A 2003 Social Accounting Matrix for Kenya: Methodological Note

Jane Kiringai
James Thurlow
Bernadette Wanjala
Nicholas Waiyaki
Clive Mutunga
Githinji Njenga
John Mutua
Nancy Nafula

Macroeconomics Division
Kenya Institute for Public Policy
Research and Analysis

KIPPRA Discussion Paper No. 72

KiPPRA
The KENYA INSTITUTE for PUBLIC
POLICY RESEARCH and ANALYSIS

KIPPRA IN BRIEF

The Kenya Institute for Public Policy Research and Analysis (KIPPRA) is an autonomous institute whose primary mission is to conduct public policy research leading to policy advice. KIPPRA's mission is to produce consistently high-quality analysis of key issues of public policy and to contribute to the achievement of national long-term development objectives by positively influencing the decision-making process. These goals are met through effective dissemination of recommendations resulting from analysis and by training policy analysts in the public sector. KIPPRA, therefore, produces a body of well-researched and documented information on public policy, and in the process assists in formulating long-term strategic perspectives. KIPPRA serves as a centralized source from which the Government and the private sector may obtain information and advice on public policy issues.

The Discussion Paper Series disseminates results and reflections from ongoing research activities of the Institute's programmes. The papers are internally refereed and are disseminated to inform and invoke debate on policy issues. Opinions expressed in the papers are entirely those of the authors and do not necessarily reflect the views of the Institute.

Table of Contents

List of Abbreviations and Acronyms

DAO	District Agricultural Officer
DLPO	District Livestock Production Officer
FAO	Food Agricultural Organization
IFPRI	International Food Policy Research Institute
GDP	Gross Domestic Product
ICT	Information and Communication Technology
IFS	Input Field Survey
IO	Input Output
ISPs	Internet Service Providers
KATA	Kenya Association of Travel Agents
KATO	Kenya Association of Tour Operators
KNBS	Kenya National Bureau of Statistics
KIPPRA	Kenya Institute for Public Policy Research and Analysis
KRA	Kenya Revenue Authority
LFS	Labour Force Survey
PAYE	Pay As You Earn
PCK	Postal Corporation of Kenya
SAM	Social Accounting Matrix
WMS	Welfare Monitoring Survey

Acknowledgements

1.   Introduction

A Social Accounting Matrix (SAM) is a consistent data framework that captures information contained in the national income and product accounts and the input-output table, as well as the monetary flows between institutions. A SAM is an ex-post accounting framework since, within its square matrix, total receipts must equal total payments for each account contained within the SAM. Since the required data is not drawn from a single source, information from various sources must be compiled and made consistent. This process is valuable since it identifies inconsistencies between Kenya's statistical sources and highlights areas where data reliability is weakest. However, the SAM is an economy-wide database that is typically used in conjunction with analytical techniques to strengthen the evidence underlying policies. The construction of the 2003 SAM was a collaborative initiative between the Kenya Institute for Public Policy Research and Analysis (KIPPRA) and the International Food Policy Research Institute (IFPRI). It formed part of a broader research project to identify potential sources of growth in Kenya.

There are two approaches to constructing a SAM. The first uses statistical methods to update a previous SAM to reflect new macroeconomic data (e.g. GDP estimates and balance of payments). This approach, therefore, assumes that the structural characteristics of the economy have not changed significantly over time. This is an important assumption since one of the strengths of a SAM is its capturing of production technology and linkages between different sectors/ institutions in the economy. Therefore, while this first approach is expedient, it is inappropriate during a period of transformation, when production technology is most likely to change and structural linkages become more complex. The second approach uses surveys to estimate the production technology underlying different sectors of the economy. These surveys collect information on input usage, labour employment, and capital assets. This second approach is expensive and time-consuming, but ensures that the SAM is an up-to-date representation of the economy.

There are mainly two steps in SAM construction. The first step is to compile information from various sources into the SAM format or framework. This is known as the 'prior SAM'. The construction of the prior SAM takes place in two stages. A 'macro SAM' is first constructed

using aggregate information from national accounts and other macroeconomic databases. This SAM is then disaggregated across sectors, factors and households to derive a detailed 'micro SAM'. Given the diversity of its data sources, the prior SAM is inconsistent (i.e., there are inequalities between receipts and payments). The second step in constructing a SAM is reconciling receipts and payments so that row and column totals are equal (i.e., 'balancing the SAM'). This is also done in two stages. The reliability of the various data sources is first assessed based on the observed inequalities between row and column accounts. The SAM is then balanced using cross-entropy econometrics.

This paper documents the construction of a 2003 Social Accounting Matrix (SAM) for Kenya. The micro SAM was constructed using survey data, which was necessitated by lack of a recent and up-to-date input-output table. Input surveys were carried out to provide the information used to construct a new input-output table.

The rest of the paper is organized as follows. Section 2 reviews the general structure of social accounting matrices and also outlines the proposed structure of the Kenya SAM. Section 3 describes the data sources used to construct the prior SAM. Section 4 describes the cross-entropy approach used in balancing the micro SAM, together with a description of the constraints imposed during the estimation procedure. Section 5 highlights some of the limitations of the SAM construction process.

2.   Structure of a Social Accounting Matrix

A SAM is an economy-wide data framework that usually represents the real economy of a single country.[1] More technically, a SAM is a square matrix in which each account is represented by a row and column. Each cell shows the payment from the account of its column to the account of its row—the incomes of an account appear along its row, its expenditures along its column. The underlying principle of double-entry accounting requires that for each account in the SAM, total revenue (row total) equals total expenditure (column total).

A SAM provides an important framework for analyzing employment, poverty, growth and income distribution. By integrating household survey data in national accounts, the SAM captures the macro and meso level transactions of an economic system and micro level transfers between all economic agents in the economy (King B., 1985; Pyatt and Round 1985; Reinert K. and Roland-Holst, 1997). In addition to providing a consistent framework of national accounts, a SAM also incorporates income distribution dimensions by disaggregating households by socio-economic characteristics.

A SAM has six standard accounts: activities/commodities accounts, factors of production, institutions, government, rest of the world, and the capital account. The dimensions of the matrix are determined by the level of disaggregation of the six standard accounts. Table 1 shows the structure of an aggregate SAM (with verbal explanations in place of numbers).

Activities and Commodities

The SAM distinguishes between 'activities' (the entities that carry out production) and 'commodities' (representing markets for goods and non-factor services). SAM flows are valued at producers' prices in the activity accounts and at market prices (including indirect commodity taxes and transaction costs) in the commodity accounts. The commodities are activity outputs, either exported or sold domestically, and imports. In the activity columns, payments are made to commodities

---

[1] For general discussions of SAMs, see Pyatt and Round (1985) and Reinert and Roland-Holst (1997); for perspectives on SAM-based modeling, see Pyatt (1988), and Robinson and Roland-Holst (1988).

Table 1: Basic structure of the Kenya SAM

	Activities	Commodities	Factors	Households	Enterprises	Government	Investment	Rest of the World	Total
Activities		marketed outputs		home-consumed outputs					activity income
Commodities	intermediate inputs	transactions costs		private consumption		government consumption	investment	exports	total demand
Factors	value-added								factor income
Households			factor income to households		surplus to households	transfers to households		transfers to households from rest of the world	household income
Enterprises			factor income to enterprises			transfers to enterprises		transfers to enterprises from rest of the world	enterprise income
Government		sales taxes, import tariffs	factor income to government	direct household taxes	surplus to govt. enterprise taxes			transfers to government from rest of the world	government income
Savings				household savings	enterprise savings	government savings	change in stocks	foreign savings	savings
Rest of the World		imports			surplus to rest of the world	govt. transfers to rest of the world			Foreign exchange outflow
Total	activity expenditures	total supply	factor expenditures	household expenditures	enterprise expenditures	government expenditures	investment	foreign exchange inflow	

(intermediate demand), and factors of production (value-added comprising of operating surplus and compensation of employees). In the commodity columns, payments are made to domestic activities, the rest of the world, and various tax accounts (for domestic and import taxes). This treatment provides the data needed to model imports as perfect or imperfect substitutes *vis-à-vis* domestic production.

The SAM is derived from a symmetric input-output (I-O) table, with activity-commodity identity. This implies that each activity produces only one principal commodity and, therefore, by-products and secondary products do not exist. The supply table is, therefore, a diagonal matrix. Even though symmetric I-Os are not realistic, they are easier to construct at first attempt given data limitations. Future attempts can be made towards construction of a non-symmetric SAM.

Trade and Marketing Margins

Domestic and international trade flows in the SAM are explicitly associated with transactions (trade and transportation) costs, also referred to as marketing margins. For each commodity, the SAM accounts for the costs associated with domestic, import, and export marketing margins (i.e., each commodity purchases other trade and transport commodities). For domestic marketing of domestic output, the marketing margin represents the cost of moving the commodity from the producer to the domestic consumer. For imports, it represents the cost of moving the commodity from the border to the domestic market, while for exports it shows the cost of moving the commodity from the producer to the border.

Government Income and Payments

Government is disaggregated into a core government account and different tax collection accounts, one for each tax type. This disaggregation is necessary, otherwise the economic interpretation of some payments would be ambiguous. In the SAM, direct payments between the government and other domestic institutions are reserved for transfers. Finally, payments from the government to factors (for the labour services provided by public sector employees) are captured in the government services activity. Government consumption demand is

a purchase of the output from the government services activity which, in turn, pays labour.

Domestic Non-Government Institutions

The domestic non-government institutions consist of households and enterprises. The enterprises earn factor incomes (a reflection of ownership of capital and/or land) and may also receive transfers from other institutions. Their incomes are used for corporate taxes, enterprise savings, and transfers to other institutions. Unlike households, enterprises do not demand commodities. It is possible to disaggregate the enterprise sector in a manner that captures differences across enterprises in terms of tax rates, savings rates, and the shares of retained earnings that are received by different household types.

Home and Final Household Consumption

The SAM distinguishes between home (own) consumption of activities and marketed consumption of commodities by households. Home consumption, which in the SAM appears as payments from household accounts to activity accounts, is valued at producer prices, i.e., without marketing margins and the sales taxes that may be levied on marketed commodities. Final household consumption of marketed commodities appears as payments from household accounts to commodity accounts, valued at consumer prices that include marketing margins and commodity taxes.

3. Constructing the Unbalanced Prior Social Accounting Matrix

The initial task in building a SAM involves compiling data from various sources into the framework outlined in Section 2. This information is drawn from national accounts, household surveys, foreign trade statistics, government budgets, balance of payments, and various other publications.

This information often uses:

(i)   different disaggregation of sectors, production factors, and socio-economic household groups;

(ii)   different years and/or base-year prices; and

(iii)   different data collection and compilation techniques.

Consequently, the initial or prior SAM inevitably includes imbalances between row and column account totals.

The underlying prior macro-SAM is based on Kenya's revised national accounts, which conform to the 1993 System of National Accounts.[2] In the disaggregated SAM, all the cells in the macro-SAM can be expanded into several accounts depending on the availability of data. The disaggregated micro-SAM is built such that the aggregate totals from the macro-SAM are preserved (i.e. shares are used from other sources rather than actual numbers).

This section explains how each macro-SAM entry is derived and then disaggregated to arrive at the prior micro-SAM. The notation for the SAM entries is row then column, and the values are in millions of Kenya Shillings.

---

[2] Note that the entries in Table 2 may not exactly match national accounts since the values in the macro SAM are 'post-reconciliation' results (i.e., estimated using the cross-entropy econometric approach outlined in section 5). In other words, national accounts are used to construct the prior macro- SAM. Although these values are largely preserved during the balancing procedure, they do deviate marginally from reported values. Furthermore, there are existing inconsistencies in Kenya's national accounts such that the same item may be assigned slightly different values in different accounts. Also, the current national accounts have a statistical discrepancy, which has to be considered when balancing the SAM.

Table 2: 2003 macro SAM for Kenya (Ksh millions)

	Activities	Commodities	Factors	Enterprises	Households	Taxes	Government	Investment	Rest of the World	Total
Activities		1,783,049			95,043					1,878,092
Commodities	867,692	117,117			772,972		202,913	196,723	281,387	2,438,804
Factors	1,010,400									1,010,400
Enterprises			544,860				41,297		4,909	591,066
Households			461,261	335,194			17,898		91,014	905,367
Taxes		131,756		37,053	33,603					202,412
Government			4,279	7,332	6,298	202,412			5,677	225,998
Savings				204,248	-2,549		-36,286	17,498	31,310	214,221
Rest of the world		406,882		7,239			176			414,297
Total	1,878,092	2,438,804	1,010,400	591,066	905,367	202,412	225,998	214,221	414,297	

(Factors, Activities)...1,010,400

This is the value of Gross Domestic Product (GDP) at factor cost or, alternatively, total value-added generated by land, labour and capital. The aggregate value is drawn from national accounts. GDP in the micro-SAM is disaggregated across 50 sectors (22 agriculture, 18 industry, and 10 services) (see Table A1 in the appendix). According to national accounts, non-agricultural GDP is disaggregated across sectors according to national accounts, which reports compensation to workers and gross operating surplus. Agricultural GDP is disaggregated according to the value of 2002/03 output reported in the KIPPRA Agricultural Data Compendium (Gitu, K. W. and Nzuma J., 2003). Sectoral value-added is further disaggregated by labour, capital and land using factor shares from the 2003 Kenya input matrix (see Box 1).

Labour is further disaggregated across three skill groups (i.e., skilled, semi-skilled, and unskilled) using occupational and income data drawn from the 1998/99 Integrated Labour Force Survey (Government of Kenya, 2003). Professional and managerial workers are classified as 'skilled'; clerical, technical and manual workers (excluding agricultural workers) are classified as 'semi-skilled'; and the remaining occupational categories (including agricultural and elementary workers) are classified as 'unskilled'.

(Commodities, Activities)...867,692

This is the value of intermediate inputs used in the production process. The aggregate value is taken from national accounts. However, the disaggregation of intermediate demand across individual commodities is derived using the sectoral GDP estimate described above and the new input coefficients (see Box 1). The ratio of intermediate demand to total value-added for each sector is calculated from the input-output matrix. This is then multiplied by the value of each sector's GDP in order to determine the value of intermediate demand for each commodity in each sector.

This sub-matrix is the largest component of the SAM, which captures the technology of production in the economy through the input-output coefficients. It is also referred to as the make/supply and use matrix (Sadoulet and de Janvry, 1995). In Kenya, the changes in the technical coefficients are likely to emanate from the following sources: (i) trade

Box 1: New technical coefficients: Input field survey methodology

The 2003 Input Field Survey (IFS) was undertaken to compute technical or factor/intermediate input coefficients for the productive sectors and to disaggregate agriculture across major crops and livestock sub-sectors. A further objective was to identify input usage across major agro-ecological zones, although the regionalized version of the SAM is not presented in this paper. The five zones included upper highlands, lower highlands, upper midlands, lower midlands, and interior/coastal lowlands.

The survey sample was selected according to the following criteria: (i) the agricultural sector should cover the major crops and agricultural activities; (ii) the sampled districts should cover the five major agro-ecological zones; (iii) the linkages between agriculture and non-agriculture should be captured; and (iv) those areas where agricultural research institutes identify significant technology changes should be captured. The sampled districts were, therefore, selected according to the above criteria. For agriculture, multi-stage purposeful sampling was used to select the field sample. In the first stage, all districts in the country were mapped onto the five agro-ecological zones using geographically referenced data (see Table A2 in the appendix). During the second stage, those districts that most closely reflected their corresponding agro-ecological zone were selected (with a minimum of two districts in each zone). The final sample included 20 districts.

In each district, the team began by interviewing the District Agricultural Officer and the District Livestock Production Officer, who assisted in ranking the top five crops and top three livestock commodities in their district based on production values. For each commodity, two farmers (one with above average yields and one below average yields) were asked to provide a detailed farm budget. In total, 38 crops and 9 livestock commodities were covered by the survey (see Table A2 in the appendix). The District Agricultural Officers also separately prepared detailed budgets for the selected commodities.

For manufacturing and services, the sampling frame was determined using a two-stage sampling process identifying industries/sectors where technology has changed considerably and where it has remained largely unchanged. For sectors where technology has changed, representative

industries were surveyed to generate new technical coefficients. After identifying the target industries, a purposeful sample was selected based on the three criteria—ensure regional representation, ensure that sectors identified in the SAM were adequately captured, and to capture linkages with agriculture.

For services, three sectors were identified based on their contribution to the overall economy: financial, tourism (hotels and restaurants), and transport services. The financial services sample was based on a random sample drawn from the Kenya National Bureau of Statistics Business Register and covered 15 banks, 15 insurance companies, three ranked SACCO and 11 micro-finance institutions. For hotels and restaurants, stratified random sampling was used to ensure that a variety of hotel 'classes' were surveyed. The selection of hotels was done systematically by choosing every 3rd hotel in the business register. However, in certain cases, purposeful selection was used in order to ensure that hotel chains in other parts of the country were included (e.g., Maasai Mara and Amboseli Samburu). For tour operators, the Kenya Association of Tour Operators membership list was used to select firms to be interviewed. Every 25th member in the list was selected. However, in cases where the list did not have up-to-date contact information, the next firm on the list was chosen. For travel agents, the Kenya Association of Travel Agents ordinary members list was used. Every 5th firm was chosen until the required numbers for Nairobi and Coast were achieved. For transport and communication, a sample of 35 firms was selected from both sectors according to the business register, although this small sample captured a significant value share of the industry. The sampled service providers in Information and Communication Technology (ICT) covered telecommunications, which includes landline, mobile telephone and informal service providers including *'simu ya jamii'*. In postal services, the sample included public service providers such as the Postal Corporation of Kenya and private mail courier providers. In internet, Internet Service Providers (ISPs), computer hardware and software firms and cyber cafes were sampled. The transport sub-sector sample included public, private and informal sector providers in public and private passenger transport, freight, pipeline and air transport.

liberalization and the reduction of both tariff and non-tariff barriers is likely to increase the share of imported intermediate inputs; (ii) changes in prices; (iii) changes in the degree of substitution as the economy changes. To capture the change in technology, the 2003 SAM used both survey and non-survey techniques in constructing the input output table (Box 1).

(Activity, Households)...95,043

The payment from households to activities represents households' consumption of own production. This production is measured at producer (or farm-gate) prices. The total level of household consumption (both own produced and marketed) was taken from national accounts. This was distributed across own and final consumption using information from the 1997 Welfare Monitoring Survey (WMS). Respondents in the WMS were asked to keep a journal of incomes and expenditures over a varied period of one week, one month and one year. Regarding consumption, households were asked to state whether the commodities that were produced by the household were also consumed in the home. According to the WMS, 10.9 per cent of the value of total household consumption was for goods produced by those households themselves.

(Activities, Commodities)...1,783,049

This represents gross output, which is derived as the sum of intermediate demand, GDP at factor cost, and activity taxes, less non-marketed production (own production, own consumption). While the SAM distinguishes between activities and commodities, and thus would facilitate interactions between single/multiple activities and single/multiple commodities, the IO table does not allow for this distinction. Therefore, the disaggregation of this cell in the micro-SAM results in single entries along the main diagonal of the activity-commodity sub-matrix (or a one-to-one mapping between activities and commodities).

(Commodities, Commodities)...117,117

The payment by commodities to commodities is a condensed version of the treatment of trade margins (transaction costs) in the final micro

SAM. In the micro-SAM, there are separate margin accounts for the trade costs incurred through the marketing of each commodity. This value of transaction costs is further disaggregated to distinguish between the costs incurred by imports, exports, and domestically produced and sold goods. Unlike most entries in the SAM, this entry was first calculated on a disaggregated level, and then aggregated to arrive at a final macro-SAM value. Total trade margins were taken from the individual commodity accounts within national accounts. These were distributed across imports, exports and domestic markets according to the value of each item in total demand or supply.

(Taxes, Commodities)...131,756

While the macro SAM in Table 2 shows only a single row and column for taxes, this account actually consists of a number of distinct tax accounts. These include specific accounts for activities (corporate), income, sales, and import taxes. The commodity tax entry can, therefore, be disaggregated to include indirect sales taxes (110,966) and import tariffs (20,790). These aggregate values of individual taxes were taken from national accounts. Indirect or sales taxes include excise duties on petroleum products, excise duties on beverages and tobacco, insurance premium tax, value-added taxes, and other taxes on products. Import tariffs, excise duties and value-added taxes for individual commodities were derived from trade data provided by the Kenya Revenue Authority. It is initially assumed that sales tax rates (excluding excise duties) as recorded in trade data can be applied to all commodities sold through markets in Kenya. Since taxes are easier to collect on commodities traded abroad, this assumption leads to a higher tax collection than is recorded in national accounts. As such, each commodity's tax collection was scaled proportionately to match the aggregate value appearing in national accounts.

(Rest of World, Commodities)... 406,882

The value of total imports of goods and services was initially taken from national accounts. The statistical discrepancy (48,516 in national accounts) was added to the value of imports, since international trade data is assumed to be less reliable. The disaggregation between goods and services was taken from national accounts. Goods' imports were then disaggregated based on import data from the Kenya Revenue

Authority, while service imports come from national accounts (i.e., for transport, travel, insurance, and financial services).

(Commodities, Households)...772,972

As already noted in the discussion of own household consumption (Activities, Households), the disaggregation of total household consumption from the national accounts used information from the 1997 Welfare Monitoring Survey. According to the survey, final household demand constitutes 89.1 per cent of total household consumption spending. Final consumption was distributed across commodities using consumption shares from the WMS.

(Commodities, Government)...202,913

The total value of government consumption spending is taken from national accounts. All government spending is for the purchase of administrative services, education and health. In this way, the government is treated as a sector producing government services as well as a demander of these services.

(Commodities, Investment)...196,723

The aggregate value and commodity-specific value of investment demand is taken from national accounts. It includes capital investment and changes in stocks. In the micro SAM, this investment is disaggregated across commodities.

(Commodities, Rest of World)...281,387

The aggregate value of export demand is taken from national accounts. This is disaggregated across commodities using the individual commodity accounts in national accounts.

(Enterprises, Factors)...544,860

The residual of capital value-added (or gross operating surplus less land rents) is paid to enterprises. Total value-added is reduced by direct transfers of operating surplus to households and government. All

remaining capital earnings are transferred to enterprises and are, therefore, subject to direct/corporate taxes.

(Households, Factors)... 461,261

This value is the sum of all land and labour value-added generated during production. The distribution of labour income across households is determined using household labour income shares as reported in the 1997 Welfare Monitoring Survey and the 1998 Labour Force Survey. Land incomes are distributed according to the amount of land households reported that they owned and currently have under cultivation. This is scaled to match national accounts, which identifies the value of gross operating surplus that households reported receiving directly from the production. Labour value-added is equivalent to the compensation paid to labour and its aggregate value is drawn from national accounts.

(Government, Factors)...4,279

This is the value in national accounts of gross operating surplus that government reported receiving directly from the production process (e.g. from profits on public enterprises or dividends from the Central Bank).

(Enterprises, Government)...41,297

This is the value of interest payments paid by the government to domestic financial enterprises. It is taken from national accounts.

(Enterprises, Rest of World)...4,909

This is property income received by domestic enterprises from abroad. It is taken from national accounts.

(Households, Enterprises)...335,194

This is the value of mixed income received by households less the allowance for depreciation of fixed assets for firms. Depreciation is a component of domestic savings and is typically assumed to form part of household savings. However, in the Kenya SAM, we assume that the

value of depreciation is a component of enterprise savings. Therefore, enterprise savings are higher in the SAM than in national accounts and less of the 'income' received by enterprises is paid to households.

Enterprise payments are distributed across the different household groups in the SAM based on information from the 1997 Welfare Monitoring Survey. The information, which acts as a proxy for the distribution of indirect capital earnings, included profits earned from operation of non-farm enterprises, dividends received, and income from renting of land and equipment.

(Government, Enterprises)...7,332

Total transfers paid by enterprises to the government are taken from national accounts. They comprise interest on property, distributed income of corporations, property rent payments, and social contributions paid by enterprises.

(Taxes, Enterprises)...37,053

These are corporate taxes paid by enterprises to the government and are derived from national accounts.

(Savings, Enterprises)... 204,248

Enterprise savings are treated as a residual and, as mentioned above, include the allowance for or consumption of fixed capital.

(Rest of World, Enterprises)...7,239

This is income from abroad on properties owned by domestic enterprises. This is taken from national accounts.

(Households, Government)...17,898

Transfers from the government to households are taken from national accounts and include social benefits and domestic miscellaneous current transfers. They are distributed across different household groups

according to the pensions and social grants that households reported receiving in the 1997 Welfare Monitoring Survey.

(Households, Rest of World)...91,014

Aggregate household income from the rest of the world is taken from national accounts and is equal to international miscellaneous current transfers. This is distributed across households according to the cash and in-kind remittances received from abroad by households in the 1997 Welfare Monitoring Survey.

(Taxes, Households)...33,603

The value of direct taxes on households is equivalent to PAYE taxes and is taken from national accounts. This value is initially distributed across households according to the 2003 statutory tax rates and households' inflation-adjusted incomes/expenditures as reported in the 1997 Welfare Monitoring Survey. The income tax that households should have paid in the absence of exemptions is then scaled down to match the value of actual direct tax collections as reported in national accounts.

(Government, Households)...6,298

Transfers from households to government are taken from national accounts. These are mainly social contributions such as pension contributions. They are distributed across different household groups according to the pensions and social grants that households reported paying in the 1997 Welfare Monitoring Survey.

(Savings, Households)...-2,549

Household savings is taken from national accounts. Savings are distributed across households according to information in the 1997 Welfare Monitoring Survey. The information, which acts as a proxy for the distribution of household savings, included cash deposits in the bank, savings over the survey period, and loans received (note that household savings in national accounts is negative, implying aggregate dis-saving).

(Government, Taxes)...202,412

The tax accounts in the micro-SAM are separated into import tariffs, sales taxes and direct taxes. Each account sums tax revenue from all sources and then transfers these funds to the government. The entries correspond to national accounts.

(Government, Rest of World)...5,677

Government income from the rest of the world is equivalent to the value of current international co-operation (outflows).

(Savings, Government)...-36,286

This is fiscal deficit and is taken from national accounts.

(Rest of World, Government)...176

Government payments to the rest of the world are equivalent to the value of current international co-operation (inflows).

(Savings, Investment)...17,498

This figure represents the change in stocks. It is taken from national accounts. There is no further disaggregation in the micro-SAM.

(Savings, Rest of World)...31,310

This is the current account deficit or the total value of foreign savings. It is derived from national accounts.

4.   Balancing Procedure

The previous section outlined the construction of the prior micro-SAM. The range of datasets used during this procedure implies that the prior micro-SAM will inevitably be unbalanced (i.e., row and column totals are unequal). Cross-entropy econometrics was used to reconcile accounts in the 2003 Kenya SAM (see Robinson *et al.*, 2001). This approach begins with the construction of the prior SAM which, as explained in the previous section, used a variety of data from a number of sources of varying quality. This prior SAM provided the initial 'best guess' for the estimation procedure. Additional information is then brought to bear, including knowledge about aggregate values from national accounts and technology coefficients. A balanced Kenya SAM was then estimated by minimizing the entropy 'distance' measure between the final SAM and the initial unbalanced prior SAM, taking into account all additional information.

Table 3 summarizes the equations defining the SAM estimation procedure. Starting from an initial estimate of the SAM, additional information is imposed in the form of constraints on the estimation. Equation 1 specifies that row sums and corresponding column sums must be equal, which is the defining characteristic for a consistent set of SAM accounts. Equation 2 specifies that sub-accounts of the SAM must equal control totals, and that these totals are assumed to be measured with error (equation 3). An example would be the estimate of GDP provided by national accounts, which is the total value of the Factor-Activity matrix in the prior SAM. The matrix G is an aggregator matrix, with entries equal to 0 or 1. The index k is general and can include individual cells, column/row sums, and any combination of cells such as macro aggregates. Equation 4 allows for the imposition of information about column coefficients in the SAM rather than cell values, also allowing for error (Equation 5).

The error specification in equations 2 and 3 describes the errors as a weighted sum of a specified 'support set' (the V parameters). The weights (W) are probabilities to be estimated, starting from a prior on the standard error of measurement of either aggregates of flows (equation 8) or coefficients (equation 9). The number of elements in the error support set (w) determines how many moments of the error distribution are to be estimated. The probability weights must be non-negative and sum to one (equations 8 and 9). The objective function is the cross-entropy distance between the estimated probability weights and their

Table 3: Cross entropy SAM estimation equations

Index	Definition
i, j	row (i) and column (j) entries
k	set of constraints
w	set of weights

Symbol	Definition
$T_{i,j}$	SAM in values
$A_{i,j}$ and $\overline{A_{i,j}}$	SAM in column coefficients
$G_{k,i,j}$	aggregator matrix for each constraint k
γ_k and $\overline{\gamma}_k$	aggregate value for constraint k
e_k	error on each constraint k
$e_{i,j}^A$	error on each cell coefficient
W and $\overline{W}$	weights and prior on error term for each constraint k or cell coefficient i,j
$\overline{V}$	error support set indexed over w for each constraint k or cell coefficient i,j

Equations	
$$\sum_j T_{i,j} = \sum_j T_{j,i}$$	(1)
$$\sum_i \sum_j G_{k,i,j} \cdot T_{i,j} = \gamma_k$$	(2)
$$\gamma_k = \overline{\gamma}_k + e_k$$	(3)
$$A_{i,j} = \frac{T_{i,j}}{\sum_i T_{i,j}} \quad \text{with} \quad \sum_i A_{i,j} = 1 \,\forall\, i$$	(4)
$$A_{i,j} = \overline{A}_{i,j} + e_{i,j}^A \quad \text{for some i, j}$$	(5)
$$e_k = \sum_w W_{k,w} \cdot \overline{V}_{k,w}$$	(6)
$$e_{i,j}^A = \sum_w W_{i,j,w}^A \cdot \overline{V}_{i,j,w}^A$$	(7)

$$\sum_{w} W_{k,w} = 1 \quad \text{with} \quad 0 \le W_{k,w} \le 1 \qquad (8)$$

$$\sum_{w} W_{i,j,w}^{A} = 1 \quad \text{with} \quad 0 \le W_{i,j,w}^{A} \le 1 \qquad (9$$

$$\min\left[\sum_{k} \sum_{w} W_{k,w} \cdot \left(\ln W_{k,w} - \ln \bar{W}_{k,w} \right) + \sum_{i} \sum_{j} \sum_{w} W_{i,j,w}^{A} \left(\ln W_{i,j,w}^{A} - \ln \bar{\bar{W}}_{i,j,w}^{A} \right) \right]$$

$$(10)$$

prior for the errors in both coefficients and aggregates of SAM flows. It can be shown that this minimand is uniquely appropriate, and that using any other minimand introduces unwarranted assumptions (or information) about the errors.

Various constraints were imposed on the model according to the perceived reliability of the Kenyan data. Certain values that appeared in national accounts were maintained in order to remain consistent with the overall macro structure of the Kenyan economy. The macroeconomic aggregates that were maintained in the micro-SAM include: total labour value-added; total capital value-added; household own demand; household final demand; government spending; investment demand; exports; government borrowing/saving; current account balance; sales taxes; import tariffs; direct taxes on households; and direct taxes on enterprises.

5.   Study Limitations

There were several problems encountered in the construction of the SAM. First, the national accounts for 2003 had a discrepancy of Ksh 48 billion on the demand side, which was reflected in the macro SAM. Assumptions were made on how to deal with this discrepancy, which were subjective. The discrepancy was mainly reflected in imports, but not domestic production. Second, data collection for the input-output coefficients was challenging. Getting highly disaggregated expenditures and incomes was difficult because of poor record keeping by most organizations. Third, the data on household consumption was inconsistent in terms of time periods. The household survey consumption data was taken in three timeframes: weekly, monthly and annually. Therefore, assumptions were made to transform the data from weekly or monthly to annually by assuming a constant consumption pattern. This is subjective because households do not have regular consumption patterns throughout the year, i.e. there are episodes of low consumption especially during drought, and also depending on the number of harvest seasons. Fourth, mapping between factors of production and household income was also challenging. This was because the two household surveys that were used (Welfare Monitoring Survey and Labour Force Survey) had a similar sampling frame within regions and clusters, but some households were different. Therefore, some households were in the Welfare Monitoring Survey but not in the Labour Force Survey and vice versa. The SAM only made use of households that were common between the two surveys, which effectively reduced the sample size. Fifth, it is recognized that the informal sector in Kenya is quite large and plays a big role in the economy in terms of output and employment. However, this sector was not adequately captured in the SAM because of lack of information on the sector, for instance, its contribution to sectoral output. However, considerable attempts were made to capture the sector through composition of factors of production (labour and capital), and also enterprises. Lastly, resource constraints limited our sample size for data collection on input-output coefficients.

6.   Conclusion

This paper has outlined the methodology used to construct the 2003 SAM for Kenya. Construction of the SAM was necessitated by the lack of an up-to-date tool for analyzing policies in Kenya, especially at the sectoral level. The 2003 SAM adopted the survey methodology, which provided an up-to-date representation of the economy. In particular, the SAM methodology utilized existing household survey data (WMS 111 and ILS 1998/99) and also primary data on inputs and outputs collected from firms in the major sectors. The available information from various sources was reconciled into a 'prior SAM'. This process involved construction of a 'macro-SAM' using aggregate information from national accounts and other macroeconomic databases. The SAM was then disaggregated and receipts and payments balanced across all the accounts. The cross entropy econometrics method was used to balance the 'micro-SAM'. The level of disaggregation of the SAM was determined by the general objective of the project, which was to analyze growth options for poverty reduction, with emphasis on agriculture, manufacturing and tourism. Data availability also limited further disaggregation.

The SAM construction process was quite challenging due to data constraints, and also because a long time had elapsed since the construction of the previous SAM in 1976. However, the resulting SAM is a reasonable reflection of the structure of the Kenyan economy in 2003. It is, therefore, hoped that this SAM will be a useful instrument for calibration of CGE models, which are important tools for policy analysis, and will also form an important base for the construction of future SAMs.

References

Gitu, K. W. and Nzuma J., (2003), *Data Compendium for Kenya's Agricultural Sector*. KIPPRA Special Report No. 4, Nairobi: Kenya Institute for Public Policy Research and Analysis.

Government of Kenya (2003), Report of the 1998/99 Labour Force Survey. Nairobi: Central Bureau of Statistics, Ministry of Planning and National Development.

King B. (1985), 'What is a SAM?', in Pyatt, G. and J. Round, *Social Accounting Matrices: A Basis for Planning*. Washington DC: World Bank, Washington.

Pyatt, G. (1988). "A SAM Approach to Modeling." *Journal of Policy Modeling*, Vol. 10.

Pyatt, G. and J. Round (1985). Social Accounting Matrices: A Basis for Planning. Washington DC: World Bank.

Reinert, K. A., and D. W. Roland-Holst (1997). "Social Accounting Matrices," in J. F. Francois, and K. A. Reinert (eds) Applied Methods for Trade Policy Analysis: A Handbook. New York: Cambridge University Press.

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

Robinson S. and D.W. Roland-Holst (1988), "Macroeconomic Structure and Computable General Equilibrium Models." *Journal of Policy Modeling*, Vol. 10.

Sadoulet, E. and A. D. Janvry (1995), 'Quantitative Development Policy Analysis', The John Hopkins University Press, Baltimore and London.

Appendix

Table A1: Mapping between SAM and national accounts sectors

Social Accounting Matrix		National Accounts	
amaiz	Maize	s1	Field crops & agriculture (see Section 3 for disaggregation method)
awhea	Wheat		
arice	Rice		
abarl	Barley		
acott	Cotton		
aogrn	Other cereals		
asugr	Sugarcane		
acoff	Coffee		
atea	Tea		
aroot	Roots & tubers		
aoils	Pulses & oil seeds		
afrui	Fruits		
avege	Vegetables		
acutf	Cut flowers		
aocrp	Others crops		
abeef	Beef	s2	Livestock (see Section 3 for disaggregation method)
adair	Dairy		
apoul	Poultry		
agoat	Sheep, goat and lamb for slaughter		
aoliv	Other livestock		
afish	Fishing	s4	Fishing
afore	Forestry	s3	Forestry and logging
amine	Mining	s5	Mining and quarrying
ameat	Meat & dairy	s6	Meat, fish, fruit, vegetables, oils and fats
amill	Grain milling	s7	Manufacture of dairy products
		s8	Milling products and prepared animal feeds
abake	Sugar & bakery, & confectionary	s10	Manufacture of bakery products
		s11	Manufacture of sugar
abevt	Beverages & tobacco	s9	Manufacture of beverages
		s12	Manufacture of coffee
		s13	Manufacture of tea
		s15	Manufacture of tobacco products
aomfd	Other manufactured food	s14	Manufacture of other food products n.e.c.
atext	Textile & clothing	s16	Manufacture of textiles
		s17	Wearing apparel an

Social Accounting Matrix		National Accounts	
d	dressing and dyeing of fur	s18	Manufacture of leather products and footwear
afoot	Leather & footwear	s19	Products of wood and cork (excl. furniture)
awood	Wood & paper	s20	Manufacture of paper and paper products
aprnt	Printing and publishing	s21	Reproduction of recorded media
apetr	Petroleum	s22	Manufacture of refined petroleum products
achem	Chemicals	s23	Chemicals and chemical products
		s24	Manufacture of rubber and plastics products
amach	Metals and machines	s26	Basic metals and metal products
		s27	Machinery, apparatus and equipment
		s28	Manufacture of transport equipment
anmet	Non metallic products	s25	Non-metallic mineral products
aoman	Other manufactures	s29	Manufacturing n.e.c.
awatr	Water	s31	Purification and distribution of water
aelec	Electricity	s30	Electricity supply
acons	Construction	s32	Construction
atrad	Trade	s33	Wholesale and retail trade (incl. repairs)
ahotl	Hotels	s34	Hotels and restaurants
atran	Transport	s35	Land transport; transport via pipelines
		s36	Water transport; harbours
		s37	Air transport
		s38	Auxiliary transport activities; travel agents
acomm	Communication	s39	Post and telecommunications
afsrv	Finance	s40	Financial intermediation
arest	Real estate	s41	Real estate, renting, and business activities
aosrv	Other services	s45	Social and personal service activities
		s46	Private households with employed persons
aadmn	Public administration	s42	Public administration and defense
aeduc	Education	s43	Education
aheal	Health	s44	Health and social work

Table A2: Selection of districts and mapping between districts and agro-ecological zones

District name	Dominant agro-ecological zones (%)					Questionnaires		Main surveyed crop and livestock enterprises
	UH	LH	UM	LM	L	DAO/DLPO	Farms	
Baringo			10	65	20			
Bungoma			40	40				
Busia				100		10	15	Banana, cotton, sweet potato, cassava, sugarcane, soyabean, local birds, beef, dairy, fodder
Embu			10	70				
Homa bay				95				
Kajiado			5	90				
Kakamega			40	60				
Keiyo, Marakwet	60		10	30		9	16	Maize, beans, wheat, tomato, mango, dairy, beef, local birds, fodder
Kericho local		70	20			9	18	Maize, tea, sugarcane, coffee, tomatoes, beans, bee keeping, birds, dairy, fodder
Kiambu, Thika local		30	50			8	17	Passion fruit, pineapples, tomato, coffee, tea, flowers, broilers, layers, dairy
Kilifi					100	5	16	Cassava, coconut, green grams, cow peas, maize, broilers, layers, dairy
Kirinyaga		10	40	40		7	21	Tea, french beans, banana, coffee, rice, tomato, layers, local birds, dairy
Kisii		30	60			10	18	Maize, tea, banana, coffee, kales, bee keeping, dairy, poultry (local birds), fodder
Kisumu				95				
Kitui				70	25	8	16	Cow peas, maize, pigeon peas, green grams, beans, beef, goats, poultry (local birds)

District name	Dominant agro-ecological zones (%) Questionnaires						Main surveyed crop and livestock enterprises	
	UH	LH	UM	LM	L	DAO/DLPO	Farms	
Kwale					100	8	15	Bixa, oranges, cashew nut, cassava, coconut, beef, dairy, layers
Laikipia	50		40			8	17	Wheat, maize, beans, tomato, tomato, dairy, sheep, beef
Machakos			40	50				
Meru	5	5	30	60		8	17	Banana, maize, french beans, coffee, tea, poultry, dairy, dairy goat
Murang'a, Maragua		20	70					
Nakuru	30	20	40			7	7	Potato, beans, wheat, maize, cabbage, dairy, poultry
Nandi		80	10					
Narok	10	20	70			8	16	Maize, beans, barley, wheat, potato, dairy, beef, shoats
Nyandarua	90	10				8	16	Carrot, cabbage, potato, wheat, maize, dairy, layers, sheep
Nyeri	20	30	30					
Samburu		10		70	10			
Siaya				99				
Taita Taveta				30	50			
Tana River			45		100	8	16	Tomato, banana, mango, maize, rice, goats, sheep, beef
Trans Nzoia	10	45						
Uasin Gishu	30	60	10			7	17	Maize, wheat, passion fruit, irish potato, tomato, dairy, poultry (broiler), bee keeping, fodder
West Pokot			25	40	25			

Note: Dominant agro-ecological zones is the percentage distribution of land in the upper highlands (UH), lower highlands (LH), upper midlands (UM), lower midlands (LM), and interior and coastal lowlands (L); Questionnaires is the number of questionnaires completed by District Agricultural Officers (DAO) and District Livestock Production Officers (DLPO) and by district farmers.