R23 Database: Sources and Methods

Scott McDonald, Karen Thierfelder and Terrie Walmsley[1]

Addresses for correspondence:

Scott McDonald,	Karen Thierfelder	Terrie Walmsley
14 Moncrieffe Road,	Department of Economics	ImpactECON LLC
Sheffield,	US Naval Academy	
South Yorkshire,	Annapolis,	Boulder,
S7 1HR, UK.	Maryland, USA	Colorado, USA
E: jrs.mcdonald@gmail.com	E: thier@usna.edu	E: twalmsley@impactecon.com
T: +44 114 250 9904	T: +1 410 293 6887	T: +1 303 537 3237

Abstract

This paper reports on the development of the first globally consistent macroeconomic database. The database has been developed to support directly economic policy analyses, using the companion R23 model or other models, and indirectly by contributing to the development improved global databases, e.g., GTAP and WIOD. The process of developing the database has demonstrated the extent to which global macroeconomic databases are unreconciled and indicated extent of the divergence and the potential benefits from improving the underlying databases. The database, model and associated packages are open source.

Keywords: Global Social Accounting Matrix; National Accounts; Computable General Equilibrium.

JEL classification: E16; E01; D58.

[1] Scott McDonald is a Visiting Professor at Humboldt University, Berlin, Karen Thierfelder is Professor of Economics at the United States Naval Academy and Terrie Walmsley is Director at ImpactECON.

Table of Contents

1. Introduction

This paper provides a technical description of the R23 database, which is a multi-region (global) database for the R23 model, which is variant of the 123 model (de Melo and Robinson, 1989; Devarajan, *et al*., 1990) where the global dimension is based on the modelling of global trade relations used in the GLOBE model (McDonald *et al*., 2007 and 2013)[2]. The R23 model is calibrated with a database presented in the form of global Social Accounting Matrix (SAM); this variant uses data derived, with augmentation, from the World Bank and IMF databases and from the Global Trade Analysis Project's (GTAP) database. The augmentation of the database involves adding data on current account transactions that capture remittances, aid and other transfers and payments for foreign owned factors. The remittance and factor payments data are bilateral, i.e., receipts of remittances by each region are identified by source region; the aid data are both bilateral and multilateral, the latter being aid transfer effected through international institutions, e.g., development banks, etc.; and the other transfers data are only multilateral.

The structure of the database derives from the SAM approach to national accounting, e.g., Stone (1962a and b) and Pyatt (1991), and the System of National Accounts (UN, 2009) the SAM approach to modelling, e.g., Pyatt (1987), Drud *et al*., (1986). Moreover, the database adheres to the SNA's price system for the valuation of transaction, i.e., basic and purchaser prices, and SNA's production boundary. A global R23 SAM is **complete**, if all transactions are recorded, and **consistent**, if all transactions recorded as incomes are equal to the matching transactions recorded as expenditures. In theory, therefore, the transactions values reported in the R23 should conform with the national accounts data reported by countries that have compiled their national accounts on the basis of the SNA and have fully reconciled their national accounts. But, in the process of compiling the R23 database it is evident that even where national accounts have been reconciled within countries this is not

[2] While the R23 model is a member of a family of CGE models that model trade relationships using principles described in the 1-2-3 model, features in this model stem from other developments in CGE modelling. Among the models that have influenced this mode are the US Department of Agriculture model (Robinson *et al*., 1990; Kilkenny, 1991) and NAFTA models (Robinson *et al*., 1993), the IFPRI standard model (Lofgren *et al*., 2002), the PROVIDE Project model (McDonald, 2003), and the GTAP model (Hertel, 1997). The model owes substantial debts to the wider community of CGE modellers.

always reflected in the World Bank or IMF databases and that national accounts are not reconciled across regions.

A distinctive feature of the database is the use of a 'dummy' region, known as Globe, that allows for the recording of inter-regional transactions where either the source or destination are not identified. Examples of such transactions include trade and transportation margins and multilateral aid and other grant transfers remittances. The Globe construct provides a general method for dealing with any transactions data where full bilateral information is missing.

The rest of this paper is organised as follows. Section 2 provides an overview of the structure and requisite features of a Global Social Accounting Matrix (SAM) so as to give context to the rest of the paper. This is followed in section 3 by a description of the estimation principles and procedures adopted to develop the database, which provides a rationale for the development of the estimated prior matrices, discussed in section 4 and the practical estimation issues detailed in section 5. The limitations and data constrains are discussed in section 6 while section 7 provides concluding comments. A series of appendices provide additional information about the process and the resultant global R23 SAM.

2. Global Macro Social Accounting Matrix

The Global R23 SAM can be conceived of as a series of single region macro SAMs that record all macroeconomic current account transactions and are linked through the trade accounts; the regions are linked directly is through commodity trade transactions and bilateral transfers (remittances, aid, etc.) and indirectly through the demand and supply of trade and transport services and multilateral transfers that all go through the Globe regions. Specifically, the value of exports, valued free on board (*fob*) from source *x* to destination *y* must be exactly equal to the value of imports valued *fob* to destination *y* from source *x*, and since this holds for all commodity trade transactions the sum of the differences in the values of imports and exports by each region must equal zero. However, the resultant trade balances reported in the R23 database do not fully accord with current account balances because of other inter regional transactions (see McDonald and Sonmez, 2004).[3] This version of the R23 database addresses the implicit 'misspecification' of the current account balance in the GTAP database[4] by increasing the degree of current account detail with respect to factor payments and incomes, remittances and aid and other transfers. A description of the transactions recorded in a representative SAM for a typical region in the database is provided in Table 2.1.

A SAM is a transactions matrix; hence each cell in a SAM simply records the values of the transactions between the two agents identified by the row and column accounts. The selling agents are identified by the rows, i.e., the row entries record the incomes received by the identified agent, while the purchasing agents are identified by the columns, i.e., the column entries record the expenditures made by agents. As such a SAM is a relatively compact form of double entry bookkeeping that is complete and consistent and can be used to present the National Accounts of a country in a single two-dimensional matrix (see UN, 1993, for a detailed explanation of the relationship between conventional and SAM presentations of National Accounts). A SAM is complete in the sense that the SAM should record ALL the transactions within the production boundary of the National Accounts, and consistent in the sense that income transactions by each and every agent are exactly matched by expenditure transactions of other agents. A fundamental consequence of these conditions is that the row

[3] Extensions of the R23 model will include augmented data to include aid and remittance flows

[4] In reality the GTAP database records the trade account balance with all other current account transactions bundled together with the capital account.

and column totals of the SAM for each region must be identical, and hence the SAM provides a complete characterisation of current account transactions of an economy as a circular (flow) system. In the context of a global SAM the complete and consistent conditions need extending to encompass transactions between regions; this simply requires that each and every import transaction by a region must have an identical counterpart export transaction by another region. This is enough to ensure that the resultant global SAM provides a characterisation of current account transactions of the global economy as a circular (flow) system.

Given these definitions of a SAM the transactions recorded in a SAM are easily interpreted. In Table 2.1 the row entries for the commodity accounts are the values of commodity sales to the agents identified in the columns, i.e., intermediate inputs are purchased by activities (industries etc.,), final consumption is provided by households, the government and investment demand and export demand is provided by the all the other regions in the global SAM and the export of margin services. All transactions recorded in the commodity account rows are valued at purchaser prices. The commodity column entries deal with the supply side, i.e., they identify the accounts from which commodities are purchased so to satisfy demand. Specifically, commodities can be purchased from either domestic activities – the domestic supply matrix valued at basic prices, i.e., exclusive of domestic trade and transport margins and non-rebated domestic commodity taxes – or they can be imported – valued at basic prices when they are valued inclusive of international trade and transport margins, i.e., *cif*, and import duties. In addition to payments to the producing agents – domestic or foreign – the commodity accounts need to make expenditures with respect to the trade and transport services needed to import the commodities and any commodity specific taxes, in which case they are valued a purchaser prices.

Table 2.1 Macro Social Accounting Matrix for a Region in the Global Macro Social Accounting Matrix

	Commodities	Activities	Factors	Households	Government	Capital	Margins	Rest of World	Totals
Commodities	(Domestic Margin Services)[5]	Combined Intermediate Use Matrix	0	Private Consumption	Government Consumption	Investment Consumption	Exports of Margins (*fob*)	Exports of Commodities (*fob*)	Total Demand for Commodities
Activities	Domestic Supply Matrix	0	0	0	0	0	0	0	Total Domestic Supply by Activity
Factors	0	Expenditure on Primary Inputs	0	0	0	0	0	Factor incomes from other regions	Total Factor Income
Households	0	0	Distribution of Factor Incomes	0	0	0	0	Remittance income	Total Household Income
Government	Taxes on Commodities	Taxes on Production Taxes on Factor Use	Direct/Income Taxes	Direct/Income Taxes	0	0	0	Aid incomes	Total Government Income
Capital	0	0	Depreciation Allowances	Household Savings	Government Savings	0	Balance on Margins Trade	'Current account' balance	Total Savings
Margins	Imports of Trade and Transport Margins	0	0	0	0	0	0	0	Total Income from Margin Imports
Rest of World	Imports of Commodities (*fob*)	0	Factor payments to other regions	Remittance expenditures	Aid expenditures	0	0	0	Total Income from Imports
Totals	Total Supply of Commodities	Total Expenditure on Inputs by Activities	Total Factor Expenditure	Total Household Expenditure	Total Government Expenditure	Total Investment	Total Expenditure on Margin Exports	Total Expenditure on Exports	

[5] When 'domestic margin services' are recorded in the commodity:commodity sub matrix they sum to zero because the supplies are recorded as positive transactions and the demands as negative transactions, and by definition supply equals demand. For these transactions basic and purchaser prices are equal.

The R23 database provides complete coverage of bilateral transactions in commodities – these are valued free on board (*fob*) - but only provides partial coverage of transactions in trade and transport margins. Specifically, the imports of trade and transport margins by each region are associated directly with the imports of specific commodities, hence for commodity imports valued *fob* the source and destination regions are identified and the value of trade and transport (margin) services used are identified. The sum of the values of trade and transport services and the *fob* value of the commodity imports represent the carriage insurance and freight (*cif*) paid value of each imported commodity. But the source regions of the trade and transport services are NOT identified, and similarly the values of exports of trade and transport services by a region do NOT identify the destination regions. To overcome this lack of information the artificial region Globe is included in the database. This region collects together all the exports of trade and transport services by other regions as its imports and then exports these to other regions to satisfy their demand for the use of trade and transport services associated with commodity imports. By construction the value of imports by Globe for trade and transport margin services must exactly equal the value of exports for the corresponding trade and transport service. However, this does not mean that the trade balance between Globe and each and every region must exactly balance, rather it requires that the sum of Globe's trade balances with other regions is exactly equal to zero. Imbalances between Globe and a particular region are recorded in the capital account for that region, indicating borrowing or lending associated with trade and transport margins.

The augmented data are a mix of bilateral transactions and with a partial coverage of transactions. The remittance data are fully bilateral; each transaction records source and destination regions and values (in US dollars). The data used in the R23 database treats remittances as payments from the household in one region to the household in the destination region. The data on payments for and receipts from factor services sold to other regions is also fully bilateral and recorded as payments by each factor type in one region to the matching factor type in the recipient region. The data on bilateral aid transfers by the Donor Assistance Committee members to recipient regions are also fully bilateral bilateral. However recipient regions also receive aid transfers from multilateral/international institutions for which the ultimate source region is unidentified. To overcome this lack of information the Globe region is used as a 'clearing' house: multilateral aid is paid by Globe to the recipient region and the aid budget for Globe is financed by transfers to Globe form other regions. Similarly, the data

on payments for and receipts of other grants does not explicitly identify the source and destination regions: again the Globe region is used as a 'clearing' house.

An important feature of the construction of a SAM can be deduced from the nature of the entries in the commodity account columns. By definition the column and row totals must equate and these transaction totals can be expressed as an implicit price times a quantity, and the quantity of a commodity supplied must be identical to the quantity of a commodity demanded. The column entries represent the expenditures incurred in order to supply a commodity to the economy and hence the implicit price must be exactly equal to the average cost incurred to supply a commodity. Moreover, since the row and column totals equate and the quantity represented by each corresponding entry must be same for the row and column total the implicit price for the row total must be identical to average cost incurred to supply the commodity. Hence the column entries identify the components that enter into the formation of the implicit prices in the rows, and therefore identify the price formation process for each price in the system. Typically, a SAM is defined such that the commodities in the rows are homogenous and that all agents purchase a commodity at the same price.

Total income to the activity account is identified by the row entry. In the simple representation of production in the database the activity in each region makes a single commodity and given differences in cost structures the commodities in each region are heterogeneous. The expenditures on inputs used in production are recorded in the activity columns. Activities use intermediate inputs, which in this version of the database are record as composites of domestically produced and imported commodities, primary inputs and pay taxes on production and factor use. For each region the sum of the payments to primary inputs and on production and factor use taxes by activity is equal to the activity's contribution to the value added definition of GDP while the sum over activities equals the region's value added measure of GDP.

The remaining accounts relate to the institutions in the SAM. All factor incomes are distributed to the single private household after making allowance for depreciation of physical capital, the payment of direct (income) taxes on factor incomes and payments for foreign owned factors. Incomes from factor sales, domestic and foreign, and remittances are the sources of income to the household account. Four categories of expenditures by the household account are recorded; direct (income) taxes, savings, remittances and consumption. The government receives incomes from commodity taxes, production taxes and direct taxes on

factor and household incomes, and aid transfers and uses that income to pay for consumption, aid transfers and for savings. In the R23 database government savings are set equal to the balance on the government's current account, i.e., the internal balance. There are therefore five sources of savings in each region: depreciation, household/private savings, government savings, balances on trade in margin services and balances on the 'current' account, but only a single expenditure activity – investment (commodity) demand.

As should be apparent from the description of the SAM for a representative region the database is strong on inter regional transactions but relatively parsimonious on intra-regional transactions.

3. Estimation Procedure

> "The issue of whether the SAM is deterministic or stochastic is crucial as the SAM provides the underlying data set upon which simple SAM-multiplier analyses and more complex Computable General Equilibrium Models (CGEs) are calibrated. Increasingly, these models are used to explore and simulate the impact of policies and exogenous shocks on the whole socio-economic system. An erroneous or inaccurate SAM invalidates the results obtained from these models" (Thorbecke, 2003, p 186)

At the heart of all quantitative analyses of economic systems, be it a modern macroeconomic model and/or some other form of whole economy model, will be found **estimates** of national accounts. Indeed, so central are such national accounts to the work of economists it is easy to forget how short is the history of (formal/institutionalised) national accounting, especially since the 'wealth of a nation' appears to be a concept that has lain at the very root of economic analyses for more than two centuries (Stone, 1977, provides a brief historical review). One of the enduring legacies of Richard Stone's contribution to economics is the fact that the use of national accounts for economic analyses was central to the conventions for the compiling of national accounts. This has meant that national accounts, if compiled in line with SNA guidelines, adopt definitions and conventions that ensure they can be used meaningfully as a basis for economic analyses and not solely as a mechanical accounting exercise that describes an economy at a particular point in time. Nevertheless, the process of compiling national accounts, of which the most general form of disaggregated national accounts is a Social Accounting Matrix (SAM), remains an estimation process that must in some way or another address imperfections and inconsistencies in the data available to statisticians.

It is important to recognise that the process of constructing a SAM requires the reconciliation of data that are subject to both sampling and measurement errors. The early compilers of SAMs adopted strategies that involved confronting data from different sources with each other; taking a subjective view on the reliability of the different sources and then attempting to satisfy the accounting constraints of a SAM (see Pyatt *et al*., 1977). Stone (1977) responded to this laudable but "laborious method" by asking whether "still better results could not be obtained by applying a formal, mathematical treatment rather than *ad hoc* manipulations to our subjective assessment of reliability" (p xxi). There appears to have been a response to this 'call' in that there are now many seemingly different techniques available for 'balancing' SAMs. More recently, Thorbecke (2003) has argued, correctly, that while

completed SAMs are deterministic, in the sense that each cell has a unique value, it is important to recognise that the process of constructing a SAM still involves the reconciliation of data that are subject to both sampling and measurement error. In essence Thorbecke is moving the debate about estimation strategy on beyond the mere development of techniques by arguing that it is not enough for the techniques to provide a 'mathematical' solution, but rather they must also incorporate recognition that each cell in a SAM is "an estimate arrived at on the basis of data containing sampling and measurement errors" (Thorbecke, 2003, p 185), i.e., they must also provide a statistical solution.

The method for estimating the R23 SAM addresses Erik Thorbecke's challenge by adopting procedures that explicitly define the process as involving estimation where there are errors in the underlying data. It is argued that while RAS based methods, and other similar mathematical methods, achieve the objective of balancing a SAM they satisfy neither Stone's nor Thorbecke's challenges. Further it is argued that while the Stone-Byron method (see Stone (1974); Byron (1976 and 1996) and, for a recent summary, Round (2003)) is a major advance on RAS based methods, in that subjective judgements enter "at a second-order rather than first-order level" (Round, p 177, 2003), it also fails to fully satisfy Stone's stated objective, because, as Stone recognised at the time (p xxii), the required variance matrix "can only be based on subjective impressions of the investigator".

3.1 Estimation Principles

3.1.1 The SAM Estimation Problem

The estimation of a SAM requires the identification of an efficient way to incorporate and reconcile information from a number of different sources that may or may not have been originally collected for purposes of compiling national accounts. In essence the cells of a SAM are unknown parameters whose values must be estimated from observed data; hence the process of compiling a SAM can be classified as an estimation problem. But this is generally an ill posed estimation problem since there are typically more cells/parameters to estimate than available data, which means there are typically negative degrees of freedom and consequently conventional statistical/econometric methods are not strictly appropriate.

Information theory provides one means of addressing the problem of parameter estimation as opposed to prediction. The consequent estimation principles can be defined as:

1. use all the information available; and

2. do not use, or make assumptions about, information that are not available.

Hence, it is not appropriate to make assumptions about either the error generating process or error distribution, e.g., the variance matrix of the Stone-Byron method. Moreover, information theory provides a theoretical framework within which parameters can be estimated when data are scarce and/or incomplete. This accords with Zellner's 'efficient information processing rule' and has close links with Bayesian estimation.

This exemplifies why the terms updating and balancing are arguably inappropriate in the context of mechanical methods used to generate new SAMs. All too often updating has referred to the derivation of a SAM for a later period primarily based upon new estimates of the total incomes/expenditures for accounts and the previous transaction data, while balancing has typically referred to the removal of inconsistencies that mean the calculated row and column totals differ from exogenously known totals. All known mechanical methods involve the use of variants of updating and balancing approaches in that they require the imposition of exogenous assumptions so as to render the problem solvable, e.g., the biproportionality assumption that underpins ALL variants of the RAS method.[6]

Ultimately the SAM estimation problem can be regarded as constituting two related sub problems; the accounting problem, i.e., how to deal with the accounting issues, and the economic problem, i.e., how to ensure that the solution(s) to the accounting problem does/do not undermine the economic content of the system

3.1.2 The Accounting Problem

The essence of the accounting problem is how to reconcile data from different sources. Unless the entire data gathering process for disaggregated national accounts is integrated this problem cannot be avoided since it will be necessary to use data gathered for different purposes. At its simplest the reconciliation process would involve deriving concordances between data collected using different classification schemes, e.g., trade transactions classified using Harmonised System (HS) commodity codes, production data where commodities are classified using a Standard Industrial Classification (SIC) system and household expenditure data where commodities are classified to reflect consumption patterns. To a great extent these difficulties could be partially resolved by ensuring that the different

[6] See Bacharach (1970) on the properties of the RAS method and Lynch (1979) on the limitations of the RAS, and implicitly other, methods for updating matrices. Some early approaches to the problems are discussed in Allen and Lecomber (1975) and Lecomber (1975).

surveys used a common commodity classification scheme for which requisite concordances were defined as part of the data gathering process. But this is unlikely to address all the key problems, since classification schemes often need to satisfy different criteria that may not always coincide, e.g., HS codes need to meet internationally defined criteria whereas an SIC needs to reflect the structures of a national economy.

Even if the problems presented by differences in classification schemes can be resolved this does not solve the reconciliation problem. The sources of data are typically censuses and surveys, and such data raise a series of related difficulties. Surveys face problems associated with the definition of the sample frame, which means that they may not always be perfect representations of populations, while censuses may not be complete. In addition, both surveys and censuses may fail to fully record all transactions, e.g., consumers typically understate expenditures on tobacco and alcohol. But, each transaction is simultaneously an expenditure by one agent and income to another agent, hence it may be the case that there are substantive differences in the recorded values of transactions by sellers and purchasers, e.g., beverage and tobacco activities will typically record higher sales values than expenditures reported by consumers.

Consequently, it is inevitable that there will be errors in measurement and a fundamental aspect of reconciliation is to address the problem of measurement error, which is not a problem of mathematics.

This highlights an important point about the method used to develop the R23 database. The necessity to confront data from different sources and make judgements about their reliability is not avoided. All the available information needs to be challenged and no information should be regarded as sacrosanct; even if this means that the resultant estimates raise doubts about published official data. This task may be laborious but it remains essential. Even the development of the most sophisticated estimation techniques does not alter the requirement for data gatherers to critically evaluate the reliability of conflicting data and to consider how different data sources should be used in the process of compiling the prior estimate of the SAM. Inevitably the judgements entered into in this process risk being subjective, but whereas the pioneers were often required to make firm decisions about the value of the transaction/cell and the Stone-Byron method required the determination of variance and the initial value, this method requires the determination of an initial estimate for the transaction and error bounds.

A critical consideration is the definition of information; in particular, do any macroeconomic totals that may be available constitute information when compiling disaggregated national accounts. It is arguable that a theoretical ideal is that estimates of macroeconomic totals should be derived from micro level data, e.g., estimates of private consumption should be based on survey evidence and population estimates. Such an approach is arguably consistent with the principles of the SNA and the concept of using supply and use data to 'benchmark' national accounts. In the context of a national statistical agency this approach is eminently sensible. But for non-government compilers of SAMs it may not be practical to follow this theoretical ideal, rather such compilers may need to adopt a more pragmatic approach. Since the databases compiled by non-government agencies will rarely if ever influence the published estimates of aggregated national accounts, it is often appropriate to treat the main macroeconomic totals as binding constraints such that the disaggregated accounts are consistent with the published national accounts. At first sight the requirement of consistency with exogenous macroeconomic total may seem to make the process easier, but whereas the bottom up approach places its emphasis upon reconciling micro level data, with macroeconomic totals then being defined as quasi deterministic aggregates, a top down approach requires that the micro level data and the macroeconomic total must be reconciled, i.e., an additional set of constraints must be satisfied.

On advantage of this approach is that it can use as many or as few (macroeconomic) aggregates as are available, or the compiler wishes to use. This facilitates the use of the estimation method by a wide range of agencies that may or may not be acting with full access to base data

3.1.3 The Economic Model Problem

In addition to addressing the accounting problem it is necessary to ensure that the economic problem is incorporated within the estimation process. This requires that the resultant SAM must be fully consistent with economic logic. There are six key aspects of the SAM that need to be recognised

1. there cannot be negative factor demands or negative final demand from households or government, or negative intermediate demand by activities, i.e., negative demands are in fact supplies;
2. the column coefficients must be logically consistent, i.e., the cost shares in production must reflected the costs in the economy, e.g., the share of intermediate

inputs in gross output should be consistent with that found in the underlying data, and the distribution of the product ('value added') between factors should accurately reflect the underlying data;

3. the institutional accounts should reflect accurately the flows of incomes and expenditures between domestic and foreign institutions;
4. sources of savings should be fully articulated, e.g., economic not accounting, depreciation should be included[7];
5. the tax system should be well articulated and the revenues, and hence the implied (not published) rates, should be accurately estimated since they are key determinants of the price formation process, e.g., basic to purchaser prices, and key policy instruments; and
6. similarly, trade and transport costs (domestic and international), and hence implied unit costs, should be accurately estimated since they are key determinants of the price formation process, e.g., basic to purchaser prices.

While the economic problem is relatively easily stated the evaluation of the economic content of a SAM is much more difficult, since to some extent the process is subjective and depends on the knowledge of the person compiling the raw data. More objective methods can be achieved by using models. Price multiplier models can be used to determine the extent of implied differences in price formation processes, e.g., typically developing countries will be less 'roundabout' in their production technologies. Simulations using a CGE model can oftentimes produce results that indicate peculiarities in the model's database.

3.2 Estimation Stages

The estimation of the global R23 SAM was conducted as a sequential exercise that is a hybrid of top-down and bottom-up processes, which emphasises the importance of the core components of (expenditure side) GDP while facilitating the detail of inter-regional transactions. A major problem is that the SAMs for each region must be complete and consistent, i.e., fully reconciled, with respect to the data available for each region, AND the global SAM must also, and simultaneously, be complete and consistent, i.e., transactions between regions must also be fully reconciled with both the intra-regional and inter-regional

[7] In the SNA depreciation is known as 'consumption of fixed capital' to emphasise the difference between economic and accounting depreciation.

data. The sequential approach is a pragmatic solution to the problem of simultaneously estimating and reconciling the macro SAMs for each region and all bilateral transactions.

The estimation process can be categorised as a three-stage process. First, the estimation of macro SAMs for each region; second, the estimation of bilateral transaction matrices that detail the inter-regional transactions; and third, the estimation of the global R23 SAM. The description of the procedure proceeds as if estimation was sequential, however the estimation process was iterative both by design and necessity[8].

The estimation algorithms are adaptions of a stochastic entropy metric, see Golan *et al*., (1994); Robinson *et al*., (1998); Robinson and McDonald (2004); PROVIDE (2006).

3.2.1 Single Region Macroeconomic SAMs

This stage involved the collection of data required to estimate simple prior macroeconomic SAMs for each region consistent with the SAM structure outlined in Table 2.1 (above), where the Rest of the World account was a single row and column. IF the source data national accounts for a country have been fully reconciled the macroeconomic SAM should be easily derived and be complete and consistent. However, the prior macro SAMs reported substantial differences in row and column totals, which demonstrated the extent to which the source data were not reconciled.[9]

This was anticipated from the experience of compiling macro SAMs for the 123 model using World Bank data. For the 123 SAMs it was not uncommon to resolve the problem by using specific transactions as balancing items.[10] But for purposes of the R23 SAM the requirement to reconcile transaction both within and between regions necessitated avoiding using specific transactions as balancing items because this would create problems when reconciling the global R23 SAM.[11]

---

8 The extent to which both the raw intra-regional and inter-regional data were inconsistent made the estimation process far more difficult than anticipated.

9 Note that the intermediate transactions data could be netted out at this stage.

10 It is not uncommon in SAMs used to calibrate comparative static CGE models to find the savings/investment account used to achieve reconciliation, e.g., through the internal balance and/or stock changes.

11 It is trivial to demonstrate that distorted estimates of a transaction in a SAM will have implications for other transactions in a SAM. The art of using specific transactions as balancing items is choosing transactions that minimise distortions. In a disaggregate SAM this arguably best achieved through the savings/investment account but in a macro SAM using any transaction introduces potentially large errors.

The SAM entropy estimation metric was used to produce the estimated macroeconomic SAMs. The information contents of these SAM were evaluated and, where there was reason to doubt the estimates of transactions in the prior macro SAMs, additional information was sought. For some countries the data were deemed too poor to retain the country as a separate region within the global SAM.

3.2.2 'Bilateral' Transaction Matrices

The next stage was to estimate prior 'bilateral' current account transaction matrices for trade in goods and services, remittances, bilateral aid transfer and factor payments to foreign owned assets. For some inter-regional transactions, it was not possible to derive estimates of the bilateral transactions because of gaps in the source data. However, the available data demonstrated that bilateral transactions data were incomplete, and hence it was necessary to include estimates of the missing data. The solution chosen was to include the Globe region within the 'bilateral' transaction matrices. Where inter-regional transactions on the current account were recorded as incomes and expenditures by countries without defined sources and destination these were recorded as payments from or to the Globe region.

The SAM entropy estimation metric was used to produce the estimated bilateral transaction matrices subject to constraints derived from the estimated aggregate totals of transactions by each country with the rest of the world accounts in the macro SAMs. Again the information contents of these bilateral transaction matrices were evaluated and, where there was reason to doubt the estimates of transactions in the prior bilateral transaction, additional information was sought. For some countries the data were deemed too poor to retain the country as a separate region within the global SAM.

5.2.3 Global Macroeconomic (R23) SAM

The total imports and exports to and from each country recorded in the bilateral transaction matrices were then used to revise the prior macro SAMs for each country and the macro SAMs were then re-estimated. The combination of the macro SAMs and the bilateral transaction matrices produced the global R23 SAM.

However, the process did require iteration. Typically, the estimation of the bilateral transactions matrices indicated that some of the estimated aggregate totals of transactions by each country with the rest of the world accounts in the macro SAMs were unreliable. This

required iteratively re-estimating the macro SAMs and the bilateral transactions matrices until the resultant estimates were deemed reliable.

3.3 Problems Posed by Ill-defined Priors

A particular issue is how to address problems posed by ill-defined priors, where, typically, an ill-defined prior is a consequence of unreliable or missing data/information.

The absence of prior estimates for transactions that economic logic indicates must occur, e.g., direct/income tax revenue, total imports etc., is an obvious problem. It is self-evidently not appropriate to assume the transactions were zero because such an assumption will introduce substantial biases in the prior and all known mathematical estimation techniques require a non-zero prior for a transaction to appear in the final matrix. In such cases the solution is relatively simple; a non-zero estimate with large error bounds is preferable to a zero estimate without error.

Unreliable prior estimates in fact present a more substantial set of problems. If prior estimates of transactions are judge to be unreliable then the logic of estimation techniques dictates that large error bounds should be attached to the priors. If the entropy metric returns a large marginal value for the transaction in the estimated macro SAM, then the statistician is directed to seek a better prior whereas is the marginal value is low so are the incentives to seek a better prior. But, mathematical estimation techniques inevitably result in larger absolute changes in large transactions and visa-versa. Unfortunately, this can produce serious estimation problems where the range of transaction values is large; the problems can be particularly severe for relatively small transactions values. In the context of the R23 SAM this means that the estimated SAMs for small countries can report large differences relative to the priors and small transactions, often tax revenues, in large countries can report large differences relative to the priors.

The resolution of these problems depends critically on the judgement of the statistician compiling the prior transaction estimates and error bounds.

4. Prior Matrices

In this section we outline the sources used for constructing the priors for the social accounting matrix. Priors are need for a number of matrices:

1. Macro SAMs for each of the countries, including the rest of the world.
2. A bilateral matrix of trade flows matrix: consisting of all trade flows between each country, as well as total import margins by country and supply of margin services to Globe by country.
3. A bilateral matrix of remittance flows: for all remittance flows between households from each of the countries.
4. A bilateral matrix of worker's compensation flows: for all payments made to workers from each of the countries working in one of the other countries.
5. A bilateral matrix of foreign income flows: for all payments made to capital from each of the countries but owned by residents of one of the other countries.
6. A bilateral matrix of aid flows: for all aid payments made to governments in one of the countries from governments in one of the other countries, as well as aid payments made through multilateral organizations (Globe).

4.1 Macroeconomic Data

The macroeconomic data used to provide the initial elements of the Macro SAM come from various World Bank, IMF and OECD datasets, as well as the GTAP 8 database (Narayanan *et al*. 2012). Specifically,

- The World Bank's World development indicators (WDI) database (http://data.worldbank.org/) is used to provide data on GDP by expenditure, gross value added and gross domestic savings, as well as supplement the international transactions and government data, when not available in the IMF datasets. A list of the codes and definitions of the datasets used from the WDI database is given in Table A1-1 in the appendix. WDI data for 2007 were available for up to 209 countries.
- The IMF's Balance of Payments Statistics (BOPS) database (http://www.imf.org/en/Data#data) is used to provide data on all transactions (trade and income flows) with the rest of the world. A list of the data provided

in the IMF BOPS (1993) is given in Table A1-2 in the appendix. Data for 2007 was available for up to 180 countries.

- The IMF's Government Financial Statistics (GFS) database (http://www.imf.org/en/Data#data) is used to provide data on government income, including taxes, and expenditures. A list of the data provided in the IMF GFS (2014) and their codes is given in Table A1-3 in the appendix. Data were available for up to 131 countries.
- The OECD's Development Assistance Committee (DAC) database containing data on aid payments made by DAC members to other countries either directly or through multilateral institutions. There are data for 25 DAC member countries by 142 aid recipient countries.
- The GTAP 8 database (Narayanan *et al*. 2012) for 2007, along with the underlying macroeconomic aggregates and trade data for all 244 countries underlying the GTAP database and supplied by the Centre for Global Trade Analysis. These were used to provide information for intermediate inputs, depreciation, trade, margins, factors splits, as well as providing a last-resort source of information for priors when IMF or World Bank data were unavailable.[12] Since the GTAP 8 database includes only 134 countries, the database was first aggregated to one commodity and then expanded from 134 to 244 countries using the original mapping and GDP shares used in the original construction of the GTAP database. The resulting macro SAMs of countries not included in the GTAP database therefore reflect the macro SAMs of their composite region in GTAP weighted by GDP with additional data for trade and macroeconomic aggregates replaced by the disaggregated GTAP data supplied by the Centre for Global Trade Analysis. These data were only used in the prior for the balancing program when no other macroeconomic data were available from any of the other sources.

[12] GTAP data also provided many of the final factor splits, which is discussed latter in this documentation.

Table 4.1: Sources used to create initial Macro SAMs

	Commodities	activities	Labour	Other factors	Commodity taxes (import, export VAT and sales*	Factor taxes (labour and other factor)*	Production taxes	direct taxes	households	government	Grants	KAP	ROW
Commodities	-	GTAP	-	-	-	-	-	-	WDI NE_CON_PRVT_CN	WDI NE_CON_GOVT_CN	-	WDI NE_GDI_TOTL_CN	BOPS credit, goods and services
activities	Residual	-	-	-	-	-	-	-	-	-	-	-	-
Labour	-	WDI NY_GDP_FCST_CN	-	-	-	-	-	-	-	-	-	-	BOPS credit, income, compensation of employees
Other factors	-		-	-	-	-	-	-	-	-	-	-	BOPS credit, income, investment income
Import tax	GFS 1151 + % of 1153	-	-	-	-	-	-	-	-	-	-	-	-
Export tax	GFS 1152 + % of 1153	-	-	-	-	-	-	-	-	-	-	-	-
VAT	GFS 11411	-	-	-	-	-	-	-	-	-	-	-	-
Sales tax	GFS 114exVAT	-	-	-	-	-	-	-	-	-	-	-	-
factor use tax labour	-	GFS 112	-	-	-	-	-	-	-	-	-	-	-
factor use tax other factors	-	GFS 113	-	-	-	-	-	-	-	-	-	-	-
Production taxes	-	GFS 116 less 25	-	-	-	-	-	-	-	-	-	-	-
direct taxes	-	-	-	-	-	-	-	-	GFS 111	-	-	-	-

	Commodities	**activities**	**Labour**	**Other factors**	**Commodity taxes (import, export VAT and sales***	**Factor taxes (labour and other factor)***	**Production taxes**	**direct taxes**	**households**	**government**	**Grants**	**KAP**	**ROW**
households	-	-	Residual	Residual	-	-	-	-	-	GFS 24+27+28-12-14	-	-	BOPS credit, Income, Current Transfers, Other Sectors, Workers' Remittances
government	-	-	-	-	total of each tax row				-	-	-	-	-
Grants	-	-	-	-	-	-	-	-		Residual	-	-	BOPS credit Income, Current Transfers, General Government and Income, Current Transfers, Other Sectors, Other Transfers
KAP	-	-	-	GTAP	-	-	-	-	WDI NY_GDS_TOTL_CN (less KAP,GOVT and KAP, fofac)	GFS 1 less 2 or residual	-	-	Residual
ROW	BOPS debit goods and services	-	BOPS debit, income, compensat-ion of employees	BOPS debit, income, investment income	-	-	-	-	BOPS debit, Income, Current Transfers, Other Sectors, Workers' Remittances	-	BOPS debit Income, Current Transfers, General Government and Current Transfers, Other Sectors, Other Transfers	-	-

* Tax columns aggregated to improve illustration.
See Appendix Tables A1-1, A1-2 and A1-3 for code definitions

Table 4.1 provides an overview of the primary or preferred data source used for each cell. While certain sources are preferred, the sources were not always complete and other data was therefore drawn upon to supplement. Aggregates, such as GDP, absorption, total tax revenues, total government expenditure etc. were also collected to provide macro totals for the balancing program and to check the reliability of underlying data.

In addition to the preferences stipulated in Table 1, data in US dollars were also preferred over local currency units (LCU). Where data in US dollars are not available conversion factors, based on the World Bank GDP data in local currency units relative to US dollars are used to convert to US dollars. We also examined alternative conversion factors – based on absorption and final expenditure – and found these to be equivalent.

Data on exports and imports were available from numerous sources: the IMF BOPS; the World Bank's recordings of IMF BOPS data, which were not always the same as the IMFs data; the World Bank's national accounts numbers, and the GTAP trade data developed by Gehlhar, *et al*. (2008). In general, the ratio of these numbers were close to one. Any differences were generally because of large discrepancies between the World Bank's national account data and the IMF's BOPS data. In those countries where large differences existed, the World Bank's balance of payments data were either closer to the IMF BOPS data or the World Bank national account numbers, but neither source stood out consistently despite the fact that the World Bank state that the data come from the IMF. It is difficult to tell if these differences are the result of lags in updating numbers between the two institutions or differences in opinion about the numbers. In general, we selected the IMF BOPS data first, as we believed this was the original and hence most up-to-date source, followed by the World Bank's balance of payments data, which generally included data for some extra countries. That said, where large discrepancies between the IMF and World Bank numbers existed, we chose to more closely match the GTAP trade data, which was built up from the underlying bilateral trade data.

Table A1-3 in the appendix shows how World Bank data were used to supplement the IMF GFS data, when data were not available in the IMF GFS data for all government incomes and expenditures. When information was not available in either of these datasets, the GTAP database was used to fill in missing tax incomes. On the other hand, the World Bank data were only used to supplement foreign income flows in the IMF BOPS data in a few cases,

where the expected relationship between the World Bank and IMF GFS data could be confirmed in the data. The explicit separation of remittances and worker's compensation in our macro SAM as incomes to households and labour respectively, was more consistent with the definitions used in the IMF BOPS data.

Data on grants/transfers between the government and the rest of world can be obtained from either the IMF GFS accounts (grants to/from foreign governments and international organizations) or from the IMF BOPS accounts (general government transfers and other transfers), as well as from the World Bank WDI database and the OECD DAC database (specific Development Assistance Committee (DAC) grants). Examination of the datasets however revealed that neither data source was consistently good for obtaining data on all recipients and senders of grants/aid, with inconsistent reporting across the sources by countries, leading to total inflows and outflows from each source failing to match globally. After some examination it was found that when the maximum values of grants in and out across sources were used, the data were globally consistent. Hence maximum values were used.

4.2 Trade Data

The initial bilateral trade matrix for 244 countries was obtained from the Centre for Global Trade Analysis and is the basis for the GTAP 8 database. This bilateral trade matrix was developed by Gehlhar *et al*. (2008). Both *f.o.b.* and *c.i.f.* values were obtained for an aggregate commodity (including both goods and services).

Since the R23 database has many more countries than the GTAP database extra attention was paid to the additional countries not in the GTAP 8 Database, but included in the 244 bilateral trade matrix. In particular the supply and demand of international transport margins was missing for many of these bilateral flows. Missing margins were filled assuming average margin rates for trade between countries within geographical regions. Hence, if margin rates were unknown for trade between Peru and Zimbabwe then Peru's average margin rates on goods to/from Africa was used; or Zimbabwe's average margin rate on goods to/from South America; or in the worst case South America's average margin rate to/from Africa. Using the initial trade data an estimated total values of margins (demand) were derived. The GTAP database provided the initial values of the supplies of those margin services.

4.3 Transfers Data

4.3.1 Aid and Other Grants

The bilateral DAC database developed by the OECD must first be adjusted to remove any negatives and ensure that aid from multilateral organizations is supplied by donor countries. The negatives in the 2007 OECD DAC database were removed and replaced with a 3-year averages and an adjustment made to ensure balance. Supplies by donors to the multilateral organizations are assumed to come from the DAC donor countries in the same shares with which they supply aid to DAC countries.

Once adjusted and balanced, DAC aid data are subtracted from total grants and included separately. The bilateral DAC database is used to provide both the source and destination of the DAC aid, with multilateral aid flowing through the globe account. The remaining non-DAC grants are also assumed to go through Globe and are therefore not bilateral.

4.3.3 Remittances

The initial estimates for determining the source and destination of remittance data comes from data developed as part of the GMig2 database (Walmsley *et al*. 2005) which estimates the numbers of workers, wages and remittances by home and host country of migrant labour.

4.3.4 Factor Payments

The GMig2 bilateral remittance data (Walmsley *et al* 2005) is also used as the initial estimates for worker's compensation by source and destination. Finally, the bilateral FDI database developed by Guimbard, Gouel and Laborde from CEPII and adapted and documented in Lakatos and Walmsley (2010) was used to provide the initial estimates of factor payments to capital by source and destination.

5. Practical Estimation

5.1 Individual Macroeconomic SAMs

With data collected and the priors compiled from the various sources, a number of countries were removed or aggregated due to insufficient data leaving 203 countries plus the rest of world. A country was removed if a figure for GDP could not be obtained or there were no numbers in the macro SAM supplied by World Bank and IMF sources. A few larger countries were kept despite having insufficient data provided in the macro SAMs derived from World Bank or IMF data. In these cases, GDP figures were obtained from GTAP.[13] The share of cells containing World Bank or IMF data for each country are given in Table A1-4 of the Appendix. The countries removed were either dependencies of larger countries, in which case they were aggregated with the country of sovereignty, of they were aggregated into a rest of world. Table A1-4 in the appendix provides the final list of countries in the R23 database.

Before proceeding to the balancing of each of the individual macro SAMs adjustments were made to the international transactions to ensure global balance across all countries. Exports, imports, remittances, workers' compensation, foreign capital flows and aid by country were all adjusted so that total global inflows were equal to global outflows. These adjustments to balance the data were applied equally to both inflows and outflows and resulted in changes of 1 to 3 percent. The decision to balance the trade and current account income flows globally, before balancing the macro SAMs was taken after testing both methods. We found that if each country's Macro SAM was balanced first, without considering global balance of the trade and other current account flows, then the resulting trade and current account flows were so far from global balance that re-balancing the macro SAMs led to even larger changes in other elements in the Macro SAMs to restore balance. Hence ensuring global balance first and reducing the standard errors applied on these international transactions assisted in reducing the number of stages in the overall R23 construction process.[14]

[13] GTAP generally use alternative sources such as the CIA worldfact book to find GDP and population figures for missing countries (Hussein and Aguiar, 2012).

[14] Note that in a handful of countries these adjusted macro SAMs did not balance and hence the initial macro SAMs were used.

Each element of the Macro SAM was then given a standard error reflecting our belief in the underlying data. Data obtained from the World Bank or IMF were given the lowest standard error and those filled with GTAP data were given larger standard errors. In order to reduce issues later with the balancing of the international datasets, the standard errors on international transactions were also set very low. Other rules applied, raised the standard error on household savings and government transfers, relative to tax revenues. Household and government expenditure and investment were also given lower standard errors; and real GDP was targeted.

The Macro SAMs were then balanced individually. Following this, the international transactions were again checked for balance and any difference was applied to the rest of the world. The rest of the world was then rebalanced assuming the differences between inflows and outflows of each of the international transactions were fixed. The resulting 204 macro SAMs were both internally consistent and all international transactions balanced globally.

Finally, some minor adjustments were made to smooth out changes. First, private consumption, government consumption and investment were assumed to be move together so as to maintain their share in absorption. Second, in some cases the balancing process would place all of the change on one particular tax (e.g., export taxes) rather than spread the change uniformly across all commodity taxes with the same standard error. Adjustments were made to smooth changes in taxes where they came from the same source and appeared in the same column of the Macro SAM.

5.2 Trade Data

With total exports and imports obtained from the macro SAMs (IMF, World Bank and GTAP data), the bilateral trade matrix was balanced assuming fixed row and column totals. The bilateral trade matrix has dimensions 206 by 206, where the first 204 rows and columns represented trade from country *r* to country *s* at *f.o.b* prices. The 205th row (named *glo*) included total margins on all imports to country *s*; and the 205th column represented the supply of margin services by country *r*. The totals value of margins (the 205th row) was then allocated across each element of bilateral trade based on the margin rate and the total value of margins. The 206th row and column contained the global balances between exports and imports, which were assumed to be fixed in order to match the macro SAM import and export data.

5.3 Transfers Data

Finally, the other international transactions were also included. Aid provided by DAC members was already balanced and was removed from totals grants provided in the balanced macro SAMs. Since grants had been balanced globally the residual was also balanced and were all assumed to be allocated through Globe.

Like the trade data, total remittances in and out were taken from the macro SAMs and the bilateral remittances data were balanced to match fixed row and column totals, assuming fixed difference contained in the "*kap*" row/column. Similarly, bilateral factor payments to capital and labour were obtained in the same way using totals obtained from the macro SAMs and bilateral estimates from the priors discussed above.

5.4 Global (R23) Macroeconomic SAM

Finally, the macro SAMs and the international transactions were brought together to form the final R23 database. GTAP value added shares were used to separate value added into capital, land, unskilled and skilled labour. Workers' compensation was also separated into skilled and unskilled based on value added shares in labour of the host country. All foreign investment incomes were assumed to be related to capital and hence land received a zero share. Finally, factor taxes on labour and capital were also separated into taxes on skilled and unskilled and on land and capital respectively using GTAP shares.

Trade taxes (import and export taxes) were applied across source and destination based on GTAP tax rates applied to the balanced bilateral trade data and subject to the new total tax revenues obtained from the macro SAMs. Globe was also created to deal with margin supply and demand, DAC multilateral grants and other grants.

6. Limitations and Data Constraints

Table A1-5 provides the ratio of row to column totals in the initial Macro SAMs thereby providing an indication of the extent to which the initial World Bank and IMF data are unbalanced.[15] These ratios are calculated before any adjustments are made to ensure the trade and foreign incomes data are globally balanced and before any adjustments made to reflect any large differences between the IMF trade data and the GTAP trade data. The extent to which the IMF and World Bank Macro SAMs do not balance is considerable and of great concern.

The large differences between the row and column totals in the initial macro SAMs in all countries (see Table A1-5) requires large movements to occur in order to balance the macro SAMs. Even in countries where macro SAM data might be expected to balance, large differences were found. Most of the large differences result from inconsistencies between the World Bank and IMF database. For example:

- Government spending recorded by the World Bank for the calculation of real GDP can differ considerably from spending, compensation of employees and consumption of fixed capital[16] recorded in the IMF GFS accounts. For example, the World Bank records government spending in the USA of $2,209.7B, while the IMF records spending, workers' compensation and consumption of fixed capital as $737.6B. The World Bank estimate is larger than the sum of all expenses recorded in the IMF GFS. It is therefore unsurprising that government row/column in the USA Macro SAM shows income (row) as much lower than expenditure (column).
- As discussed above, there are also differences between imports and exports obtained from World Bank's WDI national account numbers verses those obtained from the IMF BOPS and used for current account calculations.
- Differences in grants recorded in IMF GFS accounts compared to those recorded in IMF BOPS accounts.

15 Not all rows and columns are shows since the other row and columns balance, due to the fact that there are residuals in the row or column that are derived so as to match the row and column totals.

16 The term 'consumption of fixed capital' is used in the SNA to distinguish between accounting and economic depreciation. The term 'depreciation' in the text always refers to economic depreciation.

That said, at least some attempt is made by the IMF to ensure that the government (IMF GFS) and current accounts (IMF BOPS) balance individually, even though they may not balance with each other. No attempt appears to be made to check that savings (from domestic and foreign sources) equals investment in the country. It is therefore not surprising that the ratio of the "*kap*" row and column differs from one in all countries (Table A1-5).

These large differences in the unbalanced macro SAMs, particularly in the "*kap*" row/column, and the selection of the small standard errors on absorption and foreign transactions, mean that the resulting changes in elements of the macro SAMs relative to the original World Bank data tend to be very large for household and government savings, depreciation, the current account balance17 and taxes. As mentioned previously, allowing larger changes on foreign transactions caused issues with second round balancing of the macro SAMs once the international transactions were balanced globally.

Table A1-5 in the appendix ranks the countries according to how much each country's Macro SAM changed during the balancing process. The ranking is based on the sum of the absolute differences between initial World Bank and IMF data in the macro SAM and the final value in the macro SAM relative to the country's GDP. Note that this ranking only takes into account cell entries that were initially sourced from World Bank or IMF databases, adjusted to take account of the number of cells sourced from these datasets. Any changes in cell entries obtained from the GTAP data are not considered in the ranking, since many of the countries where GTAP data were used are not explicit in the GTAP 8 database and hence were given values obtained from the relevant "rest of region" to which they were mapped, rather than being data that truly reflected the country data, hence why higher standard errors were applied.

[17] Adjustments made to the international transactions, to balance them globally and the tightening of these standard errors, means that the current account balance is essentially determined as a residual.

7. Concluding Comments

This paper reports on the development of the R23 database and the rationale behind some of the methods used. The database has a companion CGE model, R23 (McDonald *et al*., 2015), that uses all the information in the database to generate globally consistent macroeconomic estimates of responses to policy shocks. The database, model and associated packages are open source[18].

During the process of compiling the database it emerged that a large part of the international macroeconomic data available from the World Bank and the IMF were inconsistent, i.e., unreconciled, at both a national and international level, and that for many countries there were non-trivial gaps in the databases. The data compilation process therefore required substantial amount of estimation in order to reconcile the data. The extent to which the available data were not reconciled indicates that the scope for inaccuracies in the available data is substantial and that therefore the potential for misleading economic analyses in not trivial. Hence, one conclusion from this exercise is that greater attention could/should be devoted to the reconciliation of global macroeconomic data in order that the reliability of economic analyses is enhanced.

The estimation processes were careful to recognise the distinction between the data being **complete** and **consistent** and to ensure that both conditions were, to the extent possible, met. In particular, it was deemed inadequate that the incomes and expenditures to all accounts were equal – often described as balanced – since just because the data were consistent does not mean they are complete. A concern that emerges from this process is the inherent danger of emphasising the balancing/consistency of SAMs while potentially neglecting the completeness of the SAMs; if SAMs are consistent but incomplete the missing transactions require that other transactions in the database must have been distorted to achieve consistency.[19]

The R23 SAM is the first known attempt to develop a globally reconciled set of macroeconomic data together with detailed aggregate inter-regional trade and current account

[18] http://cgemod.org.uk/r23.html

[19] This very important when top-down SAM estimation methods are used; if the underlying macro SAM is incomplete the resultant disaggregated SAM may be seriously distorted.

transactions.[20] The data have been developed to support policy analyses that incorporate the implications of inter-regional transactions and to provide globally consistent database that includes developing countries not included in other databases, e.g., GTAP.

The case study conducted as part of this project demonstrates some of the potential worth of the database and associated model. The R23 database has the potential to enhance other global databases, e.g., GTAP, WIOD, if the data are taken up to enhance the macroeconomic control totals underlying those databases. However, given the extent of estimation required to derive the R23 SAM it is suggested, strongly, that any macroeconomic total from the R23 should be treated are estimates of the true underlying parameters.

[20] While the GTAP database is globally consistent the exclusion of detailed inter-regional current account transactions and limited intra-regional macroeconomic aggregates are limitations.

8. Bibliography

Allen, R.I.G. and Gossling, W.F., (eds) (1975). *Estimating and Projecting Input-Output Coefficients*. London: Input-Output Publishing Company.

Allen, R.I.G. and Lecomber, J.R.C., (1975). 'Some Tests on a Generalised Version of RAS' in Allen, R.I.G. and Gossling, W.F., (eds) *Estimating and Projecting Input-Output Coefficients*. London: Input-Output Publishing Company.

Bacharach, M., (1970). *Biproportional Matrices and Input-Output Change*. Cambridge: Cambridge University Press.

Byron, R.P., (1976). 'The Estimation of Large Social Account Matrices', *Journal of the Royal Statistical Society*, A, Vol 141, pp 359-367.

Byron, R.P., (1996). 'Diagnostic Testing and Sensitivity Analysis in the Construction of Social Accounting Matrices', *Journal of the Royal Statistical Society*, A, Vol 159, pp 133-148.

de Melo, J. and Robinson, S., (1989). 'Product Differentiation and the Treatment of Foreign Trade in Computable General Equilibrium Models of Small Economies', *Journal of International Economics*, Vol 27, pp 47-67.

Devarajan, S., Lewis, J.D. and Robinson, S., (1990). 'Policy Lessons from Trade-Focused, Two-Sector Models', *Journal of Policy Modeling*, Vol 12, pp 625-657.

Drud, A., Grais, W. and Pyatt, G., (1986). 'Macroeconomic Modelling Based on Social-Accounting Principles', *Journal of Policy Modeling*, Vol 8, pp 111-145.

Gehlhar, M., et al. (2008). Reconciling Merchandise Trade Data. Global Trade, Assistance, and Production: *The GTAP 7 Data Base*. B. Narayanan and T. L. Walmsley. West Lafayette, IN, Center for Global Trade Analysis, Purdue University.

Golan, A., Judge, G. and Robinson, S., (1994). 'Recovering Information from Incomplete or Partial Multisectoral Economic Data', *Review of Economics and Statistics*, Vol 76, pp 541-549.

Hussein, Z., and A. Aguiar (2012). "GTAP 8 Data Base Documentation - Chapter 6: Macroeconomic Data." In Narayanan, G. B., et al. (2012). Global Trade, Assistance, and Production: The GTAP 8 Data Base. West Lafayette, Indiana, Center for Global Trade Analysis, Purdue University

International Monetary Fund (1993). *Balance of Payments Statistics Yearbook*, International Monetary Fund. Statistics Department, Washington, D.C., USA.

International Monetary Fund (2014). *Government Finance Statistics Yearbook*, International Monetary Fund. Statistics Department, Washington, D.C., USA.

International Monetary Fund (IMF) (2015): *Balance of Payments Statistics* (BOPS) database (http://www.imf.org/en/Data#data). Washington, D.C., USA.

International Monetary Fund (IMF) (2015): *Government Financial Statistics* (GFS) database (http://www.imf.org/en/Data#data). Washington, D.C., USA.

Lakatos C., and T. L. Walmsley (2010). "A Global Multi-sector Multi-region Foreign Direct Investment Database for GTAP." *GTAP Research Memorandum 18*, West Lafayette, IN: Center for Global Trade Analysis.

Lecomber, J.R.C., (1975). 'A Critique of Methods of Adjusting, Updating and Projecting Matrices' in Allen, R.I.G and Gossling, W.F., (eds) *Estimating and Projecting Input-Output Coefficients*. London: Input-Output Publishing Company.

Lynch, R.G., (1979). 'An Assessment of the RAS Method for Updating Input-Output Tables', in Sohn, I., (ed) *Readings in Input-Output Analysis*. Oxford: Oxford University Press.

McDonald, S. and Sonmez, Y., (2004). 'Augmenting the GTAP Database with Data on Inter-Regional Transactions', Sheffield Economics Research Paper 2004:009. The University of Sheffield

McDonald, S., and Thierfelder, K., (2004a). 'Deriving a Global Social Accounting Matrix from GTAP version 5 Data', GTAP Technical Paper 23. Global Trade Analysis Project: Purdue University.

McDonald, S., Thierfelder, K. and Robinson, S. (2007) 'Globe: A SAM Based Global CGE Model using GTAP Data', Department of Economics Working Paper 14, United States Naval Academy. http://ideas.repec.org/s/usn/usnawp.html and www.cgemod.org.uk/globe1.pdf

McDonald, S. andThierfelder, K., (2013) 'Globe 2: A SAM Based Global CGE Model using GTAP Data', Department of Economics Working Paper 14, United States Naval Academy. www.cgemod.org.uk/globe2.pdf

Narayanan, G. B., et al. (2012). Global Trade, Assistance, and Production: The GTAP 8 Data Base. West Lafayette, Indiana, Center for Global Trade Analysis, Purdue University.

OECD (2015): DAC International Development Statistics, OECD.Stat (http://www.oecd.org/dac/stats/oecdstat-faq.htm). Paris, France.

PROVIDE (2006). 'IterSamBal: A Multistage Matrix Estimation Method and Framework Using Cross Entropy', PROVIDE Project Technical Paper 2006:2. PROVIDE Project: Elsenburg.

Pyatt, G., (1987). 'A SAM Approach to Modelling', *Journal of Policy Modeling*, Vol 10, pp 327-352.

Pyatt, G., (1991). 'Fundamentals of Social Accounting', *Economic Systems Research*, Vol 3, pp 315-341.

Pyatt, G., Roe, A.A. and Associates (1977). *Social Accounting for Development Planning with Special Reference to Sri Lanka*. CUP: Cambridge.

Robinson, S. and McDonald, S. (2004). 'A General Information-Theoretic Approach to Estimating Disaggregated National Accounts, Including Input-output and SAM Accounts', International Conference on Input-Output Techniques, Brussels, Sept. http://cgemod.org.uk/samest.html

Robinson, S., Cattaneo, A., and El-Said, M., (1998). 'Estimating a Social Accounting Matrix Using Cross Entropy Methods', *mimeo*.

Round, J.J., (2003). 'Constructing SAMs for Development Policy Analysis: Lessons Learned and Challenges Ahead', *Economic Systems Research*, 15(2), pp 161-183.

Stone, R. (1962a) *A Computable Model of Economic Growth* (with Alan Brown). No. 1 in A Programme for Growth, Chapman and Hall, London, 1962.

Stone, R. (1962b) *A Social Accounting Matrix for 1960* (with Alan Brown and others). No. 2 in A Programme for Growth, Chapman and Hall, London, 1962.

Stone, R., (1977). 'Forward' in Pyatt, G., Roe, A.R. and Associates, *Social Accounting Matrices for Development Planning with Special Reference to Sri Lanka*. Cambridge: Cambridge University Press.

Stone, R., Bates, J.M. & Bacharach, M., (1963). 'Input-Output Relationships 1954-66', Vol 3 in Stone, R, (ed) *A Programme for Growth*. London: Chapman & Hall.

UN (1968). *A System of National Accounts*. Studies in Methods, Series F, Rev 4. New York: United Nations.

UN (1993). *System of National Accounts*. New York: United Nations

United Nations (2009). *System of National Accounts 2008*. http://unstats.un.org/unsd/nationalaccount/docs/SNA2008.pdf (with World Bank, IMF, European Commission and OECD).

Walmsley, T. L., et al. (2005). The GMig2 Data Base: A Data Base of Bilateral Labor Migration, Wages and Remittances. GTAP Research Memorandum. Center for Global Trade Analysis. West Lafayette, IN.

World Bank (2004-2011); *World bank Development Indicators* (WDI). http://data.worldbank.org/data-catalog/world-development-indicators. Washington, D.C., USA.

Appendices

Table A1-1 World Bank Development Indicator codes

BM_GSR_FCTY_CD	Primary income payments (BoP, current US$)
BM_GSR_GNFS_CD	Imports of goods and services (BoP, current US$)
BM_GSR_TOTL_CD	Imports of goods, services and primary income (BoP, current US$)
BM_TRF_PRVT_CD	Secondary income, other sectors, payments (BoP, current US$)
BM_TRF_PWKR_CD_DT	Personal remittances, paid (current US$)
BN_CAB_XOKA_CD	Current account balance (BoP, current US$)
BX_GSR_FCTY_CD	Primary income receipts (BoP, current US$)
BX_GSR_GNFS_CD	Exports of goods and services (BoP, current US$)
BX_TRF_CURR_CD	Secondary income receipts (BoP, current US$)
BX_TRF_PWKR_CD	Personal transfers, receipts (BoP, current US$)
BX_TRF_PWKR_CD_DT	Personal remittances, received (current US$)
GC_REV_GOTR_CN	Grants and other revenue (current LCU)
GC_REV_SOCL_CN	Social contributions (current LCU)
GC_REV_XGRT_CN	Revenue, excluding grants (current LCU)
GC_TAX_EXPT_CN	Taxes on exports (current LCU)
GC_TAX_GSRV_CN	Taxes on goods and services (current LCU)
GC_TAX_IMPT_CN	Customs and other import duties (current LCU)
GC_TAX_INTT_CN	Taxes on international trade (current LCU)
GC_TAX_OTHR_CN	Other taxes (current LCU)
GC_TAX_TOTL_CN	Tax revenue (current LCU)
GC_TAX_YPKG_CN	Taxes on income, profits and capital gains (current LCU)
GC_XPN_COMP_CN	Compensation of employees (current LCU)
GC_XPN_GSRV_CN	Goods and services expense (current LCU)
GC_XPN_INTP_CN	Interest payments (current LCU)
GC_XPN_OTHR_CN	Other expense (current LCU)
GC_XPN_TOTL_CN	Expense (current LCU)
GC_XPN_TRFT_CN	Subsidies and other transfers (current LCU)
NE_CON_GOVT_CD/CN	General government final consumption expenditure (current US$ or current LCU)
NE_CON_PETC_CD/CN	Household final consumption expenditure, etc. (current US$ or current LCU)
NE_CON_PRVT_CD/CN	Household final consumption expenditure (current US$ or current LCU)
NE_DAB_TOTL_CD/CN	Gross national expenditure (current US$ or current LCU)
NE_EXP_GNFS_CD/CN	Exports of goods and services (current US$ or current LCU)
NE_GDI_FTOT_CD/CN	Gross fixed capital formation (current US$ or current LCU)
NE_GDI_STKB_CD/CN	Changes in inventories (current US$ or current LCU)
NE_GDI_TOTL_CD/CN	Gross capital formation (current US$ or current LCU)
NE_IMP_GNFS_CD/CN	Imports of goods and services (current US$ or current LCU)
NY_GDP_DISC_CN	Discrepancy in expenditure estimate of GDP (current LCU)
NY_GDP_FCST_CD/CN	Gross value added at factor cost (current US$ or current LCU)
NY_GDP_MKTP_CD/CN	GDP (current US$ or current LCU)
NY_GDS_TOTL_CD/CN	Gross domestic savings (current US$ or current LCU)

Source: http://data.worldbank.org/indicator

Table A1-2 IMF Balance of Payments Statistics (BOPS) descriptions

Current Account, Net (BPM5)
Imports/Repayments of income/liabilities/outflows
Goods and Services, Debit (BPM5)
Income, Compensation of Employees, Debit (BPM5)
Income, Investment Income, Debit (BPM5)
Current Transfers, Debit (BPM5)
Income, Current Transfers, General Government, Debit (BPM5)
Income, Current Transfers, Other Sectors, Debit (BPM5)
Income, Current Transfers, Other Sectors, Workers' Remittances, Debit (BPM5)
Income, Current Transfers, Other Sectors, Other Transfers, Debit (BPM5)
Exports/Drawings of income/liabilities/inflows
Goods and Services, Credit (BPM5)
Income, Compensation of Employees, Credit (BPM5)
Income, Investment Income, Credit (BPM5)
Current Transfers, Credit (BPM5)
Income, Current Transfers, General Government, Credit (BPM5)
Income, Current Transfers, Other Sectors, Credit (BPM5)
Income, Current Transfers, Other Sectors, Workers' Remittances, Credit (BPM5)
Income, Current Transfers, Other Sectors, Other Transfers, Credit (BPM5)

Source: IMF BOPS, 1993

Table A1-3 IMF Government Financial Statistics (GFS) descriptions

	IMF taxes	Related to World Bank code
	Tax incomes	
Totals taxes	11 – total tax revenue	GC_TAX_TOTL_CN
Direct taxes	111 – taxes on income and profits (income tax)	GC_TAX_YPKG_CN
Other taxes (factor use and other)	112 – taxes on payroll (factor use tax) 113 – taxes on property (factor use tax) 116 – other taxes	GC_TAX_OTHR_CN
Goods and services taxes	114 – goods and services sales taxes (Includes VAT as a sub-category 11411)	GC_TAX_GSRV_CN
International Trade	115 – taxes on international trade (Import duties and export sub). Includes 1151 – imports duties 1152 – export taxes 1153 – rents on international trade	GC_TAX_INTT_CN includes: GC_TAX_IMPT_CN GC_TAX_EXPT_CN difference
	Other government Income	
Social security contributions	12 - social contributions total	GC_REV_SOCL_CN
Other government reventue	13 - Grants income 131 - Grants income (from foreign govts) 132 - Grants income (from Int Orgs) 133 - Grants income (other govt units) 14 - Other revenue 141 - Property income 142 - sales of G and S 143 - fines, penalties and forfeits 144 - Transfers nec 145 - Premiums, fees and claims etc	GC_REV_GOTR_CN
	Government expenses	
Compensation of employees (current LCU)	21 - compensation of employees	GC_XPN_COMP_CN
Goods and services expense (current LCU)	22 - use of G&S	GC_XPN_GSRV_CN

Consumption of fixed capital	23 - consumption of fixed capital	
Interest payments (current LCU)	24 - Interest paid	GC_XPN_INTP_CN
Other transfers	25 - subsidies 26 - grants expense 261 - grants expense (to foreign govts) 262 - grants expense (to int orgs) 263 - grants expense (to govt units) 27 - social benefits	GC_XPN_TRFT_CN
Other expense	28 - other exp 281 - property exp 282 - Transfers nec (exp) 283 - Premiums, fees and claims etc (exp)	GC_XPN_OTHR_CN

Source: adapted from IMF GFS (2014). World bank codes added.

Table A1-4 Countries in R23 Database and data sources used in prior

Country code	Country	Number of cells filled using:		Total number of cells filled	Share of cells filled with World Bank or IMF data
		World Bank data	IMF data		
abw	Aruba	7	11	41	44%
afg	Afghanistan	14	13	40	68%
ago	Angola	12	20	39	82%
aia	Anguilla	6	17	38	61%
alb	Albania	7	13	43	47%
ado	Andorra	0	0	36	0%
ant	Netherland Antilies	0	11	39	28%
are	United Arab Emirates	8	0	38	21%
arg	Argentina	7	13	43	47%
arm	Armenia	10	26	43	84%
atg	Antigua and Barbuda	14	21	43	81%
aus	Australia	22	11	43	77%
aut	Austria	8	28	43	84%
aze	Azerbaijan	7	13	43	47%
bdi	Burundi	6	9	41	37%
bel	Belgium	7	27	43	79%
ben	Benin	6	30	41	88%
bfa	Burkina Faso	7	27	41	83%
bgd	Bangladesh	7	31	43	88%
bgr	Bulgaria	12	22	41	83%
bhr	Bahrain	10	14	39	62%
bhs	Bahamas, The	14	18	41	78%
bih	Bosnia and Herzegovina	10	24	43	79%
blr	Belarus	8	24	43	74%
blz	Belize	13	21	43	79%
bmu	Bermuda	5	6	41	27%
bol	Bolivia	10	26	43	84%
bra	Brazil	20	13	43	77%
brb	Barbados	10	26	43	84%
brn	Brunei Darussalam	5	6	37	30%
btn	Bhutan	14	18	43	74%
bwa	Botswana	10	26	43	84%
caf	Central African Republic	8	2	40	25%
can	Canada	9	20	43	67%
che	Switzerland	9	27	43	84%

chl	Chile	8	20	43	65%
chn	China	9	26	42	83%
civ	Cote d'Ivoire	7	29	41	88%
cmr	Cameroon	7	11	41	44%
zar	Congo, Dem. Rep.	26	0	38	68%
cog	Congo, Rep.	8	26	41	83%
col	Colombia	7	28	42	83%
com	Comoros	7	2	40	23%
cpv	Cabo Verde	6	24	41	73%
cri	Costa Rica	13	25	43	88%
cub	Cuba	8	2	40	25%
cyp	Cyprus	12	20	41	78%
cze	Czech Republic	10	26	43	84%
deu	Germany	8	21	43	67%
dji	Djibouti	7	12	43	44%
dma	Dominica	9	25	43	79%
dnk	Denmark	8	21	40	73%
dom	Dominican Republic	9	28	43	86%
dza	Algeria	12	24	43	84%
ecu	Ecuador	7	13	43	47%
egy	Egypt, Arab Rep.	14	22	43	84%
eri	Eritrea	7	2	40	23%
esp	Spain	10	24	43	79%
est	Estonia	8	20	41	68%
eth	Ethiopia	8	25	43	77%
fin	Finland	8	25	43	77%
fji	Fiji	7	13	43	47%
fra	France	8	28	43	84%
fro	Faeroe Islands	1	9	39	26%
fsm	Micronesia, Fed. Sts.	2	2	40	10%
gab	Gabon	6	6	38	32%
gbr	United Kingdom	10	22	43	74%
geo	Georgia	8	24	43	74%
gha	Ghana	8	21	41	71%
gin	Guinea	7	10	41	41%
gmb	Gambia, The	10	24	41	83%
gnb	Guinea-Bissau	4	11	41	37%
gnq	Equatorial Guinea	9	17	38	68%
grc	Greece	10	26	43	84%
grd	Grenada	14	21	43	81%
grl	Greenland	0	0	36	0%
gtm	Guatemala	9	26	43	81%

guy	Guyana	7	12	43	44%
hkg	Hong Kong SAR, China	20	9	41	71%
hnd	Honduras	9	26	43	81%
hrv	Croatia	10	26	43	84%
hti	Haiti	4	10	41	34%
hun	Hungary	8	26	41	83%
idn	Indonesia	8	28	41	88%
ind	India	13	25	43	88%
irl	Ireland	10	26	43	84%
irn	Iran, Islamic Rep.	16	12	41	68%
irq	Iraq	6	7	39	33%
isl	Iceland	8	25	43	77%
isr	Israel	5	25	39	77%
ita	Italy	9	25	43	79%
jam	Jamaica	13	23	43	84%
jor	Jordan	10	25	43	81%
jpn	Japan	20	13	43	77%
kaz	Kazakhstan	11	25	42	86%
ken	Kenya	10	24	43	79%
kgz	Kyrgyz Republic	7	25	43	74%
khm	Cambodia	8	26	43	79%
kir	Kiribati	4	2	40	15%
kna	St. Kitts and Nevis	14	22	43	84%
kor	Korea, Dem. Rep.	20	13	43	77%
kwt	Kuwait	8	15	37	62%
lao	Lao PDR	14	20	41	83%
lbn	Lebanon	8	27	43	81%
lbr	Liberia	11	20	39	79%
lby	Libya	7	12	43	44%
lca	St. Lucia	13	22	43	81%
lie	Liechtenstein	0	0	36	0%
lka	Sri Lanka	7	27	43	79%
lso	Lesotho	9	24	43	77%
ltu	Lithuania	12	20	41	78%
lux	Luxembourg	10	24	43	79%
lva	Latvia	8	22	41	73%
mac	Macao SAR, China	13	17	40	75%
mar	Morocco	9	27	43	84%
mco	Monaco	0	0	36	0%
mda	Moldova	9	26	43	81%
mdg	Madagascar	21	6	40	68%
mdv	Maldives	10	21	43	72%

mex	Mexico	7	11	43	42%
mkd	Macedonia, FYR	9	28	42	88%
mli	Mali	7	23	41	73%
mlt	Malta	10	24	41	83%
mmr	Myanmar	1	10	41	27%
mne	Montenegro	8	9	43	40%
mng	Mongolia	9	27	42	86%
moz	Mozambique	7	11	41	44%
mrt	Mauritania	8	2	40	25%
mus	Mauritius	13	21	43	79%
mwi	Malawi	7	11	41	44%
mys	Malaysia	10	21	41	76%
nam	Namibia	9	27	43	84%
ncl	New Caledonia	1	11	39	31%
ner	Niger	5	29	39	87%
nga	Nigeria	11	16	43	63%
nic	Nicaragua	8	23	43	72%
nld	Netherlands	8	26	43	79%
nor	Norway	11	24	42	83%
npl	Nepal	12	23	43	81%
nzl	New Zealand	21	10	43	72%
omn	Oman	8	18	39	67%
pak	Pakistan	13	21	43	79%
pan	Panama	7	12	43	44%
per	Peru	8	21	43	67%
phl	Philippines	7	23	43	70%
plw	Palau	4	2	40	15%
png	Papua New Guinea	1	13	41	34%
pol	Poland	12	22	41	83%
pri	Puerto Rico	6	0	36	17%
prk	Korea, Rep.	0	2	38	5%
prt	Portugal	10	26	43	84%
pry	Paraguay	8	21	43	67%
wbg	West Bank and Gaza	23	2	40	63%
pyf	French Polynesia	1	11	39	31%
qat	Qatar	7	8	37	41%
rou	Romania	10	24	41	83%
rus	Russian Federation	12	16	41	68%
rwa	Rwanda	5	11	39	41%
sau	Saudi Arabia	6	8	39	36%
sdn	Sudan	7	12	43	44%
sen	Senegal	7	13	43	47%

sgp	Singapore	14	14	40	70%
slb	Solomon Islands	1	10	39	28%
sle	Sierra Leone	11	17	41	68%
slv	El Salvador	8	26	43	79%
smr	San Marino	2	13	35	43%
som	Somalia	0	2	38	5%
srb	Serbia	12	22	43	79%
stp	Sao Tome and Principe	12	19	41	76%
sur	Suriname	6	26	41	78%
svk	Slovak Republic	10	22	41	78%
svn	Slovenia	10	24	41	83%
swe	Sweden	8	27	43	81%
swz	Swaziland	11	26	42	88%
syc	Seychelles	2	24	41	63%
syr	Syrian Arab Republic	12	20	43	74%
tcd	Chad	8	2	40	25%
tgo	Togo	6	27	41	80%
tha	Thailand	6	24	41	73%
tjk	Tajikistan	6	13	43	44%
tkm	Turkmenistan	8	2	40	25%
tmp	Timor-Leste	10	2	40	30%
ton	Tonga	6	13	43	44%
tto	Trinidad and Tobago	20	10	43	70%
tun	Tunisia	7	31	43	88%
tur	Turkey	10	23	42	79%
tuv	Tuvalu	2	2	40	10%
twn	Taiwan	0	0	36	0%
tza	Tanzania	7	11	41	44%
uga	Uganda	7	26	43	77%
ukr	Ukraine	8	26	43	79%
ury	Uruguay	7	28	43	81%
usa	United States	11	22	41	80%
uzb	Uzbekistan	8	2	40	25%
vct	St. Vincent and the Grenadines	9	23	41	78%
ven	Venezuela, RB	7	13	43	47%
vnm	Vietnam	8	5	41	32%
vut	Vanuatu	7	11	41	44%
wsm	Samoa	3	13	43	37%
yem	Yemen, Rep.	1	12	41	32%
zaf	South Africa	8	25	43	77%
zmb	Zambia	9	24	41	80%
zwe	Zimbabwe	7	2	40	23%

xrw	Rest of world including: American Samoa, Antarctica, French Southern Territories, Bouvet Island, Cook Islands, Cayman Islands, Western Sahara, Falkland Islands (Malvinas), Guernsey, Gibraltar, French Guiana, Guam, Isle of Man, British Indian Ocean Territory, Jersey, Marshall Islands, Northern Mariana Islands, Montserrat, Mayotte, Niue, Nauru, Pitcairn, South Georgia and the South Sandwich Islands, Saint Helena, Saint Pierre and Miquelon, Turks and Caicos, Tokelau, United States Minor Outlying Islands, Holy See (Vatican City State), Virgin Islands (U.S.), British Virgin Islands, and Wallis and Futuna	0	2	38	5%

Table A1-5 Ranking of changes to World Bank and IMF data resulting from balancing of Macro SAMs and international transactions

Country code	Country	Share of cells filled with World bank or IMF data	Ratio of row to column total				
			ccom	hous	govt	Row	kap
abw	Aruba	44%	1.07	0.86	1.00	0.96	0.75
pri	Puerto Rico	17%	0.98	0.91	1.00	1.00	1.87
dji	Djibouti	44%	0.84	1.81	1.00	1.01	1.31
lca	St. Lucia	81%	0.99	0.85	1.04	1.00	1.55
irq	Iraq	33%	1.47	0.41	1.00	0.68	1.30
tmp	Timor-Leste	30%	1.08	6.97	1.00	1.10	-15.99
lbr	Liberia	79%	0.87	0.48	11.85	0.98	-0.56
xrw	Rest of world	5%	0.88	1.00	1.00	1.00	2.45
slb	Solomon Islands	28%	0.95	1.00	1.00	0.68	4.86
stp	Sao Tome and Principe	76%	0.99	-0.96	1.02	0.35	1.24
prk	Korea, Rep.	5%	0.93	1.00	1.00	1.00	1.60
rwa	Rwanda	41%	1.05	0.94	1.00	1.08	0.73
lso	Lesotho	77%	0.91	1.82	1.86	0.94	-1.12
bmu	Bermuda	27%	0.88	1.00	1.00	1.08	2.49
tjk	Tajikistan	44%	0.81	-5.90	1.00	0.86	-0.04
tkm	Turkmenistan	25%	0.88	1.46	1.00	1.00	1.46
kwt	Kuwait	62%	1.02	0.74	2.41	1.43	1.31
hkg	Hong Kong SAR, China	71%	1.00	0.92	0.90	1.01	1.40
omn	Oman	67%	1.00	0.70	1.89	1.12	1.64
tgo	Togo	80%	1.11	0.61	2.27	0.96	0.84
wbg	West Bank and Gaza	63%	0.96	0.98	2.89	1.67	-0.61
sgp	Singapore	70%	1.00	1.14	1.17	0.87	1.72
dma	Dominica	79%	1.01	0.78	1.09	1.02	1.51
uzb	Uzbekistan	25%	0.90	1.36	1.00	1.00	1.57
kir	Kiribati	15%	1.17	1.00	1.00	1.00	0.35
cog	Congo, Rep.	83%	0.93	0.37	0.33	0.88	4.51
pan	Panama	44%	0.93	0.94	1.00	0.99	1.96
gab	Gabon	32%	1.00	0.70	1.00	1.00	2.02
plw	Palau	15%	0.89	1.00	1.00	1.00	2.74
mne	Montenegro	40%	0.98	0.86	1.00	1.02	1.43
smr	San Marino	43%	1.01	1.00	2.50	1.00	0.65
tcd	Chad	25%	0.90	1.06	1.00	1.00	1.80
aia	Anguilla	61%	0.91	1.00	4.05	1.00	0.81
hti	Haiti	34%	1.03	1.28	1.00	0.84	0.48
nor	Norway	83%	1.00	0.64	4.72	0.96	1.52
bhr	Bahrain	62%	1.06	0.64	3.30	1.09	1.33
mmr	Myanmar	27%	1.08	1.00	1.00	0.91	-0.02
sau	Saudi Arabia	36%	1.04	0.76	1.00	1.12	1.43

fro	Faeroe Islands	26%	0.92	1.00	1.00	1.01	2.00
ton	Tonga	44%	0.97	2.04	1.00	0.81	0.10
gmb	Gambia, The	83%	1.15	0.61	1.96	0.93	1.03
cpv	Cabo Verde	73%	1.06	0.70	1.90	0.93	0.92
zwe	Zimbabwe	23%	0.95	1.18	1.00	1.00	0.94
gin	Guinea	41%	0.96	0.95	1.00	0.94	2.16
gnq	Equatorial Guinea	68%	0.97	0.63	0.96	1.00	1.63
mdv	Maldives	72%	0.79	1.00	0.94	1.17	1.95
qat	Qatar	41%	1.00	0.75	1.38	1.00	1.27
swz	Swaziland	88%	0.97	0.66	1.73	1.02	2.51
ago	Angola	82%	1.06	0.50	1.08	0.85	3.55
ant	Netherland Antilies	28%	0.96	1.00	1.00	1.00	1.37
gnb	Guinea-Bissau	37%	0.80	2.12	1.00	0.91	1.68
sle	Sierra Leone	68%	0.98	1.03	3.00	0.59	0.85
isr	Israel	77%	1.05	0.75	1.13	1.07	1.20
syc	Seychelles	63%	1.02	1.00	0.72	0.96	1.28
wsm	Samoa	37%	0.94	1.00	1.00	0.90	3.45
pry	Paraguay	67%	0.96	0.87	1.53	1.08	1.85
nga	Nigeria	63%	1.04	0.96	2.89	0.73	0.51
brn	Brunei Darussalam	30%	0.99	0.88	1.00	1.08	1.85
mda	Moldova	81%	0.99	1.14	2.12	0.93	0.33
lux	Luxembourg	79%	1.02	0.72	1.31	0.94	2.56
vct	St. Vincent and the Grenadines	78%	1.00	0.85	1.26	0.95	1.38
nic	Nicaragua	72%	0.91	0.95	2.44	1.05	1.17
usa	United States	80%	1.10	0.79	0.63	0.97	1.39
kgz	Kyrgyz Republic	74%	1.00	0.97	2.18	0.93	0.46
dza	Algeria	84%	0.89	0.98	1.24	0.99	1.55
pyf	French Polynesia	31%	1.06	1.00	1.00	1.01	0.37
afg	Afghanistan	68%	0.99	1.01	2.81	1.00	-0.25
tto	Trinidad and Tobago	70%	0.95	0.75	1.71	1.09	2.64
irl	Ireland	84%	1.00	0.70	1.64	0.97	1.88
are	United Arab Emirates	21%	0.96	0.96	1.00	1.00	1.50
mlt	Malta	83%	1.00	0.73	2.16	0.98	1.64
prt	Portugal	84%	1.01	0.77	1.65	0.94	1.57
zar	Congo, Dem. Rep.	68%	1.01	0.90	1.42	1.17	0.91
aze	Azerbaijan	47%	0.99	0.92	1.00	0.70	2.06
mac	Macao SAR, China	75%	1.01	0.72	0.89	1.09	1.42
com	Comoros	23%	0.97	1.04	1.00	1.00	1.44
nld	Netherlands	79%	1.00	0.90	0.86	0.98	1.66
ncl	New Caledonia	31%	1.01	1.00	1.00	0.99	0.92
moz	Mozambique	44%	0.99	1.08	1.00	0.82	1.59
gbr	United Kingdom	74%	0.97	0.80	1.96	1.02	1.53
hrv	Croatia	84%	1.00	0.74	2.38	0.98	1.40

cyp	Cyprus	78%	0.98	0.83	1.31	1.01	1.58
mys	Malaysia	76%	1.01	0.76	1.12	1.03	1.67
tha	Thailand	73%	1.00	0.75	1.59	1.03	1.58
cze	Czech Republic	84%	1.01	0.74	1.12	0.95	1.71
mng	Mongolia	86%	0.97	0.80	2.74	0.97	1.22
nzl	New Zealand	72%	0.97	0.79	1.55	0.95	1.74
est	Estonia	68%	0.97	0.80	2.24	0.96	1.44
bih	Bosnia and Herzegovina	79%	0.94	1.38	1.81	0.94	0.49
lao	Lao PDR	83%	1.06	0.93	1.13	0.92	0.82
lby	Libya	44%	0.94	0.99	1.00	1.07	1.48
hun	Hungary	83%	1.01	0.77	1.04	0.96	1.96
bwa	Botswana	84%	0.98	0.68	2.60	0.95	1.31
ner	Niger	87%	1.07	0.85	0.76	0.91	1.29
bdi	Burundi	37%	0.97	1.37	1.00	0.87	0.56
dom	Dominican Republic	86%	1.00	0.81	2.02	1.05	1.25
zaf	South Africa	77%	1.00	0.80	1.19	0.96	1.83
bhs	Bahamas, The	78%	0.94	0.89	1.75	0.93	1.69
kna	St. Kitts and Nevis	84%	1.00	0.77	2.16	0.99	1.21
mkd	Macedonia, FYR	88%	0.99	0.89	1.89	0.92	1.01
gha	Ghana	71%	0.96	0.81	1.80	1.17	1.47
ven	Venezuela, RB	47%	0.99	0.87	1.00	1.09	1.35
kaz	Kazakhstan	86%	1.00	0.84	0.87	0.86	1.72
grd	Grenada	81%	1.00	0.79	1.18	1.07	1.26
alb	Albania	47%	0.96	1.30	1.00	0.92	0.84
ben	Benin	88%	1.07	0.72	1.07	0.99	1.44
fji	Fiji	47%	0.98	0.98	1.00	0.99	1.43
grc	Greece	84%	1.00	0.76	2.06	0.94	1.48
syr	Syrian Arab Republic	74%	0.94	0.96	1.52	0.98	1.56
atg	Antigua and Barbuda	81%	1.00	0.82	1.05	1.00	1.36
brb	Barbados	84%	1.03	0.82	1.17	0.97	1.50
slv	El Salvador	79%	1.00	0.89	2.54	0.99	0.48
cub	Cuba	25%	1.00	0.87	1.00	1.00	2.11
swe	Sweden	81%	0.99	0.90	0.94	1.07	1.45
bol	Bolivia	84%	1.03	0.79	1.76	0.90	1.27
can	Canada	67%	0.98	0.92	1.01	1.05	1.47
vnm	Vietnam	32%	1.03	0.95	1.00	0.89	1.12
rou	Romania	83%	1.02	0.68	2.27	0.93	1.70
per	Peru	67%	0.99	0.81	1.56	0.85	1.77
jor	Jordan	81%	0.97	1.08	1.48	0.96	0.75
phl	Philippines	70%	0.98	0.88	2.06	0.97	1.31
dnk	Denmark	73%	1.00	0.82	1.17	1.00	1.48
isl	Iceland	77%	1.01	0.79	1.02	0.96	1.57
hnd	Honduras	81%	1.00	0.90	2.19	0.98	0.78

mar	Morocco	84%	0.98	0.93	1.67	0.92	1.05
blr	Belarus	74%	0.95	0.96	1.56	0.99	1.27
svn	Slovenia	83%	1.00	0.87	0.94	0.99	1.44
eth	Ethiopia	77%	0.96	1.08	1.72	0.98	0.68
srb	Serbia	79%	0.99	0.78	1.53	0.97	1.37
aus	Australia	77%	0.98	0.80	1.63	0.93	1.59
sen	Senegal	47%	0.99	1.07	1.00	0.94	1.00
chl	Chile	65%	1.01	0.85	1.24	0.86	1.70
zmb	Zambia	80%	0.99	0.93	1.14	0.81	1.82
png	Papua New Guinea	34%	0.97	1.00	1.00	0.97	1.58
nam	Namibia	84%	0.97	0.98	1.26	0.95	1.11
mdg	Madagascar	68%	1.00	0.88	1.56	1.00	1.11
bel	Belgium	79%	1.00	0.87	0.97	1.01	1.56
mli	Mali	73%	0.95	0.96	1.50	0.88	1.50
lka	Sri Lanka	79%	0.95	1.06	1.24	0.93	1.18
esp	Spain	79%	1.03	0.87	0.65	0.97	1.47
caf	Central African Republic	25%	0.99	1.04	1.00	1.00	0.92
mus	Mauritius	79%	0.99	0.84	1.30	1.06	1.32
lbn	Lebanon	81%	1.01	0.87	1.81	1.05	0.68
mwi	Malawi	44%	0.98	0.99	1.00	0.92	1.40
bgd	Bangladesh	88%	0.97	0.99	2.63	0.84	1.04
jam	Jamaica	84%	0.98	0.87	1.29	0.93	1.54
civ	Cote d'Ivoire	88%	0.97	0.85	1.11	1.05	2.78
fin	Finland	77%	1.00	0.87	0.88	1.01	1.55
blz	Belize	79%	1.00	0.81	1.32	0.94	1.73
npl	Nepal	81%	0.95	1.11	1.77	0.93	0.62
deu	Germany	67%	0.95	0.90	1.39	1.03	1.64
pak	Pakistan	79%	0.98	0.92	1.85	0.90	1.35
gtm	Guatemala	81%	0.99	0.97	2.23	0.87	0.81
bgr	Bulgaria	83%	0.99	0.80	1.20	0.90	1.73
geo	Georgia	74%	0.96	0.96	1.59	0.98	1.10
fra	France	84%	1.01	0.92	0.67	1.03	1.50
fsm	Micronesia, Fed. Sts.	10%	1.07	1.00	1.00	1.00	0.43
tza	Tanzania	44%	1.00	1.11	1.00	0.84	0.98
tur	Turkey	79%	1.00	0.81	1.40	1.04	1.56
egy	Egypt, Arab Rep.	84%	1.01	0.90	1.44	0.97	1.22
svk	Slovak Republic	78%	1.01	0.90	0.71	0.99	1.45
guy	Guyana	44%	1.03	0.97	1.00	0.96	0.94
aut	Austria	84%	1.00	0.83	1.09	1.00	1.67
arm	Armenia	84%	0.99	0.91	1.77	1.03	0.92
tuv	Tuvalu	10%	0.82	1.00	1.00	1.00	5.38
che	Switzerland	84%	1.00	0.95	0.64	0.94	1.59
kor	Korea, Dem. Rep.	77%	0.98	0.88	1.26	1.01	1.31
cmr	Cameroon	44%	0.98	0.97	1.00	0.85	1.91

sur	Suriname	78%	0.94	1.00	2.09	1.13	0.97
sdn	Sudan	44%	0.94	1.06	1.00	0.86	1.52
jpn	Japan	77%	0.96	0.95	1.02	1.08	1.50
bfa	Burkina Faso	83%	1.00	1.13	0.61	0.97	1.21
ukr	Ukraine	79%	0.97	0.89	1.78	0.98	1.26
khm	Cambodia	79%	0.98	0.91	1.88	0.98	1.35
pol	Poland	83%	1.00	0.85	1.09	0.95	1.61
lva	Latvia	73%	0.99	0.88	1.23	0.96	1.31
idn	Indonesia	88%	1.01	0.85	1.50	0.95	1.41
ken	Kenya	79%	1.01	0.88	1.37	0.97	1.13
ita	Italy	79%	1.00	0.85	0.98	1.00	1.64
mex	Mexico	42%	0.96	0.97	1.00	1.00	1.47
btn	Bhutan	74%	0.98	1.16	0.80	0.98	0.98
chn	China	83%	1.01	0.98	0.54	1.01	1.20
irn	Iran, Islamic Rep.	68%	1.01	0.87	0.68	1.00	1.33
uga	Uganda	77%	0.99	0.95	1.52	1.03	0.92
cri	Costa Rica	88%	1.00	0.89	1.01	0.97	1.51
mrt	Mauritania	25%	0.97	1.03	1.00	1.00	1.15
rus	Russian Federation	68%	0.97	0.96	1.23	0.95	1.32
ltu	Lithuania	78%	1.00	0.92	0.95	0.96	1.38
ind	India	88%	1.04	0.87	1.10	0.94	1.13
ury	Uruguay	81%	1.02	0.86	1.21	1.01	1.36
col	Colombia	83%	1.01	0.90	1.37	0.88	1.25
eri	Eritrea	23%	0.93	1.24	1.00	1.00	1.00
bra	Brazil	77%	1.02	0.83	0.93	0.95	1.70
vut	Vanuatu	44%	0.98	1.03	1.00	0.89	1.33
tun	Tunisia	88%	1.00	0.91	0.96	0.94	1.52
arg	Argentina	47%	0.97	0.95	1.00	1.05	1.40
ecu	Ecuador	47%	0.99	0.97	1.00	0.91	1.33
yem	Yemen, Rep.	32%	1.03	1.00	1.00	0.94	0.80
som	Somalia	5%	0.94	1.00	1.00	1.00	1.74
ado	Andorra	0%	0.81	1.00	1.00	1.00	2.61
grl	Greenland	0%	1.09	1.00	1.00	1.00	0.24
lie	Liechtenstein	0%	0.99	1.00	1.00	1.00	1.06
mco	Monaco	0%	1.00	1.00	1.00	1.00	1.03
twn	Taiwan	0%	1.01	1.00	1.00	1.00	0.81