

## Estimating a Social Accounting Matrix: Online Course

*“[I]t is **impossible** to establish by direct estimation a system of national accounts free of statistical discrepancies, residual error, unidentified items, balancing errors and the like since the information available is in some degree incomplete, inconsistent and unreliable. Accordingly, the task of measurement is not finished when the initial estimates have been made and remains incomplete until final estimates had been obtained which satisfy the constraints that hold between their true values.” (Stone, 1982, p 186, emphasis added).*

|                                                                 |    |
|-----------------------------------------------------------------|----|
| Introduction.....                                               | 2  |
| Course Description.....                                         | 3  |
| Software .....                                                  | 6  |
| Course Aims and Objectives.....                                 | 6  |
| Estimating a Social Accounting Matrix Course: Outline .....     | 7  |
| Module M1: Theory of Social Accounting.....                     | 7  |
| Module M2: Developing a Prior SAMs.....                         | 8  |
| Module M3: Matrix Estimation Methods: RAS .....                 | 8  |
| Module M4: Matrix Estimation Methods: Entropy Theory.....       | 9  |
| Module M5: Matrix Estimation Methods: Entropy Practical 1 ..... | 9  |
| Module M6: Matrix Estimation Methods: Entropy Practical 2 ..... | 10 |

Disaggregated national accounts are core components of the System of National Accounts (SNA)

## Introduction

This course is a specialist course in the development and estimation of Social Accounting Matrices (SAMs). The course is orientated to the estimation of SAMs designed to support computable General Equilibrium (CGE) models, but SAMs so estimated can be used by other whole economy models. The course is designed for individuals who have a well-developed background in economics and single-country or global CGE modelling, and who have extensive experience using the General Algebraic Modelling System (GAMS) software. The GAMS programmes used in this course are more complex than many of those used for CGE models. The course uses GAMS intensively and users will need a full GAMS license. During the course participants will typically need to expand their knowledge and understanding of GAMS.

The transactions data used in all CGE models must conform the principles underpinning a SAM; if they do not then either the data or the (CGE) model or both are wrong.

The course emphasises the development of an understanding of the properties and structure of SAMs, with a strong emphasis on understanding the price system embedded in a SAM and its role in whole economy models. The course follows the principles of national accounting developed by Richard Stone, and associates, during the 1950s and 1960s, and that are now cornerstones of the System of National Accounts.<sup>1</sup> The estimation of SAMs that are appropriate to support meaningful policy analyses depend critically upon the effort devoted to compiling a high-quality *PRIOR* (initial) SAM from national accounts data, including various surveys. The mathematical techniques, explored in this course, for converting the prior SAM into a final SAM are **NOT** a substitute for the efforts devoted to compiling the prior SAM.

During the course users will use different mathematical techniques to estimate SAM using data provided for the course. The worked examples for the course use data for Bhutan developed by Arndt Furbacher.

---

<sup>1</sup> “The fact that others have not had to reinvent the architecture of the national accounts in particular is perhaps the most telling measure of the importance of Richard Stone's contributions and their enduring significance.” (Pyatt, 2005).

## Course Description

All published national accounts are estimates derived from limited data and imperfect information, this is especially the case with disaggregated national accounts, e.g., Input-Output Tables (IOT) and Social Accounting Matrices (SAM). This is inevitable because estimated national accounts depend heavily on information collected from surveys, e.g., production, household and labour force surveys, where point estimates are measured with error. Moreover, for some data points prior estimates are missing and/or depend on ‘inferences’, i.e., data are limited. These are not novel observations. Even from the beginning of the development of modern national accounts the issue of limited data and imperfect information was recognised.<sup>1</sup>

A review of the literature on the estimation of national accounts, especially disaggregated national accounts, demonstrates that the fundamental problems of estimation with limited data and imperfect information have been recognised from the beginning. However, it is arguable that practitioners (economists?) outside of national accounts agencies have all too often been unaware of the historic attempts to address the problem of estimation with limited and imperfect information. Some of the historic attempts are review in this course to emphasise the importance of addressing information gaps and how an entropy metric addresses information gaps.

The course starts with a module on Social Accounting Matrices (SAM). This is not an introduction to SAMs as single-entry accounting systems. The module explores the

- Importance of the complete and consistent requirement for a SA,
- the (implicit) price system that is embedded in any SAM,
- the importance of the System of National Accounts (SNA) production boundary,
- the role of Supply and Use (SUT) (interindustry) accounts in the SNA,
- the relationship between SUT and Input Output Tables (IOT),

---

<sup>1</sup> Stone, R., Champernowne, D.G. and Meade, J.E., (1942). ‘The Precision of National Income Estimates’, *The Review of Economic Studies*, Vol 9(2), pp 111–125.

- the interpretation of column coefficients in a SAM, and
- the contribution of satellite accounts.

The second module explores how prior estimates of complete but inconsistent SAMs can be derived from aggregate national accounts data and information of surveys/censuses. The emphasis is on information gathering and assembly. Early practitioners place great importance on confronting information from different sources to derive initial point estimates of transactions. The importance of reliable prior estimates was important because the implementable mathematical techniques available in the 1960s and early 1970 were constrained by computing power. The third module reviews these techniques and their developments for two reasons: first, they transparently provide insights into the difficulties of SAM/matrix estimation, and second, despite the advances in computing power RAS techniques are still commonly used<sup>1</sup>.

But RAS is deficient as a tool for SAM estimation; this is demonstrated by replicating a study by Lynch to estimate an Input-Output Table (IOT) for 1968 using the 1963 IOT as a prior. A modified RAS method is used to demonstrate how adding information can greatly improve the performance of the RAS method. The modified RAS method provides a basis upon which users can appreciate how adding even small amount of additional information can greatly improve the performance of any mathematical SAM estimation technique

Modules M4, M5 and M6 concentrate on using the entropy metric to estimate matrices, of which SAMs are a subset. Module 4 concentrates on information theory from which entropy estimation methods and objective function are derived. The relationship with the Stone-Byron method is explored and how entropy estimation methods are a generalisation of the insights developed by Stone et al., (1942) (see reference in footnote on page 4).

Modules M5 and M6 are practical. In module M5 a simple SAM is estimated using the method of sequential estimation whereby as the prior SAM is extended using the maximum entropy method, without ‘tuning’, to develop increasingly disaggregated SAM. Module 6

---

<sup>1</sup> The available evidence indicates that the FIT programme used to reconcile the GTAP database is a RAS routine (see James, M., and McDougall, R., (1993). ‘FIT: An Input-Output Data Update Facility for SALTER’, *SALTER Working Paper No. 17*. Australian Industry Commission.

concentrates on estimating a large(ish) SAM while progressively tuning the maximum entry programme (SAMEST). This is essentially an iterative process, where the programme is used to learn about the properties of the estimated SAM so as to guide tuning the programme.

The main SAMEST programme used in this course, to derive/estimate the final SAM, uses a maximum entropy metric (developed by Sherman Robinson and associates) that is based on information theory. During the course the RAS method, developed by Stone and associates, is also used: it is demonstrated that the RAS method is good, given the computing constraints when it was developed, for its intended purposes. The entropy programme used in this course is designed to allow those estimating a SAM to add additional layers of information over and above the *prior* SAM. The entropy method has another major advantage: the estimation results provide information about the potential benefits of improved prior estimates of transaction values.

This course requires user to have substantial experience with using GAMS; the techniques used are more complex than those used in GAMS based CGE models. The maximum entropy metric is Bayesian, so users need a good knowledge of statistical theories.

The course assumes that users have an in-depth knowledge of graduate level economics and a reasonable understanding of mathematics and statistics for economics. The methods used in this course require an understanding of Social Accounting Matrices (SAMs), and an understanding of the relationships between SAMs and CGE models will be helpful.

The materials are organised in 6 modules. The first three modules contain 22 components that will require some 3 to 5 hours, on average. The second three modules contain 16 components. Those in module M4 will require, on average, 2 to 3 hours each, while the components of modules M5 and M6 will require, on average, 4 or more hours each.

User should plan a SAM estimation project of their own that will be undertaken immediately following completion of this course: experience indicates that embedding the skills explored in this course is important.

The course is delivered from a website ([www.cgemod.org.uk/training.html](http://www.cgemod.org.uk/training.html) ).

## Software

This course does NOT use a GUI (Graphical User Interface) to access GAMS. The GAMS programmes used in this course require that the participants work in ‘native’ GAMS using data that are accessed from Excel and GDX. Subsequently, when estimating a SAM using national surveys users may need to master other computer programmes, e.g., SPSS, SQL, Stata, R, etc., to extract the data; the requisite programmes will depend on the format in which the data have been stored. These programmes are not used in this course.

A full GAMS licence (see [https://www.gams.com/buy\\_gams/](https://www.gams.com/buy_gams/)) with the CONOPT solver is recommended for this course. While CONOPT is adequate for this course user should also consider having access to the KNITRO solver. (PATHNLP may be useful and since PATH is used for the CGE courses there is no marginal costs).

For a license details contact GAMS Corporation ([sales@gams.com](mailto:sales@gams.com))

## Course Aims and Objectives

### Course Aims

To develop the SAM estimation skills of participants (using GAMS) so they

- i) understand the structure and content of complete and consistent SAMs,
- ii) understand the price system in a SAM,
- iii) can organise the data required for *prior* SAMs and satellite accounts,
- iv) understand the strengths and weaknesses of SAM estimation techniques, and
- v) can critically evaluate the information content of a SAM.

### Course Objectives

On completion of the course the participants will be able to:

- i) evaluate critically the information content of a SAM,
- ii) identify and understand the data requirements to create a *prior* SAM,
- iii) use SAMEST, & RAS, programmes to estimate complete & consistent SAMs,
- iv) interpret the results from the SAMEST programme to identify those cells of a *prior* SAM that would benefit from improved prior estimates, and
- v) identify appropriate satellite accounts consistent with the estimated SAM.

## Estimating a Social Accounting Matrix Course: Outline

### Module M1: Theory of Social Accounting

|      | Topic                          | Tasks                                                                  | Exercises                                                               |
|------|--------------------------------|------------------------------------------------------------------------|-------------------------------------------------------------------------|
| M1:1 | Introduction                   | Introduction.                                                          | None                                                                    |
| M1:2 | Introduction to SAMs           | SAMs a single-entry accounting system. Top down & Bottom-up Approaches | Data Requirements; Interpreting transaction values (TV)                 |
| M1:3 | Macro SAM                      | Macro SAMs                                                             | Deriving a simple Macro SAM from macro-T-accounts                       |
| M1:4 | Prices in a SAM                | The price system in a SAM and the Law of One Price (LOOP)              | Deriving implicit prices from TVs<br>Building your own macro-SAM        |
| M1:5 | SNA Production boundary        | The SNA and 'general' production boundaries                            | Deriving a macro-SAM from aggregate national accounts data              |
| M1:6 | SAMs and the SNA               | Interindustry Tables, SAMs and the System of National Accounts         | The relationship between Supply and Use Tables and Input-Output Tables. |
| M1:7 | Interpreting a SAM             | Using coefficients to interpret the information in a SAM               | A SAM interpretation exercise                                           |
| M1:8 | Satellite Accounts and the SNA | Satellite Accounts and SAMs in the SNA                                 | Interpreting the data in satellite accounts                             |

## Module M2: Developing a Prior SAMs

|      | Topic                                | Tasks                                        | Exercises                           |
|------|--------------------------------------|----------------------------------------------|-------------------------------------|
| M2:1 | Role of a macro-SAM                  | Macro SAMs and top-down method               | Deriving implicit prices from TVs   |
| M2:2 | Macro SAM                            | Data requirements                            | Collecting data for a macro-SAM     |
| M2:3 | A prior Macro SAM                    | Organising data for a <i>prior</i> Macro-SAM | Building your prior macro-SAM       |
| M2:4 | Micro SAM                            | Data requirements and organisation           | Collecting data for a micro-SAM     |
| M2:5 | A first prior Micro SAM              | Combining macro-SAM and SUT data             | Building your first prior micro-SAM |
| M2:6 | Sequential estimation of a Micro-SAM | Micro-SAMs and control totals                | Deriving implicit prices from TVs   |
| M2:7 | Prior SAM                            | Using data to derive a prior SAM             | Ex M2.1 Building a prior SAM        |

## Module M3: Matrix Estimation Methods: RAS

|      | Topic                    | Tasks                                                        | Exercises                                         |
|------|--------------------------|--------------------------------------------------------------|---------------------------------------------------|
| M3:1 | Matrix Estimation        | Review of matrix estimation methods                          |                                                   |
| M3:2 | Estimation methods       | Estimation tools; estimation vv 'balancing'; control totals. |                                                   |
| M3:3 | RAS                      | The simple mechanics of RAS                                  | Ex M3.1 RAS in Excel<br>Ex M3.2 RAS in GAMS.      |
| M3:4 | RAS Reconciled Macro SAM | Using RAS to reconcile a Macro SAM                           | Ex M3.3 A Reconciled Macro SAM                    |
| M3:5 | Modified RAS             | RAS estimation with additional information                   | Ex M3.4 RAS with Extra and Incomplete Information |
| M3:6 | RAS Reconciled Micro SAM | Using RAS and Modified RAS to reconcile a micro-SAM          | Ex M3.5 A Reconciled Micro SAM                    |
| M3:7 | Evaluating RAS           | Comparing RAS estimated SAM with a known SAM                 | Ex M3.6 Comparison of SAMs                        |

## Module M4: Matrix Estimation Methods: Entropy Theory

|      | Topic                   | Tasks                                          | Exercises                          |
|------|-------------------------|------------------------------------------------|------------------------------------|
| M4:1 | Information Theory      | Economics & Information                        | None                               |
| M4:2 | Estimation with error   | Stone-Byron and Entropy estimation methods     | None                               |
| M4:3 | Info-metrics            | Estimation with imperfect information          | None                               |
| M4:4 | Maximum & Cross Entropy | Specification of objective functions           | None                               |
| M4:5 | Error Specification     | Error specification with imperfect information | None                               |
| M4:6 | SAM Estimation          | Samest_5                                       | Ex M4.1 Entrop Estimation mini-SAM |

## Module M5: Matrix Estimation Methods: Entropy Practical 1

|      | Topic                         | Tasks                                                 | Exercises                                     |
|------|-------------------------------|-------------------------------------------------------|-----------------------------------------------|
| M5:1 | SAM Estimation                | A First SAM Estimation                                | Ex M5.1 Estimating a first micro (Bhutan) SAM |
| M5:2 | Adding Institutional Accounts | Augmenting a first Micro-SAM with multiple households | Ex M5.2 A micro-SAM with multiple households  |
| M5:3 | Adding Factor Accounts        | Augmenting a Micro-SAM with multiple factor account   | Ex M5.3 A micro-SAM with multiple factors     |
| M5:4 | Adding Factor Accounts        | Augmenting a Micro-SAM with multiple factor account   | Ex M5.4 A micro-SAM with multiple factors     |
| M5:5 | Adding information            | Going beyond row and column controls                  | Ex M5.5 Adding control totals                 |

## Module M6: Matrix Estimation Methods: Entropy Practical 2

|      | Topic                       | Tasks                                                     | Exercises                                               |
|------|-----------------------------|-----------------------------------------------------------|---------------------------------------------------------|
| M6:1 | SAM Estimation              | SAM Estimation with SAMEST_5                              | Ex M6.1 Estimating a full SAM for Bhutan                |
| M6:2 | Error specification         | Setting error specification                               | Ex M6.2 Adjusting error specification                   |
| M6:3 | Adding error information    | 'Tuning' entropy estimation; refining error specification | Ex M6.3 Adding information about error terms            |
| M6:4 | Evaluating an Estimated SAM | Statistical measures to evaluate SAM/matrix estimates     | Ex M6.4 Evaluating estimated SAMs/matrices              |
| M6:5 | Matrix Estimation           | Matrix Estimation with MATEST_5                           | Ex M6.5 Estimating bilateral transfer matrices for GTAP |