stages. First, you need to decide which accounts will be endogenous and which will be exogenous; for this exercise, we will specify those accounts that are exogenous. Second, having made that decision you need to calculate the multiplier matrix. And third, you need to use the multiplier matrix to carry out two simple experiments. There are two template worksheets for these exercises 'RSA Inc Mult' and 'RSA Price Mult Expt' in the workbook 'SAM Exercises M1.xlsx'.

The exogenous accounts will be the government, including tax accounts, the capital accounts, including stock changes, the enterprise account and the rest of the world account.

Deriving the Multiplier Matrix

The following are instructions for the stages you need to follow using the worksheet 'RSA Inc Mult Expt'.

1. Populate the endogenous accounts in the matrix C7:AV52.
2. Calculate the totals for the row sums (recorded in AX7:AX52) and the column sums (recorded in vector C54:AV54) for the exogenous accounts.
3. Calculate the matrix $[I - A]$ using the matrix C64:AV109) to record the result. (NB: An appropriately dimensioned identity matrix has been created for you in the worksheet 'RSA SAM Identity Matrix'.)
4. Calculate the inverse of the $[I - A]$ matrix using the matrix C116:AV161 to record the result. This needs the Excel function MINVERSE.

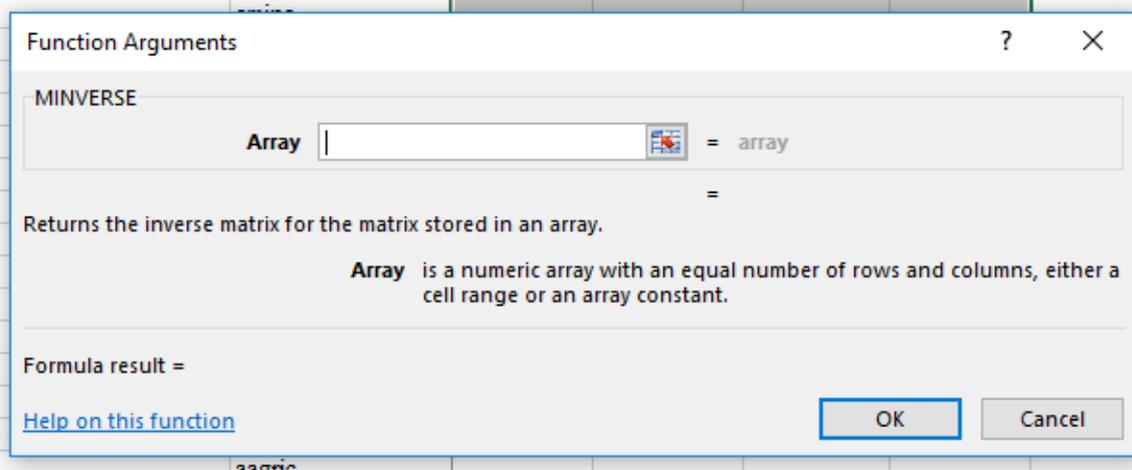
Excel's function library contains a wide range of matrix operations. Select Formulas>Math&Trig and select the function MINVERSE; you will need to scroll down. This will open the function arguments box shown below. This tells you what to do and there is a help system.

There are a few things worth knowing that may save you time:

1. You must select/highlight the into which the result will be recorded BEFORE selecting the function. This must be of the correct dimensions, i.e., to invert an $n \times n$

matrix you must select an $n \times n$ matrix for the result. Excel may not provide a meaningful error message if the dimensionality is wrong.

2. The array referred to in the Function Arguments box is the array/matrix that is to be inverted.
3. For ALL array operation hold down the **Ctrl** and **Shift** keys when pressing **OK**. (Experienced users may insert the formula in other ways; in those cases, hold down the **Ctrl** and **Shift** keys when pressing **Enter**.)



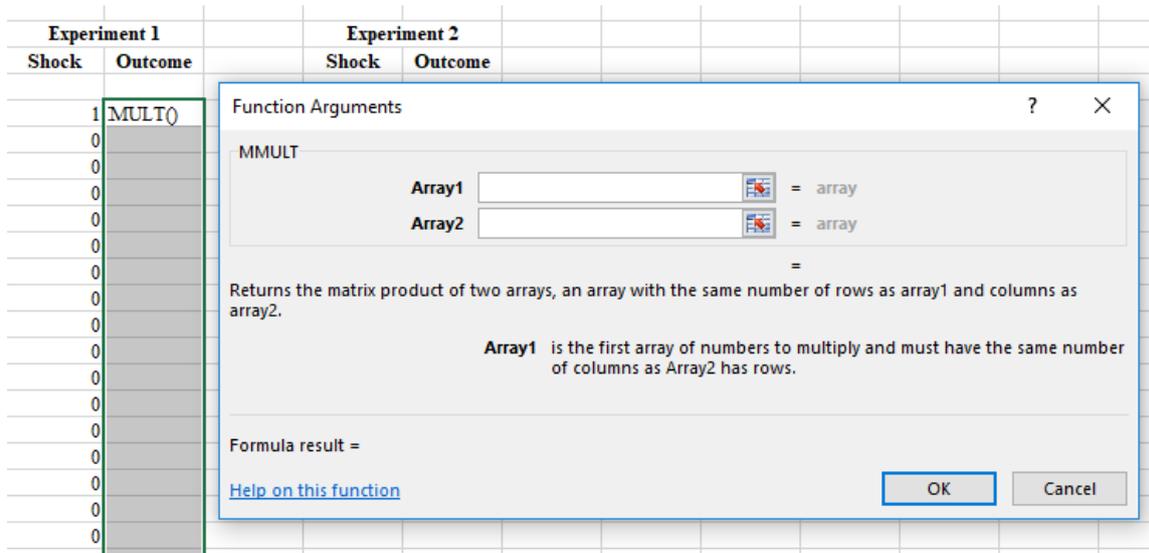
Simple Income Multiplier Experiments

A very simple approach to multiplier experiments is to evaluate the implications of a one-unit shock in exogenous demand for different accounts. In matrix algebra terms this effects a summation that picks out the column of the multiplier matrix that is paired with the respective row of the shock. Economically, such an experiment tells the analyst by how the output for every endogenous account must increase to satisfy the increased exogenous demand, e.g., export demand.

We will consider and compare two such unit shocks. The first, is an increase in the exogenous demand of the commodity agriculture (*'cagric'*) and the second for the commodity minerals (*'cmins'*). The template worksheet for this exercise is *'RSA Inc Mult Expt'*. The steps to follow are:

1. In the worksheet *'RSA Inc Mult Expt'*. Enter 1 (one) in the cell C8, and enter 0 (zero) in the cells C9:C53 for the experiment for *'cagric'*. For the experiment for *'cmins'*, enter 1 (one) in the cell F9 and 0 (zeros) in the other cells of column F.

2. In the cells D8:D53 calculate the value of $[I - A] * \Delta d$ where Δd is the shock vector (C8:C53). This needs the Excel function MMULT. Select Formulas>Math&Trig and select the function MMULT (see below for an illustration; note there is a help function).
3. Note that the operation MMULT requires two arrays. These must be entered in the correct order and have correct dimensions. The same operation of holding down the Ctrl and Shift keys when pressing OK is required.



Having computed the outcomes for both experiments you can compare the results.

1. Why is the total increase in outputs so much greater for 'cagric' than for 'cmins'. (HINT: remember the Keynesian injections = leakages in the new equilibrium.)
2. What are the shares of exogenous transactions in the columns of the SAM for 'cagric' than for 'cmins'?

6. Ex M1.4 SAM Price Multiplier Exercises

The price multiplier exercise using the South African SAM involves three stages. First, you need to decide which accounts will be endogenous and which will be exogenous; for this exercise, we will assume these accounts are the same as for the income multipliers (this will reduce the number of calculations needed). Second, having made that decision you need to calculate the multiplier matrix; this will only need us to transpose the income multiplier matrix. And third, you need to use the multiplier matrix to carry out two simple experiments. There are two template worksheets for these exercises 'RSA Inc Mult' and 'RSA Price Mult Expt' in the workbook 'SAM Exercises M1.xlsx'.

Deriving the Multiplier Matrix

This only requires that we transpose the multiplier matrix derived for the income multiplier case. Check out the mathematics in the SAM Multiplier document and PPTs. We will use the worksheet 'RSA Price Mult' in the workbook 'SAM Exercises M1.xlsx'.

Select the matrix C66:AV111.

1. Use the array function TRANSPOSE. Select Formulas>Lookup & Reference and select the function TRANSPOSE. And then follow the instructions noting that this is an array function.

This ensures that we have a correctly dimension price multiplier matrix. The alternative would be to transpose the shock vector, but this would produce the outcomes as a row in Excel, which is less easily read.

Simple Price Multiplier Experiments

A very simple approach to multiplier experiments is to evaluate the implications of a one-unit shock in exogenous **costs** for different accounts. In matrix algebra terms this effects a summation that picks out the column of the transposed multiplier matrix that is paired with the respective row of the shock. Economically, such an experiment tells the analyst by how the costs for every endogenous account must increase to satisfy the increased exogenous costs, e.g., an increase in commodity taxes.

We will consider and compare two such unit shocks. The first, is an increase in the exogenous cost of the commodity agriculture (*'cagric'*) and the second for the commodity minerals (*'cmins'*). The template worksheet for this exercise is 'RSA Price Mult Expt'. The steps to follow are:

1. In the worksheet 'RSA Price Mult Expt.' Enter 1 (one) in the cell C8, and enter 0 (zero) in the cells C9:C53 for the experiment for *'cagric'*. For the experiment for *'cmins'*, enter 1 (one) in the cell F9 and 0 (zeros) in the other cells of column F.
2. In the cells D8:D53 calculate the value of $[\mathbf{I} - \mathbf{A}]' * \Delta \mathbf{v}$ where $\Delta \mathbf{v}$ is the shock vector (C8:C53), using Excel's matrix multiplication function.

Having computed the outcomes for both experiments you can compare the results.

1. What are the differences in the implications for the costs faced by the different accounts from the differences in the two shocks?
2. How much have the costs of living gone up for different households because of the two shocks?
3. Explain why the changes in the cost of living for the households *'hafflow'* and *'hwhhigh'* differ. Are these differences consistent with expectations?

7. Ex M1.5 Collecting National Accounts Data

Data collection is time consuming and, usually, frustrating. Typically, national account agencies report data according to long established formats that often do not coincide with your interests/needs. Hence a time-consuming activity is finding where the data are reported, and another involves learning how the data are reported.

Data collection is time-consuming and requires careful organisation; simply accessing a data repository, e.g., the web site for a national account agency, is not enough. You need a plan and a framework within which to work. A NAM/SAM provides a framework within which to work. The plan needs to include the keeping of detailed records of the databases accessed, e.g., the url, date accessed, etc., the sources of individual transactions included in the SAM, e.g., a record of the codes used in the databases, and recognition that you will probably need to revise the databases you are compiling when you identify your mistakes.

You should start this exercise before moving on to Module M2; you will need some of the data to progress through Module M2.

Data Requirements

The minimum data requirements are:

1. Consolidated national accounts for a 'normal' year: ideally these will have been compiled in accordance with the System of National Accounts (SNA) and be fully reconciled. If they are not consistent with the SNA the development of the NAM/SAM may be more time-consuming; reconciliation will not be an issue.
2. Supply and Use tables (SUT): ideally for the same year as the consolidated national accounts or a recent 'normal' year.

These data should be available with varying degrees of accuracy/completeness for all OECD countries, most middle-income countries, and some lower income countries: consolidated national accounts and regular/annual SUT are standard parts of SNA compliance. If the consolidated national accounts are presented in annual form, it is easier to assess the data before compiling a NAM/SAM. If present for multiple years for different accounts, as in the UK's Blue Book, any gaps in the data are only likely to be found when

compiling the NAM/SAM. However, the full extent to which on-going revisions to national accounts data are fully reflected in all the published data is an open question.

For EU and OECD member states SUT are often available even if they cannot be accessed from the national accounts' agency. **If NO SUT data are available for your chosen country, you will need to collect data for another country for this course.** Additional guidance about estimating SUT tables from limited data will be provided later in the course.

NB: Countries, e.g., UK, are increasingly suppressing data in the published SUPPLY matrix by not reporting the value of secondary products produced by activities. Some comments about ways around this problem are explored later in the course.

Additional data that will be needed for further developing a NAM/SAM include

1. Household income and expenditure surveys:
 - a. ideally access to the raw survey data, but if not then access to detailed reports,
 - b. questionnaire and coding framework
2. Labour Force surveys:
 - a. ideally access to the raw survey data, but if not then access to detailed reports.
 - b. questionnaire and coding framework
3. Tax data:
 - a. estimates of applied tax rates and subsidies for trade, VAT, Excise, production, and any other taxes that are deemed important.
 - b. aggregate tax revenues by broad categories, e.g., taxes on products, subsidies on products, total tax revenues, is essential to provide a residual

In an ideal world you will be able to access the 'raw' data for the Household Income and Expenditure and Labour Force surveys. If you can access these data, you may need specialist software to access the data and software to process the data. The course assumes that you will need to use these data in Module M6, so collecting the data now is intended to give you time to access required software.

It may be wise to collect these data at the same time as collecting the minimum data requirements: it will save time and allow you to review these data during the course. The surveys should be reasonably accessible. The big problem is usually the tax data; often total

revenues by different tax instruments are available but detailed tax revenue by commodity, activity, factor, and household are not available; they may take time to find.

Data Handling and Storage

Everyone has their own preferences for data handling and storage. Whatever those preferences, the following guidelines are reasonable

1. Keep unchanged copies of the base data – you will make mistakes so do not work directly on the base data.
2. Keep detailed records of the data sources, e.g., the url and date accessed, any codes used by the source agency, etc., - there will be mistakes and you then need to know where to look.
3. Working in a programming language rather than Excel will make it easier to revise the derived data given additional or corrected data, and correct calculation errors.

Large amounts of national accounts data are now made available in Excel (xlsx, csv), so the first choice for aggregate national accounts and SUT is Excel. Historically, I have used Excel to store collected data that are then transformed into GDX and then used GAMS to assemble the data. This is despite these tasks not being a strength of GAMS, because the output from the assembly process is the input to `SAMEST.gms`. Recent developments at GAMS indicate that embedded Python may be a better future choice.

Survey data comes in various formats. In the past it was often best to use the software used by the source agency, e.g., SPSS. But nowadays most of the software packages can read data in multiple formats so the choice is more flexible and usually reflects the user's experience. A package with statistical capabilities, e.g., SPSS, Stata, R, is essential for the survey data to produce moment data when deriving Representative Household Groups and labour categories.