Constructing SAMs for Development Policy Analysis: Lessons Learned and Challenges Ahead

Jeffery I Round[1]

February 2003

Abstract:

The aim of this paper is to appraise a few of the key innovative features of the early work in compiling SAMs for development policy analysis; to set out and review some recent methodological advances; and to identify those areas where compilation continues to be problematic. It briefly re-visits the features of the SAM as an integrating framework and sets out its relationship to the SNA 1993. The main compilation problems faced in practice arise from assembling the household accounts from household survey data where income data are especially unreliable and are difficult to link to the factor accounts and to income transfers. Experience is drawn from the construction of a Ghana SAM. Relatively more attention has been devoted to balancing and data reconciliation methods, which are briefly reviewed, although these are second order adjustments and much still depends on the quality of the initial estimates.

KEYWORDS: Social Accounting Matrices, SAM, household accounts, balancing methods

---

[1] Department of Economics, University of Warwick, Coventry CV4 7AL, U.K. E-mail: J.I.Round@warwick.ac.uk. An earlier version of this paper was presented at a seminar 'Data Analysis for Poverty and Income Distribution Analysis' in honour of Professor Graham Pyatt, Institute of Social Studies, The Hague, Amsterdam, November 29-30, 2001. I wish to thank André Lemelin, Rob Vos and the seminar participants for very helpful comments. The usual disclaimers apply.

1. Introduction

In April 1973 a small team of economic statisticians led by Graham Pyatt embarked on a mission to compile a social accounting matrix (SAM) for Sri Lanka. While this was not the first ever SAM to be compiled, either for a developed or developing country[2], nor was it even the first experience with SAMs for many of the team members[3], it did prove to be a landmark event. A great deal of subsequent work and literature on SAMs followed, including the publication of the conceptual framework by Pyatt and Thorbecke (1976). It is fair to say that the SAM concept has had a significant impact on data analysis and modelling and on development policy analysis more generally. It is a fact that even before the Sri Lanka mission, a new momentum had already been mounting to re-focus the data and information base in support of policy analysis away from almost exclusively 'production-oriented' aggregate measures and towards 'people-oriented' data and information strategy[4]. However, and in common with other early SAM studies for Iran, Colombia and Swaziland in quite different settings, the Sri Lanka exercise served to show what could be achieved with relatively limited source data, intensive effort by a dedicated team, and considerable local expertise on the economy in question.

The aim of the present paper is to appraise a few of the key innovative features of the early work on SAMs (including Sri Lanka) three decades on; to set out and review some recent methodological advances in compilation; and to identify those areas where compilation continues to be problematic. The paper begins with a brief discussion of the SAM as an integrating framework, especially in the light of the emergence of the 1993 SNA (SNA, 1993). The particular question raised here is: to what extent would an implementation of the 1993 SNA fulfil all the needs for assembling a SAM? This is the subject of the Section 2. Section 3 then sets out some compilation issues mainly based on recent experience in constructing a SAM for Ghana. These are potentially important issues and

[2] Most references are to the U.K. SAM produced by Sir Richard Stone and associates (Cambridge, 1962) as the most comprehensive example of an earlier SAM.
[3] Graham Pyatt and Alan Brown were associates of Sir Richard Stone in the Cambridge Growth Project, and Alan Roe and I had worked together on a regional version of the U.K. SAM at the University of Wales, Aberystwyth in the mid 1960s.
[4] Pyatt and Thorbecke (1976; p 1) refer to their respective studies for Iran and Colombia which they carried out under the aegis of the International Labour Office (ILO) World Employment Programme led by Dudley Seers, and to Seers' more general influence on their work.

mainly concern the use of household surveys, which do not seem to have been adequately addressed in the recent literature. Section 4 reviews some technical issues to do with re-balancing inconsistent data in a SAM framework. In the early SAMs for Iran, Sri Lanka, Swaziland and Colombia, expert judgement was almost always the way in which inconsistencies were eliminated. Now algorithms are much more commonly used and the aim here is to look at the current range of techniques and to make some comparative observations about them. The final section concludes the paper.

2. SAMs: an integrating framework?

It is now well-known that a SAM, a concept due to Sir Richard Stone, is a matrix representation of transactions in a socioeconomic system. It is a comprehensive, flexible, and disaggregated framework which elaborates and articulates the generation of income by activities of production and the distribution and redistribution of income between social and institutional groups. A principal objective of compiling a SAM is, therefore, to reflect various interdependencies in the socioeconomic system as a whole by recording, as comprehensively as is practicable, the actual and imputed transactions and transfers between various agents in the system. The key distinguishing features of the SAM relative to alternative accounting systems are, first, the system is represented by a set of single-entry accounts; secondly, it places relatively more importance on the factoral, household and institutional dimensions; and thirdly, the framework is complete and comprehensive. Stone's earliest writings on SAMs highlighted these features, though he focused more on the first and the third[5] - while the social dimension was certainly present in his earliest work (the term 'social' accounting matrix is significant here) actual examples of disaggregations of the factoral and household accounts were few and far between.[6] It was therefore highly significant when Pyatt and Thorbecke stressed the need to make explicit in the accounts 'what is going on in any economy and how the living standards of different groups are related both to each other and to other aspects of economic activity' (Pyatt and Thorbecke, 1976; p 5)[7]. The framework then began to be piloted in a series of empirical exercises in selected countries (Pyatt and Round, 1977).

[5] Stone employed the matrix accounting format in many of his early writings on the national accounts.
[6] See especially Cambridge (1962) for examples based on the U.K. economy.
[7] There is a related issue about some early literature which integrates income distribution and production structure in multiplier models (e.g. Miyazawa, 1976). This has been discussed by Pyatt (2001).

Basic motivations for constructing SAMs

Beyond the obvious elaboration and detailed representation of the circular flow of income, a suitably-designed SAM should provide information on how, and the degree to which, different groups in society interconnect and interact with the rest of the economic system. This primary aim underlies three main benefits that arise out of compiling a SAM. First, their *construction* requires a significant degree of detailed estimation and use of data sets that have not hitherto formed part of standard national accounting practice. These can be used to good effect in improving estimates more generally. Secondly, they are a very good way of *displaying* information; the structural interdependence at both the macro and meso levels are shown in a SAM in a simple and illuminating way. Thirdly, they represent a useful analytical framework for *modelling*; that is, they provide a direct input into a range of fixed-price multiplier models and are an integral part of the benchmark data set that supports computable general equilibrium (CGE) models. This paper focuses on the first of these motivations, although clearly it is impossible to isolate one from the other entirely as they are mutually interdependent and reinforcing.

Some preliminaries: a basic SAM

For completeness and to assist in our discussion of compilational aspects, Table 1 shows a very basic SAM. It contains many simplifications of the structure found in a full SAM (Pyatt, 1991a and 1991b). In Table 1 the ordering of the accounts reflect the emphasis on factor income generation and domestic institutions and, through further disaggregations of the current accounts of institutions it is designed to show the structure of income flows and transfers between institutions. There is no reason, in principle, why we should not also show similar amount of detail for the institutions' capital accounts but this is not shown here as it is beyond the scope of our present discussion. In spite of the simplifications Table 1 is a reasonably complete representation of all the major transactions within a socioeconomic system. If these transactions are estimated for an economy in a particular accounting period then, with suitable disaggregations of the major blocks of accounts, the resulting SAM provides useful information about that economy for a wide range of structural and policy analysis. In particular, it connects the following aspects: the levels and distributions of incomes available to institutions (in particular households); the private and public spending of these incomes on goods and services (which are part of the determination of individuals' living standards); transfer payments and savings by institutions; the production of goods and services, and the generation of factor incomes.

[TABLE 1 HERE]

This simplified structure of Table 1 conceals many complexities that make the compilation of SAMs a difficult exercise. For example, it is well-known that, until recently, very few developing countries compiled their national accounts on the basis of the income or expenditure methods. Most were compiled from the production side (Heston, 1994) and even then, the use of commodity balances (i.e. input-output tables) was rare. So, to integrate source data on the incomes and outlays of households, corporate enterprises and government within a unified consistency framework was - and still is - a non-trivial step. Compilers of the initial SAMs were confronted with a range of problems, some were conceptual (e.g. dealing with many estimation and boundary problems) while others were practical, for example, dealing with different survey practices, definitions, timing, coverage, etc, some of which only come to light when the estimates from different sources were compared side by side.

Factor and household classifications

Disaggregations of the factor and household accounts are fundamental to any SAM. Pyatt and Thorbecke (1976) set out some clear principles and guidelines for choosing these factor and household classifications. Their main recommendation, which has been taken up in many studies since, was that classifications should be chosen to introduce as much within-group homogeneity relative to between-group differences as is possible, bearing in mind the limitations on the number of classifications that can be supported by the data.

In deciding on an appropriate disaggregation of the factor accounts, the aim should be to choose classifications which identify distinct factor markets. Accounts for labour are often cross-classified by location (e.g. urban-rural, or geographical region), skill or education level attained, employment status (e.g. employee, own account worker, employer) and by gender. Mixed income (a category suggested in the 1993 SNA) is also frequently chosen as a category to represent the income of household enterprises (where it is difficult to distinguish the returns to labour from the returns to other factors) and is also cross-classified in a similar way to labour. There are fewer distinctions between different types of capital and natural resources although for modelling purposes there may be multiple accounts if these factors are assumed to be sector-specific and hence not fully mobile.

Household classifications are chosen in accordance with the overall analytical or policy focus and to a degree that can be supported by the data (Pyatt and Thorbecke, 1976, and SNA, 1993; chapter XX). Many different criteria have been selected including; geographical location (e.g. urban-rural), assets (e.g. wealth, size of land holding) and the socio-economic characteristics of a representative individual (e.g. household head or principal earner). In many recent SAMs urban households tend to have been disaggregated by socio-economic group while rural households have been disaggregated by some dimension of land holding. Income level (e.g. division by income deciles) has usually been avoided as a classification criterion, because households are potentially mobile between income groups making ex ante and ex post comparisons and policy-targeting difficult (Pyatt and Thorbecke, 1976). However, there are several examples of recent SAMs where, for the purposes of making cross-sectional comparisons especially, income percentile groups have been used.

It is clear that the case often made for 'flexibility' in guiding the choice of classifications in a SAM in order to fit the characteristics of the economy in question has more recently taken on a new meaning. In the early days, the choice of classifications was a defining moment in SAM construction. The decision was almost irrevocable; tables and matrices were produced as a consequence of that decision and users - including modellers - simply had to work with the result. One could aggregate accounts but re-classifications were very limited indeed. Nowadays, with the availability of computer software and better spreadsheet technology, it is perceived to be more important to maintain as much detail as possible so as to enable the user to aggregate the SAM to one or more alternative classifications. The old method has proved to be far too rigid. The most recent development in modelling is towards a micro-simulation approach and to model in detail the behaviour that is observed at the level of individual households and firms (Cockburn, 2001; and Robilliard, Bourguignon and Robinson, 2001)[8]. This underscores the need to avoid compromising subsequent use and analysis of the SAM by pre-selecting rigid sets of classifications, and this especially applies to factor and household classifications, as well as products and factors.

[8] Cockburn (2001) works with a microsimulation model based on a SAM for Nepal in which he treats all 3373 households in the sample survey as representative households.

Relationship with the 1993 System of National accounts

The 1993 SNA (SNA, 1993) embodies many new features compared with its predecessor. In the preface to the volume[9] four key features of the revised system are highlighted. In particular, it is claimed that the system is comprehensive, that it is flexible, that it acts as an improved guide to concepts and definitions, and that it reinforces the central role the national accounts play in relation to economic statistics more generally. To a large degree these are quite legitimate claims. The new system is quite comprehensive and there are certainly many significant novelties and improvements, not least with regard to the central focus that institutional sectors now have in the new structure[10]. The earlier work on SAMs must have been influential in this respect, by helping to shift the focus away from production accounts *per se* and towards institutions. However, SAMs have not been embraced as the core of the system and it is therefore reasonable to enquire to what extent, if a country implements the 1993 SNA, it simultaneously fulfils all the needs for constructing a SAM for analytical purposes. We shall see that there are some important lacunae and many remaining difficulties. Some of these have been discussed previously (Keuning, 1998; and Pyatt, 1999) while others merit some further discussion. So to proceed further let us briefly examine some key features of the 1993 SNA and see how it translates into a SAM.

The 1993 SNA is created around a central framework which consists of several components. For our purposes we may identify three main components[11]. These are the Supply and Use Table (SUT), the Integrated Economic Accounts (IEA), and various sets of three-way cross-classification tables of which the Cross-Classification of Industry and Sector (CCIS) is perhaps the most important. The SUT table is a fairly conventional set of matrix accounts which records the supply and use of products by activities, extended to show the generation of income by activities and the final use of products by institutional sector. The CCIS tables are potentially a fairly flexible concept. In an implementation of the 1993 SNA by the Ghana Statistical Service, the CCIS tables were conveniently incorporated into the SUT table by disaggregating activities by institutional sector[12].

[9] SNA (1993; p xxxiv)

[10] Many other previously-acknowledged advances of the new system are not referred to here: the discussion is purposely limited to be relevant to our comparisons with SAM studies.

[11] In terms of terminology it should be noted that 'sector' here refers to an institutional sector, while we shall continue to use the terms 'activity' and 'industry' interchangeably. 'Products' and 'commodities' are also synonymous terms.

[12] This is referred to in Powell and Round (1998; p 13). It has the advantage of making explicit the differences in

The IEA is a central element of system. In essence it shows sets of current, accumulation and asset accounts for each institutional sector, and for the total economy and for the rest of the world. Leaving to one side the accounts for assets, which are not a practical proposition for most developing economies, the current accounts are split further into a set of production accounts; the generation, distribution and use of income accounts; capital transactions accounts; and connecting accounts for the rest of the world. But perhaps most significant of all, the 1993 SNA is represented in a T-account format rather than the matrix-accounting format of a SAM.

A matrix representation of the 1993 SNA

A summary matrix representation of the aggregate system is included in the 1993 SNA and a link with SAMs is also discussed.[13] Table 2 shows an abridged version of the system. One can observe that Stone's fundamental (i.e. four-account) accounting structure is easily identifiable by the following blocks of accounts: production (1, 2, 3), consumption (4, 5, 6), accumulation (7), and the rest of the world (8). We may also observe that the balancing items for the accounts are recorded in a natural step-wise fashion as income 'cascades' from one account to the next in sequence. The circular flow of income, so clearly featured in the early SAMs is therefore also readily apparent in this matrix representation.

[TABLE 2 HERE]

Table 2 has many of the features of the basic SAM shown previously in Table 1. Thus, for instance, the accounts recording the generation of income (account 2) are equivalent to the factors of production accounts in Table 1. One difference is that Table 2 shows more of the *process* of income transmission between institutional sectors, through the primary allocation, secondary, and the use of income accounts. Many earlier SAMs (viz Table 1) compressed all of the various elements of income redistribution between sectors into a single submatrix, and property income and current transfers were all subsumed in one set of cell entries. So these are largely alternative presentational arrangements and Table 2 fully qualifies as a SAM.

---

activities' technology and product mixes across institutional sectors.

[13] SNA (1993; Tables 2.5 and 20.4).

What then are the major stumbling blocks in constructing a SAM from the SNA? The first and fairly obvious point to make is that there is no guarantee that the T-accounts will easily convert into a comparable set of matrix accounts. The transactions accounts that underpin the system record the origin and destination of resources and uses. So in many cases the originating and destination accounts (by sector) can easily be identified but this is not always the case. This problem featured in compiling the SAM for Ghana alongside an attempt to implement the 1993 SNA (Powell and Round, 1998). The main difficulty arose in the derivation of estimates of intersectoral property income, current transfers and capital transfers (i.e. matrices (4, 4), (5, 5) and (7, 7) in Table 1) which could not be obtained from the IEA accounts or any existing tables. They had to be compiled separately by constructing sets of transaction matrices. Obviously this is not a problem unique to the SNA; these intersectoral transfers would have to be estimated for the SAM anyway, the point is that the SNA is not, of itself, a sufficient source of information for compiling a SAM.

A second point, as noted by Keuning (1991 and 1998), is that guidance in the 1993 SNA on the 'generation of income' (i.e. factor) accounts is decidedly weak. For instance, there are no recommendations about including any disaggregations by types of labour, capital or land. In a similar vein, there is only a minimal discussion of possible disaggregations of the household accounts, although there are references to the importance of 'sub-sectoring' the household sector, including some possible classification criteria, and a discussion of the need to maintain flexibility. Whilst the specific choice of classifications has to be case-specific there is only a brief reference to the importance of disaggregations for pursuing policy analysis in developing countries (SNA, 1993; 4.152). So there is an overriding sense that the 1993 SNA will not generate sufficient impetus to collect enough information to construct even moderately useful SAMs - the basic tabulations for constructing detailed factor and institution accounts will simply not be there. A series of operational guidelines, including software systems, are currently being developed to assist in the implementation of the 1993 SNA and are being disseminated in the form of a series of Handbooks of National Accounting (UNSD, 1999 and 2000). Clearly these are helpful steps in enhancing current practice but they are not yet enough to help develop fully articulated SAMs[14].

[14] At the time of writing a Handbook on Social Accounting Matrices and Labour Accounts is being prepared.

Relationship between SAMs and other accounting systems

In spite of the twenty-five years' or more experience in constructing SAMs it is still common to find modellers and economic analysts referring to an input-output table as a 'SAM'. Those of us who have wrestled with the challenges of assembling a SAM know full well that an input-output table is just the beginning and not the end of the compilation process. Clearly, some representation of the functional and institutional distribution of income is a minimal requirement for a SAM. Beyond this, there would probably be general agreement with the view that to earn the title 'social' accounting matrix the matrix needs to display at least some further minimal disaggregation of the household and factor accounts in order to capture some higher order institutional features. These requirements are in addition to the formal properties any accounting matrix must have. Therefore the really significant steps in compiling a SAM are those that integrate information from household and labour force surveys alongside the input-output accounts, and which represent a disaggregated mapping of the circular flow of income via accounts for factors and households.

Extended input-output tables continue to be a puzzling concept. Pyatt (2001) has demonstrated how the so-called Miyazawa multipliers can be derived from a particular reduced-form SAM. However there are compilational implications of the Miyazawa system too. It is difficult to see how a reduced-form SAM can be compiled in practice without first compiling a complete SAM. For example, while it is possible to eliminate the factor accounts either by apportionment or by aggregation to show factor income mapping directly from activities to households and other institutions, data do not tend to come that way. Household surveys do not usually provide sufficient information on the activity codes for each and every component of factor income received by households. Supplementary information has to be sought from Labour force surveys and production surveys. It is therefore much more straightforward to compile the complete SAM and carry out manipulations on it afterwards, either by reducing the number of accounts or by eliminating blocks of accounts altogether.

In contrast with extended input-output tables, 'extended SAMs' are an entirely different proposition. A SAM in the sense we are considering it is representative of the whole economic system, but several extensions have been put forward to take into account other linkages. These include, for example, environmental linkages (SESAME and extensions), food supply and the food chain, financial accounts and the flow of funds, demographic linkages, and various other ways of accounting for differences in lifestyles. All of these extensions fall outside the present discussion of SAM construction issues; they require extended datasets with links and bridges to the core SAM system.

3. Compilation issues

In terms of the early published material, Pyatt and Roe, *et al* (1977) set out in detail how the 1970 Sri Lanka SAM was constructed. This study set out a basic blueprint for future studies, and it has been replicated many times since. The Sri Lanka SAM turned out to be a compromise between the desire to produce a matrix with sufficient detail to meet a range of analytical and modelling objectives while not stretching beyond what is credible given the relative paucity of information available. Access to some results from a household survey (Sri Lanka Socio-Economic Survey, 1969-1970) was crucial of course but in this study it is interesting to note that no recent input-output table was available. A much more ambitious SAM – at the time we thought it would be *the* definitive study – was the Malaysia SAM, also compiled for the year 1970 (Chander *et al*, 1980). This also benefited from the availability of a major household survey (Malaysia Household Expenditure Survey, 1973, supplemented by the Malaysia Post-Enumeration Survey, 1970) but in this case there was also a very detailed set of commodity balances. Again the compilation procedures are set out in considerable detail in Pyatt and Round (1984). Amongst the most innovative features of this study was to work with quite detailed factoral and household classifications from which subsets were chosen for eventual compilation of the SAM. At around the same time Downey and Keuning, *et al* (1982) were assembling a similarly detailed SAM for Indonesia, again based on very good household survey and commodity balance data. Several studies followed and, based on these experiences, Keuning and de Ruijter (1988) established a useful set of guidelines for constructing SAMs representing the overall design, stages in construction, potential data sources, and some discussion of error identification and data reconciliation methods. This was subsequently followed up by an extended version based on the compilation of a SAM for Ecuador in 1975 (Alarcon, *et al*, 1991).

In view of these quite well-established sets of guidelines, the remainder of this paper will focus on selected issues that arise in the SAM compilation process and which continue to be problematic. It is based on some relatively recent experience in compiling a SAM for Ghana (Powell and Round, 1998) together with some more casual observations on other recently-compiled SAMs. As already noted, in the case of Ghana the SAM was compiled alongside an implementation of the 1993 SNA and hence involved a major revision of the Ghana national accounts. So there were few statistical benchmarks to work with. However all of this effort was carried out at a time when there was plenty of other statistical activity including a series of household surveys (i.e. the Ghana Living Standards Survey, GLSS) and considerable empirical research on poverty and living standards. So the opportunity to compile a SAM for Ghana seemed unprecedented at the time. Subsequently, some policy models and other analyses have been based on it[15] although the present discussion will be confined to constructional aspects. In fact we shall concentrate on one particular aspect, the treatment of household sector activity and especially on the problem of estimating the incomes and outlays of households from household survey data, as this is so central to the construction of any SAM.

Measuring household economic activity[16]

It is well-known that in many developing countries' national accounts, especially in Africa, there is only a limited use of household survey information. Historically, consumption expenditure has been estimated as a residual even though it accounts, on average, for about 60 per cent of final expenditure. Ravallion (2001) has recently noted the marked discrepancy between estimates of mean per capita private consumption (national accounts) and mean per capita expenditure (household surveys). Of course, national accounts and household surveys are not measuring the same thing, but the discrepancy is nevertheless significant. The new SNA emphasis on compiling accounts at the institutional sector level means that eventually output, income, consumption, savings and investment all have to be estimated at the sectoral level too. While different survey instruments throw light on different sectoral contributions it is clear that household surveys are bound to be the principal and possibly only source of information for the household sector. Even so, there is some

[15] This has comprised fixed-price and accounting multiplier analysis (Powell and Round, 2000) and a poverty impact analysis based on a CGE model (Bussolo and Round, 2001).

[16] This section draws from some joint work with Harold Coulombe and Andrew McKay (Coulombe, McKay and Round, 1996).

way to go in knowing how best to tackle some difficult problems that arise in interpreting household survey results.

Clearly, production and business surveys are the basic survey instruments for measuring production activity, supplemented by dedicated surveys of agriculture, services, transport, distribution, and construction. These surveys have usually been carried out using the establishment as the basic sampling unit, often with a sampling frame based on a threshold unit size defined in terms of a minimum number of employees. This means that small-scale and informal sector activity is excluded from the surveys, and allowances or adjustments for this omission have typically had to be made often in a quite arbitrary way in order to arrive at a measure of total activity. Household surveys, such as those with the scope of the Ghana Living Standards Survey (GLSS), offer an opportunity to measure activity for this excluded segment more directly.

There are different kinds of household enterprise. They include both micro-enterprises (which hire employees) and family enterprises (which operate on own account) and either of these may engage in both formal and informal activities, though to different degrees. Also, family enterprises may undertake a range of subsistence and non-market activity as well as market-based activity and to try to capture this creates an extra challenge for compilers of SAMs. Even on the basis of established rules on the boundary of production, an imputed value of all such activity should be included in the estimates as part of the economic output of households, and this generates imputed incomes and expenditures for those households engaged in this activity. The extent to which earlier SAMs – or the national accounts even – have taken proper account of informal and/or non-market activity is uncertain and unclear in most cases. But this problem was tackled head-on in constructing the Ghana SAM.

Household level accounts

Most households in developing countries do not keep formal records of their production, consumption, saving or investing activities. However, based on the GLSS and similar multi-topic household surveys, it is possible to produce estimates of these components by carefully reconstructing the production, income and outlay and, to some extent, the accumulation accounts for each household in the sample. The process is obviously problematic and we shall illustrate some of the problems encountered and sketch some possible solutions based on our experience for Ghana.

[TABLE 3 HERE]

Based on the GLSS data estimates of ten separate components of incomes and expenditures were compiled at the individual household level (Johnson, McKay and Round, 1990). These components are listed in Table 3. Although the components are fairly aggregative, a distinction is maintained here between agricultural and non-farm enterprises, imputed and non-imputed items, food and non-food items, as well as rent, and various transfer incomes and outlays. The components are set out in the form of income and outlay accounts, and the imputed items are set side by side to highlight the fact that a matching item appears on both sides of the accounts for a particular household.

Total income is therefore the sum of components 1 to 6 and total current expenditure is the sum of components 7 to 10. As there are no direct estimates of household savings in the GLSS item 11 is derived as a balancing item and our expectation is that this should be non-negative in most cases. Hence, inclusive of savings, total income should equal total outlay in an accounting sense.

Information can be drawn from different sections of the GLSS survey to provide alternative estimates for some of the income components. A particular example is non-farm enterprise income which can be estimated in three ways; first, from responses to questions on self-employment income; second, on enterprise profits; and third by calculating operating surplus derived by subtracting costs from revenues of the household enterprise (Coulombe, McKay and Round, 1996; and Vijverberg, 1991). The problem is that the estimates differ widely[17] and, moreover, there is no clear basis for preferring one method of estimation to another. For employee compensation, perhaps the most consistently reported income component, LSMS surveys report information on the main job much better than secondary jobs.

As regards the imputed components, not surprisingly, the values of these income and expenditure

[17] On the face of it the preferred estimate of operating surplus would be the one compiled from revenues net of costs. However, household respondents have difficulty in separating out the intermediate costs of enterprises from consumer expenditures, so in the GLSS, as in LSMS surveys of other African countries, the majority of the estimates turned out to be negative.

components were found to be substantial, especially the imputed food expenditures of households in rural areas. But even the imputed non-food components were significant, underlining the general importance of non-market activity, especially in rural areas. Having established a measure of the volume of non-market activity there is still a problem of choosing the appropriate imputations to convert these quantities into values. For Ghana market price equivalents were used but valuation remains a potentially significant issue.

Remittances received and remittances paid out are listed as separate aggregates in the household incomes and outlays in Table 3. From the point of view of compiling a SAM the problem here is that there is rarely any indication of the source or destination of these flows (by sector) or even what they constitute. Moreover there is good reason to believe that the amounts reported may be very unreliable.[18] 'Other income', including some investment/asset incomes, are often reported in an inconsistent and erratic manner.

Reliability, income estimates, and negative household savings

Aside from the practical difficulties arising from gleaning information from the GLSS and arriving at estimates, there is a more general problem about reliability. We know that some income and expenditure components may be more reliable than others and in this regard there is a particular concern about the income components especially (McKay, 2000).

The reliability issue is most clearly manifested when the estimates of total incomes are compared with total expenditures at the individual household level. To illustrate this, just consider the following results. In the case of the GLSS survey for 1988-89, the shortfall of estimated household income relative to consumption across the sample for Ghana as a whole amounted to 32.3 per cent. If the imputed components are excluded, which are of course common to both the income and expenditure aggregates, then the estimated shortfall rose to 41.2 per cent. Put another way, this suggests that 81.5 per cent of all households in the sample have effectively reported negative savings. Now while it is perfectly reasonable to expect individual households, or even the household sector as a whole, to dissave in some years, it is unlikely to be on this scale, especially as this level of

[18] Household respondents may not know about (or may wish to conceal) income remittances received or paid by household members as they may be considered too sensitive to report.

income shortfall is replicated in other rounds of the Ghana household surveys. Overall this seems hardly credible and the conclusion is inescapable that income has been under-recorded and possibly to a substantial extent (Coulombe, McKay and Round, 1996).

There are obviously major conceptual and practical difficulties in recovering good income data from household surveys. Deaton (1997), and others, have expressed doubts about whether the effort is actually worthwhile. But it must be said that without any information on the income side there would be no basis at all for tracing the interconnection between production and income distribution, and this would be a major setback for development policy analysis. In particular, it would simply not be possible to compile a SAM. Deaton brings the survey problem to the fore by highlighting the negative savings issue. He remarks that 'although there are often good reasons to doubt the absolute accuracy of the national income estimates, the fact that surveys repeatedly show large fractions of poor people dissaving, and apparently doing so consistently, strongly suggests that the surveys underestimate savings' Deaton (1997; 32).[19]. In fact, as many countries' national accounts do not use household survey data it is unlikely that any under-reporting of income or savings is a particular contributory factor to the unreliability of the national accounts aggregates. Aggregate domestic savings can be determined via the basic macroeconomic aggregates identities, and it is only when separate institutional sector accounts are included that a problem arises. Thus, for example, if we believe aggregate savings are correct and household savings are underestimated, then it would follow that either corporate enterprise or government savings are overestimated.

This problem with underestimating household income and savings seriously affects our ability to compile a credible SAM. The problem was faced early on, for example in compiling the Malaysia SAM (Pyatt and Round, 1984). And the solution then was to raise labour incomes by a scale factor sufficient to yield positive savings for all but one of the household groups, the assumption being that wages and especially household enterprise income are under-recorded in the household survey results. In the light of the GLSS results, and the seriously high levels of implied negative savings, the problem is manifested again in the Ghana SAM. The solution adopted this time was to develop a

[19] McKay (2000) reports some evidence from 14 household surveys, most were not LSMS or 'Integrated Household' surveys (although some were) and, of these, 11 implied negative savings on average.

more formal procedure, though the principle behind it is similar to the one used for the Malaysia SAM. Scale factors were determined and applied to each of the separate income components. These were determined by a procedure based on the 'reconstructed' household accounts which ensured that total income for each household group was raised sufficiently to match total expenditure (Coulombe, McKay and Round, 1996). In a subsequent revision to the Ghana SAM the assumption of zero aggregate savings have been replaced by an aggregate savings rate based on national accounts estimates. This was also the basis of the method used in the construction of the Ecuador SAM (Alarcon, *et al*, 1991). An alternative approach would be to derive savings from the financial balance sheets (as changes in net worth) but in most cases this would also be hampered by data limitations.

4. Balancing techniques

The SAM can be a laborious and demanding exercise and there are often many inconsistencies between data sources during assembly that cannot easily be eliminated. In many of the attempts to construct SAMs statisticians have resorted to a variety of data reconciliation methods to smooth out discrepancies. Obviously, data reconciliation should always involve a detailed re-inspection of initial estimates in order to eliminate, as far as possible, any discrepancies caused by inconsistencies in timing, treatment and definition. The problem is that many discrepancies remain, and these violate the consistency requirements inherent in the basic accounting structure. Clearly, while 'consistency' does not imply 'accuracy' a virtue of the SAM is that it does provide users with a consistency check to challenge any of the estimates at any stage. Revising one transaction will have implications for other transactions in the system.

Apart from informal methods of adjustment based on judgement, informed or otherwise, several formal methods of data reconciliation have been proposed and there now exists a substantial literature on the subject. So the purpose of this section is to review the methods that have been used in practice and to make some observations on their comparative properties.

Macro and Micro SAMs

Reinert and Roland-Holst (1997) have suggested that the construction of a SAM should begin by recasting the macroeconomic accounts for the economy into a simple matrix tableau, a so-called Macro SAM. This of course assumes that the macroeconomic accounts exist and that the aggregates are to be relied on without further revision or adjustment. Either or both of these assumptions may be questionable in practice. But at the next stage the Macro SAM is followed by the construction of the detailed Micro SAM. Most of the SAMs compiled under various IFPRI modelling projects[20] have followed this procedure. Clearly, if the SAM is being compiled in tandem with the national accounts (as in the Ghana study) then the concept of a Macro SAM for benchmarking purposes, and the sequencing from Macro SAM to Micro SAM is much more fluid and uncertain. Also, there are instances where SAMs may be compiled for regional or village economies where the aggregates are not predetermined in this way. So a strict adherence to this procedure may not apply in all cases. More importantly, in compiling the earliest SAMs for Sri Lanka, Swaziland, and Malaysia, as well as in the more recent Ghana study, there were clear instances where the national accounts aggregates could be questioned as a direct result of compiling the SAM, either because of the availability of new or additional data or because one set of estimates simply did not match another. This suggests that extra caution should be exercised in a strict application of the 'from Macro to Micro' rule, especially if household survey data are used to construct the SAM and if the corresponding national accounts have not relied on these data or only to a minimal extent.

Data reconciliation methods

Informed judgement of local experts and compilers played a major part in reconciling discrepancies in some of the early SAMs (for example, Pyatt and Roe, *et al*, 1977). However, the procedure was not as arbitrary as it might at first seem. There were essentially three steps involved in the judgement approach. First, the initial data were set alongside each other in the accounting framework to take initial stock of the problem. Secondly, a qualitative judgement was taken on the relative reliability of the alternative estimates, relying on expert local advice. Thirdly, after choosing the most reliable estimates, further scaling and adjustments were made manually to achieve consistency. Gaps and missing entries were usually handled differently from inconsistent estimates.

Sometimes missing entries were estimated directly as residuals using the accounting constraints or they were eliminated by an aggregation of accounts. The whole process of smoothing data sets into a consistent set of estimates using judgement involves some iteration between the stages.

In a preface to the Sri Lanka volume Stone (1977) pointed out that these essentially subjective adjustment techniques were unscientific, and he posed the question whether more formal mathematical methods could provide 'better' estimates. In fact, a large number of algorithms had already been proposed on the adjustment of unbalanced data matrices and there have been more since. SAMs and input-output tables are but two of a whole range of practical contexts[21] where the need to balance initially unbalanced data matrices has arisen. Stone, Champernowne and Meade (1942) had already suggested one method in the context of adjusting social accounting estimates.

As a useful precursor to further discussion it can be noted that Schneider and Zenios (1990) have suggested that most matrix balancing problems fall into two categories of problem:

Problem 1: If $X = \left[x_{ij}\right]$ is an m x n nonnegative matrix and u and v are positive vectors of orders m and n respectively then determine an m x n matrix X^* 'close to' X such that

$$\sum_j x_{ij}^* = u_i$$
$$\sum_i x_{ij}^* = v_j$$

and $x_{ij}^* > 0$ if and only if $x_{ij} > 0$ (all i, j).

Problem 2: If $X = \left[x_{ij}\right]$ is an (n x n) nonnegative matrix and u and v are positive vectors of orders m and n respectively then determine an (n x n) matrix X^* 'close to' X such that

$$\sum_j x_{ij}^* = \sum_j x_{ji}^* \qquad \text{(all } i\text{)}$$

[20] See for example, M Fontana and P Wobst (2001).
[21] Schneider and Zenios (1990) cite examples in transportation and traffic flows, demography and migration flows; and estimating a range of practical problems in estimating transition matrices in applied Markov models.

and $x_{ij}* > 0$ if and only if $x_{ij} > 0$ (all i, j).

In an accounting matrix context the two problems characterise two distinct classes of matrix balancing situation. The first is often encountered in balancing (and updating) input-output tables to satisfy known row and column constraints, while the second is the more usual SAM balancing problem where, though account totals may themselves be unknown, there are accounting restrictions on corresponding row and column totals. From these fundamental problems a range of extended problems have been considered. In all cases a basic consideration is to choose a criterion to define a measure of 'closeness' and thereafter to carry out a constrained minimisation solution, where the constraint set may involve additional and more complex constraints to the basic restrictions in problems 1 and 2. For example the extensions may involve equality relationships between selected elements or groups of elements, or inequality relationships represented by lower or upper bounds on sets of elements. Against this general background we now concentrate on a subset of algorithms which have proved to be both popular and operational in SAM construction, and then we shall make some general observations on their use.

RAS method

A classic method of matrix adjustment suggested in the input-output literature is to generate a new matrix X^* from an existing matrix X (to satisfy new known row and column totals) by applying row and column multipliers, r and s respectively

$$X^* = \hat{r}\, X\, \hat{s} \qquad (1)$$

The (2n-1) unknown multipliers are determined by the (2n-1) independent row and column restrictions using an iterative adjustment procedure. Günlük-Senesen and Bates (1988) and others have shown this to be equivalent to a 'type 1' problem involving the minimisation of

$$L(X^* : X) = \sum x_{ij}^* \ln\left({x_{ij}^*} \big/ {x_{ij}} \right) \qquad (2)$$

subject to known row and column sum constraints. The RAS method has been extended to accommodate uncertainty in the row and column totals and negative elements, which would otherwise be problematic (Günlük-Senesen and Bates, 1988).

Since RAS solves a 'type 1' problem it is not an effective algorithm for balancing a SAM, although it is certainly useful for balancing submatrices of SAMs. However, Schneider and Zenios (1990) have suggested another algorithm, referred to as a 'Diagonal Similarity Scaling' (DSS) method which is formally similar to RAS but is designed to solve a 'type 2' problem. In this case

$$X^* = \hat{d}\, X\, \hat{d}^{-1} \tag{3}$$

where d is determined by iteratively eliminating the discrepancies between correponding row and column sums. Again the problem can be expressed as a constrained minimisation with the minimand

$$L(X^* : X) = \sum x_{ij}^* \left(\ln\left(x_{ij}^* \big/ x_{ij} \right) - 1 \right) \tag{4}$$

and subject to 'type 2' problem constraints. In principle this might seem to be an appropriate for balancing entire SAMs except that it relies on scaling adjustments across whole rows and down whole columns and this misses an important feature of a SAM. A SAM comprises blocks of different kinds of transactions and the estimates of each block may be derived from different sources and hence subject to quite different degrees of reliability. Therefore it is usually not appropriate to impose a uniform scaling adjustment and even this variant of RAS may not be very suitable for balancing out inconsistencies.

Stone-Byron method

Another method, which is analogous to the method of restricted least squares, was first discussed in a SAM context by Stone (1977) although it had been suggested many years previously by Stone, Champernowne and Meade (1942) for adjusting more general sets of social accounting estimates. But only with modern computing capacity has the method become a practical proposition (Byron, 1978). The method has since been utilised in compiling several SAMs. It can be described briefly as follows.

As before, let X be an initial estimate of a SAM, or a part of a SAM. Suppose also that there are known sets of desired *linear* constraints between the elements of the SAM. These may either be the standard accounting restrictions as in any 'type 2' problem, or linear restrictions on sums of subsets of elements (e.g. sums of sectoral value added to equal total GDP) or restrictions on ratios of elements (e.g. fixed savings ratios). As before, let X^* be the revised SAM which satisfies the constraints. Express the elements of X and X^* as ordered elements of the vectors x and x^* and define a grouping matrix G (mainly containing 0, 1 and -1) and a restriction vector h to express the desired linear restrictions on the elements of X^* as follows

$$G\,x^* = h \tag{5}$$

Now, let V be a variance-covariance matrix associated with the vector x (or, equivalently, a matrix of reliability or tolerance estimates of the SAM) then by choosing a quadratic loss function (i.e. weighted least squares) as the minimand, it can be shown that

$$x^* = x - V\,G'(G\,V\,G')^{-1}(G\,x - h) \tag{6}$$

This has some desirable properties, including that in an analogous statistical context x^* can be interpreted as the best linear unbiased estimator of the vector of true elements. Also, the method accommodates multiple estimates of cells, as the restrictions can ensure that revised estimates become equal. The elements of V are not observed, but the usual assumption is to set all covariances between elements to be zero and to choose the variances relative to the size of the elements (x_{ij}). More usually, coefficients of variation are chosen subjectively in accordance with the perceived relative reliability of the different components. Thus, although there is compiler judgement, it enters at a second order rather than first order level as it is the tolerance factors rather than the estimates themselves about which judgement is being exercised.

In the context of our more general representation of balancing problems, Stone-Byron can be shown to be equivalent to a solution of a 'type 2' problem, where the minimand is of the form

$$L(X^*: X, V) = \sum_{i,j} \left(x_{ij}^* - x_{ij}\right)^2 \big/ v_{ij} \tag{7}$$

where v_{ij} are analogous to the variances of the elements, and where all the restrictions are linear. Expressed in this way, the problem can easily be extended to cases where non-linear restrictions are imposed although of course the solution would not then be of the same neat analytical form as depicted in (6).

Cross-entropy method

A third balancing method, which has been used extensively by Sherman Robinson and his associates in the IFPRI group for compiling and balancing several SAMs, is the cross-entropy (CE) method (Robinson, Cattaneo and El-Said, 2001)[22]. In essence, the method is formally similar to the generalised RAS method, which we saw earlier uses an entropy-based minimand and a constraint set appropriate to a 'type 2' balancing problem (McDougall, 1999). However there are some significant differences and additional complexities. First, the minimand is based on the derivation of a coefficient structure for the SAM, A^*, where the initial column coefficients are $A = [a_{ij}]$ rather than transaction flows, X^*. Second, the minimand now has to include the estimation of a set of error weights, w_{ih}, which are part of the generation of error variables, e_i.[23]

$$L(A^*, W : A) = \sum_{i,j} a_{ij}^* \ln\left(a_{ij}^* \big/ a_{ij} \right) + \sum_{i,h} w_{ij} \ln\left(n w_{ih} \right) \tag{8}$$

The error variables, e_i , which are not part of the minimand, serve to bring corresponding row and column sums into balance. Third, the error weights and error variables are part of a more complex constraint set, which, in addition to the accounting constraints and possible additional (linear and non-linear) constraints on sets of transactions, now have to maintain the accounting relationships between coefficients and flows.

Other methods

RAS, Stone-Byron, and cross-entropy (CE) are not the only formal methods for balancing matrices, but they are representative of the methods that have been used to balance SAMs in practice. Beyond these methods, other alternatives utilise variations in the choice of minimand and two are particularly worthy of mention. The first alternative is another 'quadratic' minimand, viz

$$L(X^* : X) = \sum_{i,j} \left(x_{ij}^* - x_{ij} \right)^2 \big/ x_{ij}^{\,2} \tag{9}$$

[22] There are several antecedents to this in the literature; this reference is included as it represents a comprehensive discussion of the method.

[23] The way this works is that h is the order of the set of 'error support values', s_h , usually three (to include zero and two symmetric values) and then error variables are formed from $e_i = \sum_h w_{ih} s_h$.

In fact, this is the same as Stone-Byron for the special case where the coefficients of variation of all elements are equal; that is, where the initial estimates are judged to be of equal relative reliability. Usually our prior judgement about the relative reliability of different data sources will allow us to do better than this and therefore, in general, Stone-Byron would be preferred to the quadratic minimand.

A second alternative is suggested by the similarity of the CE method (problem 2) to RAS (problem 1) (McDougall, 1999). A simple hybrid follows if a cross-entropy minimand based on transactions (i.e. flows) rather than coefficients is combined with 'type 2' problem constraints and possibly additional (linear and/or nonlinear) constraints. However it is interesting to note that, under particular circumstances[24], the entropy function is approximated by the function

$$L(X^*:X)=\sum_{i.j} x_{ij}^* \ln\left(x_{ij}^* \big/ x_{ij}\right) \cong \sum_{i,j}\left(x_{ij}^* - x_{ij}\right)^2 \big/ x_{ij}^* \qquad (10)$$

so this in turn approximates to another special case of the Stone-Byron method if the variance were set equal to the adjusted coefficient. So the analytical correspondences between methods are quite close.

Is there a preferred method for balancing a SAM?

Robinson, *et al* (2001) carry out a range of Monte Carlo experiments which suggest the superiority of the CE method over RAS in those circumstances (under problem 1 conditions) where comparisons are valid. Günlük-Senesen and Bates (1988) also conduct experiments with several balancing methods under similar problem 1 conditions and observe more mixed outcomes. One problem in carrying out experiments is that the criteria for assessing success (the measures of closeness of an adjusted matrix to a 'true' matrix) are intimately related to the choice of minimand. Therefore there is an inherent bias built into any experimentation, which make objectivity difficult.

The relatively close analytical relationships between the most frequently-used alternative methods for

[24] The condition is that $\sum_{i,j} x_{ij}^* = \sum_{i,j} x_{ij}$

balancing SAMs suggest that if the required adjustments are relatively small then the differences between the methods are likely also to be small. Schneider and Zenios (1990) applied five methods to the unbalanced and highly aggregated SAM (n = 5) used by Stone (1977) to demonstrate an application of the Stone-Byron method. The differences between the resulting balanced SAMs were of a very small order of magnitude. Of course, for a higher dimensional SAM or where the required adjustments are large then the differences might well be greater.

In spite of the apparent preference for the cross-entropy (CE) method by many compilers of SAMs, the Stone-Byron method (possibly extended to include additional constraints) does seem to have some advantages over alternative methods. In particular it allows us to incorporate judgement on the relative reliability of data sources and is therefore closer to the spirit of the problem at hand. Also, it accommodates initial multiple estimates, a common feature in SAM compilations. In fact, the Stone-Byron method was used to balance the Ghana SAM (Powell and Round, 1998) although, as the SUT table had been balanced prior to the rest of the SAM, the dimension of the unbalanced matrix was considerably reduced.

Finally, now that matrix balancing methods are so convenient and easy to use it is important to add a cautionary note and to remind ourselves that they are unlikely to ever be an adequate substitute for the careful assembly of primary data (initial estimates). A premature recourse to mechanical balancing methods can sometimes be used as a substitute for a more careful reappraisal of the source data. There is then a danger in assuming that a balanced SAM which is based on a set of weak and possibly unrepresentative initial estimates is going to be representative of the economy in question. It is a far better strategy to concentrate on improving the initial estimates and to use the smoothing techniques only *in extremis* or as a final resort.

5. Conclusions

In the three decades since SAMs first assumed some prominence as a framework for data and development policy analysis, there have been a few developments in terms of compilational techniques. As demonstrations of what can be achieved and what to aim at, the earliest SAMs were bold and innovative constructions. 'Doing the best with what we have' and making 'the whole greater than the sum of the parts' were catch-phrases which captured the pioneering spirit of those early exercises. There has been a period of consolidation and replication which has allowed some analytical work to build further on the conceptual framework, and for extended SAMs to be developed and pursued further too. However, some longstanding and quite difficult compilation problems remain and need to be addressed quite urgently. Three main conclusions can be drawn from this paper.

First, notwithstanding the achievements of the 1993 SNA, it is not safe to assume that a SAM can easily be achieved as a bi-product of the adoption and implementation of the SNA system. Experience in compiling the Ghana SAM has demonstrated this fact. Many of the most analytically useful and interesting interrelationships would almost certainly not be captured by a straightforward application of the SNA system.

Secondly, the use of household survey information, though fundamental and crucial, remains problematic. The income side is well-known to be generally weaker and more unreliable than the expenditure side, perhaps in some regions more than others, it may be dangerous to generalise. This, too, continues to offer significant challenges in compiling SAMs. The Ghana exercise has also highlighted the importance of accounting for subsistence activity and non-cash transactions. The effect of these items could be considerable in terms of measures of living standards. Overall, the many problems to do with measuring income and the use of household surveys receives very little mention in the literature on compiling SAMs, but it is nevertheless of overriding importance.

Finally, the paper devotes some space to a topic which is perhaps afforded a correspondingly undue allocation of space in the literature - that is, the question of matrix balancing and data reconciliation methods. It would surely be preferable to devote most energy to a careful assembly of the initial estimates and to rely on mechanical methods only as a last resort. A method of smoothing weak initial estimates is unlikely to generate reliable final estimates, however efficient that method is. However, a review of the three or four methods commonly used suggests that the Stone-Byron method, possibly augmented by non-linear restrictions, is still likely to offer the most flexibility to compilers of SAMs. But faced with small adjustments to the initial estimates there is very little to choose between any of the methods on offer.

References

Alarcon Rivero, J. V., J van Heemst, S Keuning, W de Ruijter & R. Vos (1991) **The Social Accounting Framework for Development; Concepts, Construction and Applications** (Gower Press).

Bussolo, M. & J. I. Round (2001) Redistribution and Poverty: Some Experiments Based on a CGE Model for Ghana, ESRC Development Economics Study Group Conference (April), University of Nottingham, Nottingham.

Byron, R P (1978) The Estimation of Large Social Account Matrices, **Journal of the Royal Statistical Society**, Series A, 141 (3); 359-367.

Cambridge, DAE (1962) **A Programme for Growth**, 12 volumes (Chapman and Hall).

Chander, R., S. Gnasegarah, G. Pyatt & J. I. Round (1980) Social Accounts and the Distribution of Income : The Malaysian Economy in 1970, **Review of Income and Wealth**, Series 26, No.1; 67-85.

Cockburn, J. (2001) Trade Liberalisation and Poverty in Nepal: A Computable General Equilibrium Micro Simulation Analysis, Discussion Paper 01-18, CREFA, Université Laval.

Coulombe, H., A. McKay & J. I. Round (1996) Estimating the Contribution of the Informal Sector to the Ghana GDP, in Ghana Statistical Service, **Measuring Informal Activity in Ghana**, Proceedings of a Ghana Statistical Service/ Overseas Development Administration Conference, Accra, Ghana (revised).

Deaton, A. (1997) **The Analysis of Household Surveys: A Microeconometric Approach to Development Policy** (Johns Hopkins University Press: Baltimore).

Downey, R. & S. Keuning, and staff of the CBS (1982) An Indonesian SAM, Central Bureau of Statistics, Jakarta.

Fontana, M. & P Wobst (2001) A Gendered 1993-94 Social Accounting Matrix for Bangladesh, Macroeconomics and Trade Division, Discussion Paper No. 74, IFPRI, Washington DC.

Günlük-Senesen, G. & J. M. Bates (1988) Some Experiments with Methods of Adjusting Unbalanced Data Matrices, **Journal of the Royal Statistical Society**, Series A, 141; 473-490.

Johnson, M., A. McKay & J. I. Round (1990) Income and Expenditure in a System of Household Accounts: Concepts and Definitions, **Social Dimensions of Adjustment in sub-Saharan Africa Working Paper No 10** (The World Bank, Washington D C).

Heston, A. (1994) A brief account of some problems in using national accounts data in level comparisons and growth studies, **Journal of Development Economics**, 44; 29-52.

Keuning, S. (1991) Proposal for a Social Accounting Matrix which Fits into the Next System of National Accounts, **Economic Systems Research**, 3 (3); 233-248.

Keuning, S. (1998) A Powerful Link Between Economic Theory and Practice: National Accounting, **Review of Income and Wealth**, Series 44, No 3; 437-445.

Keuning, S & W de Ruijter (1988) Guidelines for the Construction of a Social Accounting Matrix, **Review of Income and Wealth**, Series 34 No 1; 71-100.

McKay, A. D. (2000) Should the Survey Measure Total Household Income?, Chapter 17 in M. Grosh & P. Glewwe (eds) **Designing Household Survey Questionnaires for Developing Countries: Lessons from 15 Years of the Living Standards Measurement Study**, Volume 2 (The World Bank, Washington D.C.).

McDougall, R. A. (1999) Entropy Theory and RAS are Friends, GTAP Working Paper No 300, Department of Agricultural Economics, Purdue University.

Miyazawa. K. (1976) **Input-Output Analysis and the Structure of Income Distribution**, Berlin, Springer.

Powell M. & J. I. Round (1998) A Social Accounting Matrix for Ghana, 1993, Ghana Statistical Service, Accra, Ghana.

Powell M. & J. I. Round (2000) Structure and Linkage in the Economy of Ghana: A SAM Approach, in E Aryeetey, J Harrigan & M Nissanke (eds) **Economic Reforms in Ghana: Miracle or Mirage** (James Currey Press, Oxford): 68-87.

Pyatt. G. (1991a) Fundamentals of Social Accounting, **Economic Systems Research**, 3 (3); 315-341.

Pyatt, G. (1991b) SAMs, the SNA and National Accounting Capabilities, **Review of Income and Wealth**, Series 17, No 2; 177-198.

Pyatt, G. (1999) Some Relationships Between T-Accounts, Input-Output Tables and Social Accounting Matrices, **Economic Systems Research**, 11 (4); 365-388.

Pyatt, G. (2001) Some Early Multiplier Models of the Relationship between Income Distribution and Production Structure, **Economic Systems Research**, 13 (2); 139-164.

Pyatt G. & A. R. Roe, with R .M. Lindley, J. I. Round and others (1977) **Social Accounting for Development Planning with Special Reference to Sri Lanka** (Cambridge University Press, Cambridge).

Pyatt, G. & J. I. Round, with J. Denes (1984) Improving the Macroeconomic Database: A SAM for Malaysia, 1970, World Bank Staff Working Paper No 646 (The World Bank, Washington D C).

Pyatt, G. & J. I. Round (1977) Social Accounting Matrices for Development Planning, **Review of Income and Wealth**, Series 23, No.4; 339-364.

Pyatt G., & J. I. Round (eds) (1985) **Social Accounting Matrices: A Basis for Planning** (The World Bank, Washington D C).

Pyatt G. & E. Thorbecke (1976) **Planning Techniques for a Better Future** (ILO, Geneva).

Ravallion, M. (2001) Measuring Aggregate Welfare in Developing Countries: How Well Do National Accounts and Surveys Agree?, Macroeconomics and Growth Working Paper No 2665 (The World Bank, Washington D C).

Reinert, K. A. & D. W. Roland-Holst (1997) Social Accounting Matrices, J. F. Francois & K. A. Reinert (eds), **Applied Methods for Trade Policy Analysis: A Handbook** (Cambridge University Press, Cambridge); 94-121.

Robilliard, A-S, F. Bourguignon & S. Robinson (2001) Crisis and Income Distribution: A Micro-Macro Model for Indonesia, mimeo, The World Bank, DIAL, and IFPRI.

Robinson, S., A. Cattaneo & M. El-Said (2001) Updating and Estimating a Social Accounting Matrix Using Cross Entropy Methods, **Economic Systems Research**, 13 (1):47-64.

Schneider, M. H. & S. A. Zenios (1990) 'A Comparative Study of Algorithms for Matrix Balancing', **Operations Research**, 38; 439-455.

SNA (1993) **System of National Accounts 1993**, Series F, No 2, Rev 4 (United Nations Statistics Division, New York).

Stone, R (1977) Forward to G Pyatt & A Roe *et al*, **Social Accounting for Development Planning** (Cambridge University Press, Cambridge); xvi-xxxi.

Stone, J. R. N., D. Champernowne & J. E. Meade (1942) The Precision of National Income Estimates, **Review of Economic Studies**, Vol IX, No 2; 111-125.

UNSD (1999) **A System Approach to National Accounts Compilation**, Handbook of National Accounting, Studies in Methods, Series F No 77, Department for Economic and Social Affairs, Statistics Division, United Nations, New York.

UNSD (2000) **Household Accounting Experience in Concepts and Compilation, Vol 1 Household Sector Accounts and Vol 2 Household Satellite Extensions**, Handbook of National Accounting, Studies in Methods, Series F No 75, Department for Economic and Social Affairs, (United Nations Statistics Division, New York).

Vijverberg, W. P. M. (1991) 'Measuring Income from Family Enterprises with Household Surveys', **Living Standards Measurement Study Working Paper No 84** (The World Bank, Washington D C).

Table 1: *A Basic Social Accounting Matrix (SAM)*

Account			(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Factors of production		(1)						Gross value added payments to factors		*Net* factor income from RoW
Institutions (Current accounts)	Households	(2)	Labour and mixed income	Inter-household transfers	Distributed profits to households	Current transfers to households				*Net* current transfers from RoW
Institutions (Current accounts)	Corporate enterprises	(3)	Operating surplus			Current transfers to companies				*Net* current transfers from RoW
Institutions (Current accounts)	Government (&NPISHs)	(4)		Direct taxes	Direct taxes		*Net* taxes on products			*Net* current transfers from RoW
Production	Goods and services	(5)		Household consumption		Government consumption		Intermediate consumption	Fixed capital formation and change in stocks	Exports
Production	Activities	(6)					Domestic sales			
Combined capital accounts		(7)		Household savings	Corporate savings	Government savings			Capital transfers	*Net* capital transfers from RoW
Rest of World (combined account)		(8)					Imports		Current external balance	
TOTALS			Factor Income payments	Current household outlays	Current corporate outlays	Current government outlays	Supply of products	Costs of production activities	Capital outlays	Aggregate receipts from RoW

Note: Row totals are not shown but they match column totals.

Table 2: *1993 SNA (abridged) in matrix format*

SNA Account			(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
0	Production: Products	(1)		Intermediate consumption				Final consumption expenditures	Fixed capital formation and change in stocks	Exports
I	Production: Activities	(2)	Domestic sales							
II.1.1	Primary income distribution: Generation of income	(3)		Domestic product						*Net* employee compensation from RoW
II.1.2	Primary income distribution: Allocation of income	(4)	*Net* taxes on products		Income generation	Property income				*Net* property income from RoW
II.2, II.3	Secondary income distribution	(5)				National Income	Current transfers			*Net* current transfers from RoW
II.4	Use of income	(6)					Disposable income			
III	Accumulation	(7)						Savings	Capital transfers	*Net* capital transfers from RoW
V	Rest of World	(8)	Imports						Current external balance	
TOTAL			Supplies (purchasers' prices)	Activity inputs	Income generated	Income allocated	Income redistributed	Use of income	Capital expenditure	External current account flows

Source: Adapted from SNA (1993) Table 20.4. Row totals are not shown but they match column totals.

Table 3 Simplified household income and expenditure components.

Incomes	**Outlays**
1. Employee compensation	
2. Agricultural enterprise (*includes* item 7 (b))	7. (a) Food (actual) (b) Food (imputed)
3. Non-farm enterprise (*includes* item 8 (b))	8. (a) Non-food (actual) (b) Non-food (imputed)
4. Rent (*includes* item 9 (b))	9. (a) Housing (actual) (b) Housing (imputed)
5. Remittances received	10. Remittances paid and other outlays
6. Other income	
	Total expenditure
	11. **Balancing item: household savings**
Total income	Total outlay