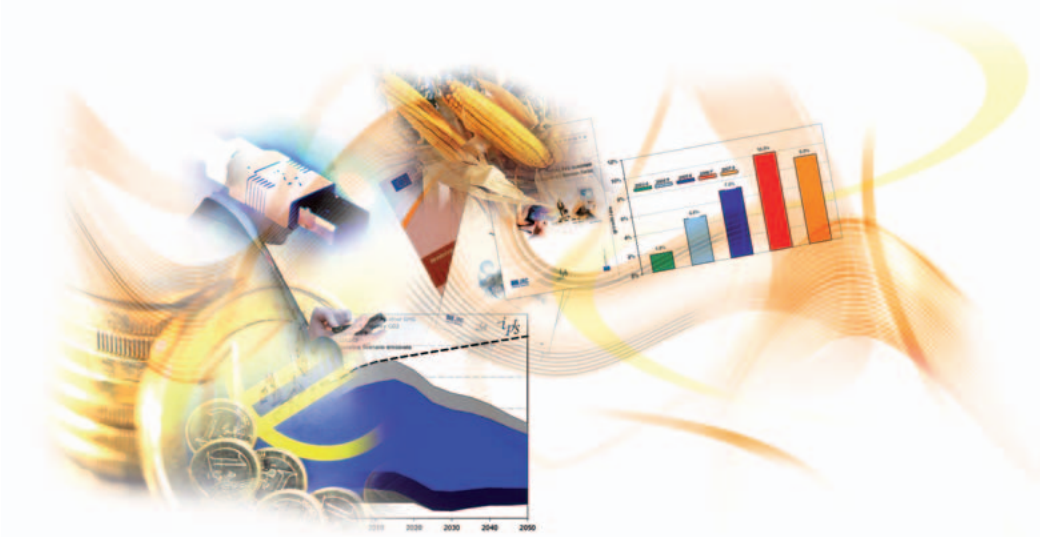
JRC Scientific and Technical Reports

Construction of Social Accounting Matrices for the EU-27 with a Disaggregated Agricultural sector (AgroSAM)

Authors: Marc Müller, Ignacio Pérez Domínguez, Stephan Hubertus Gay



EUR 24010 EN - 2009

JRC
EUROPEAN COMISSION

ipts
Institute for Prospective Technological Studies

Institute for Prospective Technological Studies

Construction of Social Accounting Matrices for the EU-27 with a Disaggregated Agricultural Sector (AgroSAM)

Authors:
Marc Müller, Ignacio Pérez Domínguez, Stephan Hubertus Gay

2009

JRC
EUROPEAN COMISSION

EUR 24010 EN

The mission of the JRC-IPTS is to provide customer-driven support to the EU policy-making process by developing science-based responses to policy challenges that have both a socio-economic as well as a scientific/technological dimension.

European Commission
Joint Research Centre
Institute for Prospective Technological Studies

Contact information
Address: Edificio Expo. c/ Inca Garcilaso, 3. E-41092 Seville (Spain)
E-mail: jrc-ipts-secretariat@ec.europa.eu
Tel.: +34 954488318
Fax: +34 954488300

http://ipts.jrc.ec.europa.eu
http://www.jrc.ec.europa.eu

JRC 53558

EUR 24010 EN
ISBN 978-92-79-13625-2
ISSN 1018-5593
DOI 10.2791/29021

Luxembourg: Office for Official Publications of the European Communities

Printed in Spain

Preface by the Authors

This study was carried out from December 2006 to March 2009 by a research group at the "Institute for Prospective Technological Studies" (IPTS), one of the seven scientific institutes of the European Commission's "Joint Research Centre" (JRC). It originated from a formal request by the Directorate General for External Trade (DG Trade) to update the input-output tables of the widely used database of the "Global Trade Analysis Project" (GTAP). Additionally, a modelling platform at IPTS, comprising several partial and general equilibrium models, was launched at this time. In this context, it became evident that the value-added of the modelling platform could be significantly increased if the different analytical tools were not only utilized individually but also built on a common database. Due to the fact that the application of general equilibrium models for analysis of the Common Agricultural Policy of the EU (CAP) is often limited by the coarse representation of the agricultural sector in national economy-wide databases, it was decided to work on the harmonization of in-house model databases combining national Supply- and Use tables from EuroStat with the database of the "Common Agricultural Policy Regional Impact" modelling system (CAPRI) at Member State level. This should allow performing general equilibrium modelling work with a highly disaggregated and consistent agricultural sector.

Already at the early stages of this project, it became clear that the available economy-wide datasets needed additional treatment to fulfil the requirements of consistency and completeness. Additionally, progress was slowed down by systematic definitional and structural deviations between the databases to be combined. As a result, a comprehensive sequence of data compilation steps had to be developed, which raised at some point the question whether it would have been faster to compile each country database individually and manually. However, it was concluded that a generic computational approach based on widely accessible databases was preferable as this would allow in the future for updates and incorporation of additional information whenever available. Therefore, the final output of this project – Social Accounting Matrices with a disaggregated agricultural sector (AgroSAM) – cannot be regarded as a fixed database but rather a dynamic product that relies heavily on constant cross-checks and feedbacks from interested user groups. Since the delivery of the first AgroSAM model version for 21 countries to the GTAP network in summer 2008, the compilation procedure has undergone numerous changes and updates. The model version presented here, and which will be made available to the public, will be therefore subject to further improvements through its application in different project consortia.

This process would not have been possible without the feedback from many external experts who checked the data and pointed to major flaws of the earlier AgroSAM model versions. The most relevant input came from an advisory group consisting of Terrie Walmsley, Martin Banse, and Scott McDonald, but the authors are also grateful to many other researchers who spent their time and efforts to provide valuable feedback.

Furthermore, we would like to encourage all interested users to send us comments and, most importantly, provide us with additional information and datasets to improve further releases of the AgroSAMs.

Preface by the Advisory Group

Empirical economic models are critically dependent on the quality of the data used by the models, but the time and energy (resources) available for the development of the requisite databases are seriously limited while all too often project funders are loath to fund data collection. Despite the limitations on resources for database compilation policy makers are placing evermore emphasis on ex ante economic analyses that support economic policy decisions. The scope of such analyses is ever widening; inter alia simulation models have been used to evaluate the implications of farm support programmes, e.g., the Common Agricultural Policy, bilateral trade agreements, e.g., Economic Partnership Agreements, multilateral trade policies, e.g., the Doha Development Agenda, energy policies, e.g., 'cap and trade' regimes, and integrated (climate) assessment models, e.g., Kyoto protocol and climate change policies. At various times there have been attempts to address the requirements of data intensive economic models; input-output databases, building on Leontief's work address issues relating to inter-industry transactions, and agricultural commodity databases, for instance FAPRI's development of linked databases and models to address issues relating to agricultural commodity markets. But attempts to develop integrated economic accounts are much less common.

The development of national accounts was largely driven by a need to give empirical content to the Keynesian economic model after the Second World War. Among the results of these efforts is the United Nations (UN) "System of National Accounts" (SNA) whose latest manifestation came out in 1993. One substantial development of the SNA between 1968 and 1993 is the prominence now given to Social Accounting Matrices (SAMs) and associated satellite accounts. SAMs are integrated representations of national accounts that are complete and consistent; complete in the sense that all economic transactions are recorded and consistent in that every expenditure transaction is matched by an identical income transaction. It is these characteristics of SAMs, together with the richness of the institutional and social information they can contain, that have made them attractive to economists as the databases for the calibration of whole economy economic models. But large amounts of time and effort are required even to develop one SAM for a single economy; a fact demonstrated by the fact that SAMs are regarded as an optional component of the SNA and that the production of SAMs, even by national accounts agencies, are, at best, irregular activities.

This research report is concerned not with the development of one SAM for one economy but rather the development of SAMs for all members of the European Union (EU). There are four distinctive features of these SAMs. First, the accounts for each SAM are identical, which means that they can be used to provide directly comparable structural information for each economy. Second, the tax accounts have been disaggregated so that the implications of changes in different tax (and subsidy) instruments can be systematically analysed. Third, the accounts for agriculture have been disaggregated so that the implications of international and domestic agricultural policies can be disentangled. And fourth, the account structure has been designed so that this integrated EU database can be used to provide consistent data for the "Global Trade Analysis Project" (GTAP). As such this research represents an important and very valuable resource for economic analyses.

The availability of a standardised series of SAMs for all EU economies has numerous potential benefits. At the simplest level a set of standardised SAMs means that a single (standardised) model can be used for all EU economies, thereby saving on costs and providing results that can evaluate the differential impacts

of common policies on all members. Moreover, because of the efforts devoted to the tax accounts the models using the SAMs can provide a much richer simulation environment for fiscal policy analyses; again something that is important as the EU seeks to move towards a series of increasingly integrated economic policies. As such the SAMs provide a valuable and cost efficient set of databases for the analyses of conventional economic policy issues – trade policies, sectoral (including agricultural) policies, fiscal policies, etc. – in a way that allows immediate and direct comparisons of the implications for the different members of the EU.

While conventional economic policy issues remain important, they are arguably increasingly less important than some emerging issues. Many economic models used to evaluate inter alia the implications of energy use, bio fuel programmes, 'cap and trade' schemes for reducing carbon emissions, land-use changes, labour migration, etc., rely on SAMs to provide the core economic transactions data. As such these SAMs provide an important resource for the development of economic models that can be used to analyse a range of emerging policy questions. While the current versions of the SAMs will require augmenting to address many of these issues most of the additional data requirements can be stored as satellite accounts, e.g., the SNA's environmental satellite accounts, that can be grafted onto the economic (transactions) core of the SAMs.

The disaggregation of the agricultural accounts provides the basis for coordinated analyses of EU and international agricultural policies. The linking of the whole-economy (SAM) data with the commodity specific CAPRI data allow for a better understanding of intermediate use of agricultural commodities within agriculture but also for an improved presentation of flows of intermediates between primary agriculture and food processing. The work in this study will also contribute to data improvements in the area of biomass and biofuel. Currently quantitative research based on the GTAP data base requires some 'ad hoc' data adjustments to capture the increasing demand of biomass in the energy and bio-refinery sectors. Linking the CAPRI data base with the GTAP data base by using these SAMs will help to update the GTAP data base for the biobased sectors in a more transparent way.

The synchronisation of the account structure with that of the GTAP database means that they can also contribute to the development of the GTAP database. This means that the EU SAMs are included with 86 other countries/regions in a globally consistent database. This database can be used in conjunction with the GTAP or an alternative global model (such as GTAPinGams or GLOBE) to examine the impact of EU or global trade and environmental issues on the production, trade and welfare of EU and non-EU countries. Furthermore these models can be linked to other models, such as the CAPRI model, to obtain detailed information on the impact of a policy on EU agriculture, while taking into account interactions with other economies.

As the authors recognise this is work in progress, but then all economic databases are, to a greater or lesser extent, works in progress. At a trivial level there are refinements to the data estimates included in the existing structure; some involve 'improvements' in estimation techniques but others require improvements in the underlying national economic data. The authors indicate many of these refinements and, by implication, issue a challenge to national account statisticians in the EU. There are obvious aspects of the SAMs that would benefit from further research, e.g., the extension of the factor accounts to include land and different types of labour and more types of (private) household, but these developments will require the authors to address large gaps in underlying economic data for many EU members. Similarly the development of environmental and demographic satellite accounts that integrate with the economic

transactions data has hardly started, in part because of the remit of this project but also in part because of the paucity of national level data relating to environmental and demographic questions. Consequently an especially important contribution of this research is the implicit challenge it issues to both national account statisticians and economic modellers. But a comprehensive and consistent set of SAMs for all EU economies now exists; hence it is incumbent upon statisticians and modellers to 'raise their game' by enhancing the quality of the economic and satellite account data and using them for economic analyses.

Dr Martin Banse, *Agricultural Economics Research Institute (LEI), Netherlands*.
Professor Scott McDonald*, Oxford Brookes University, Oxford, UK*.
Dr Terrie Walmsley, *Center for Global Economic Analysis, Purdue University, USA*.

Executive Summary

Integrated policy impact assessment at pan-European or global scale requires large-scale consolidated databases to feed economic or bio-physical models or components. A key data set for economic analysis are Social Accounting Matrices (SAM) which represent the monetary flows between productive sectors and institutions and thus may serve a large variety of quantitative tools, especially Computable General Equilibrium (CGE) models. However, the datasets underlying the SAMs, namely national Supply- and Use-Tables (SUT) or symmetric Input-Output tables (IOT) are typically highly aggregated by sectors and commodities and thus provide little detail for sub-sector specific analysis. The agricultural sector is e.g. often represented as one row and one column in the national datasets. This coarse representation is an important reason for the limited application of CGEs for analysis of the Common Agricultural Policy.

The AgroSAM project hosted at the Institute for Prospective Technological Studies of the European Commission (IPTS) addresses this issue by combining national SUT for the EU Member States with the highly disaggregated information on the agricultural sector provided by the "Common Agricultural Policy Regionalised Impact" (CAPRI) model database. The project also aimed at providing an updated version of the EU-27 IOT for the "Global Trade Analysis Project" (GTAP) database, which was asked for by DG TRADE.

One of the main challenges of compiling EU-27 SAMs with a disaggregated agricultural sector (AgroSAMs) consists in overcoming definitional and structural differences between the SUT based on the European System of National Accounts (ESA95) and the CAPRI database which is mainly structured according to the Economic Accounts for Agriculture (EAA). As such, the AgroSAM project is one example for constructing large-scale databases for impact assessment where different data sources are combined and consolidated.

This report provides an overview on the datasets used and methods applied to construct the consolidated AgroSAMs. Particularly the conceptual and definitional differences between the combined datasets, namely the CAPRI database and the SUT provided by EuroStat are addressed. As a consequence, a wide range of methods for the estimation of balanced SAMs, its sub-matrices, and related control-totals had to be applied. In principle, the procedure comprised the following steps: first, a full set of macroeconomic indicators was collected and arranged in the format of macroeconomic accounting matrices or macroeconomic SAMs (step 1). Next, SUT and data on monetary flows between domestic and foreign institutions like taxes and income transfers were used to create a set of institutional SAMs following the ESA95 classifications (step 2), where SAMs were balanced with respect to the macroeconomic totals originating from the previous step. In step 3 detailed data for the agricultural sector from the CAPRI model were mapped into a comparable SAM format and combined with the ESA95-based SAMs into an unbalanced a priori estimate for the final AgroSAMs. Again, these a priori SAMs were balanced with respect to the corresponding entries of the SAMs in ESA95 format.

The balancing procedures to be applied throughout the different compilation steps had to account for the specific needs of the problems at hand, namely to allow expressing confidence in certain datasets, to accommodate control-totals, to permit negative entries, and to ensure preservation of signs. To identify a most suitable approach, a variety of SAM balancing procedures and assumptions concerning the nature of distortions in the used datasets were compared. Although it soon became clear that there is no single preferable approach for the given problem, a Generalised Cross Entropy procedure based on the actual

values in the SAMs (rather than SAM-coefficients), combined with the formulation of multiplicative error terms, appeared to be the most adequate procedure. The fact that multiplicative error terms may result in sometimes huge deviations from the originally observed values was addressed by using different, and sometimes comparatively small, variances for some of the used datasets.

To ensure the scientific quality of the project, an advisory group of three international experts in the field of SAM and IOT compilation and CGE modelling was asked to provide scientific guidance and support. As one of the objectives of the AgroSAM project was to contribute to the database of the GTAP, it was most helpful that the advisory group members were highly involved in this project. A set of 14 national IOT has been contributed to the GTAP v7 database in the summer 2008. Furthermore, the full set of EU-27 AgroSAM has been submitted to GTAP in summer 2009, in will be included in the new release after internal evaluation.

Apart from the contribution to the GTAP network, the AgroSAM project is currently involved in two projects of the 7th framework programme, namely the CAPRI-RD and AgFoodTrade projects. Within these projects, the procedures to compile the AgroSAMs will be further refined, the database will be brought to a more recent year, and it is foreseen to use the AgroSAMs as a starting point for the compilation of a regionalised database in Europe. Therefore, the methodologies presented in this report represent the current state of a highly dynamic process, and will be subject to further change, refinement, and improvement. Consequently, this report should be read as a starting point of a process that aims at the creation of a consolidated database for a variety of partial or general equilibrium models that are hosted at the iMAP modelling platform at IPTS.

Table of Contents

List of figures

List of tables

1. Background and Motivation for this Study

Integrated policy impact assessment at Pan-European or global scale requires large-scale consolidated databases to feed economic or bio-physical models or components. A key data set for economic analysis are Social Accounting Matrices (SAM, see Pyatt and Round (1985)) which represent the monetary flows between productive sectors and institutions and, thus, may serve a large variety of quantitative tools, especially Computable General Equilibrium (CGE) models. However, the datasets underlying the SAMs, namely national Supply- and Use Tables (SUT) or symmetric Input-Output tables (IOT), are typically highly aggregated by sectors and commodities and, thus, provide little detail for sub-sector specific analysis. The agricultural sector is e.g. often represented as one row and column only in the national datasets.

This coarse representation is one reason for the limited application of CGEs for analysis of the Common Agricultural Policy. The AgroSAM project hosted at the Institute for Prospective Technological Studies of the European Commission (IPTS) addresses this issue by combining national SUT for the EU Member States with the highly disaggregated information on the agricultural sector provided by the database of the "Common Agricultural Policy Regional Impact" modelling system (CAPRI) (Britz and Witzke, (2008)). One of the main challenges for AgroSAM consists in overcoming definitional and structural differences between the SUT based on the European System of National Accounts (ESA95) and the CAPRI database which is mainly structured according to the Economic Accounts for Agriculture (EAA). As such, the AgroSAM project is one example for constructing large-scale data bases for impact assessment where different data sources are combined and consolidated.

The main output of this project is the construction of AgroSAMs for the EU-27 Member States which allow analysing the economic effects of the CAP reform within and beyond agriculture: The AgroSAM constitute a database that is sufficiently large and detailed to serve as main input for general equilibrium models like the "Global Trade Analysis Project's" model (GTAP) or the GLOBE model (McDonald et al, (2007)), which are part of the model portfolio of IPTS.

With this objective in mind, the following aspects are highlighted:

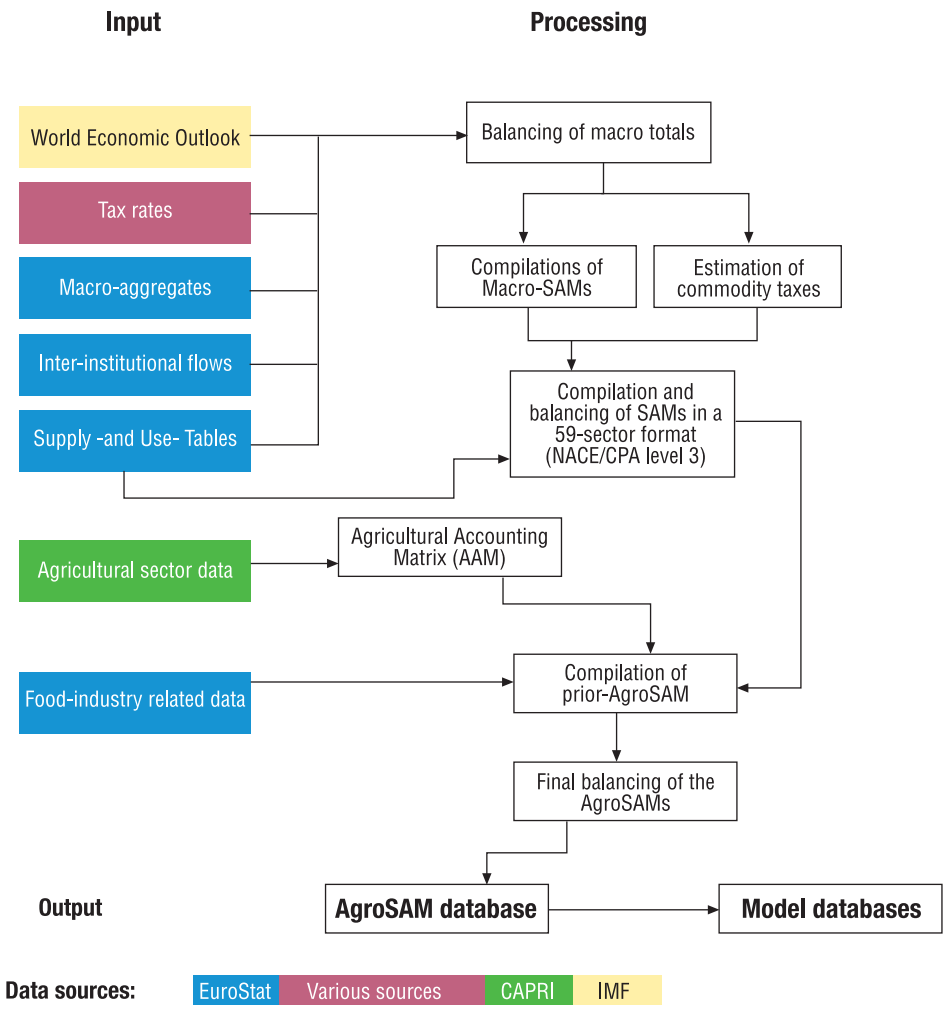
a) AgroSAMs should allow an adequate analysis of agricultural policies. Taxes and subsidies on commodities, production activities, and enterprises should be included in as much detail as possible.

b) The number of agricultural sub-sectors should allow:
   - the incorporation of datasets from already existing economic models (e.g. CAPRI);
   - the reusability by other modelling systems (e.g. GTAP, GLOBE);
   - the utilisation of readily available datasets from statistical departments (e.g. EuroStat, FAOSTAT).

c) A transparent and automatised routine should allow for the extraction, transformation and incorporation of new datasets, so that the update costs of the AgroSAMs are kept at a low level.

This technical documentation follows the compilation-sequence of the AgroSAMs as outlined in Figure 1. Firstly, the desired properties and structure of the final database are highlighted in section 2. The used datasets are described in section 3, giving particular attention to the data sources and their location on the sites of the respective providers. Section 4 provides a brief methodological discussion on the general approach used here to consolidate information drawn from different sources. Subsequently, the necessary steps to compile, adjust, and balance the used datasets are described, beginning with the compilation of macroeconomic indicators which will serve as control totals at later stages (section 5). As building ground for model databases, the final AgroSAMs should permit the distinction of different types of taxes on commodities (value-added taxes (VAT), excise taxes, import duties). Since this information is not available in the required detail and format, but nevertheless relevant for the formulation of policy scenarios, its estimation was deemed necessary. The applied method is discussed in section 6.2. After estimating the tax rates, SAMs are compiled according to the ESA95 classification scheme by re-arranging SUT and data on monetary flows between institutional sectors into a SAM format (ESASAMs, section 6), and balancing them based on and subject to the intermediate results from the previous sections. The balanced ESASAMs are then used, together with detailed agricultural sector data, to compile a prior dataset (section 7), which is then again balanced with a second cross-entropy procedure in section 8. Section 9 summarises the achievements and addresses still unresolved problems.

Figure 1. AgroSAM Compilation Flow



2. Target Structure of the AgroSAMs

The structure of the AgroSAMs is largely determined by the available data and the desired compatibility with the classifications used in prominent modelling systems, namely the "Global Trade Analysis Project" (GTAP)[1] and the "Common Agricultural Policy Regional Impact" (CAPRI)[2] models. The "Complete and Consistent Database" of CAPRI (COCO, see Britz and Witzke (2008)), which is shared also by the CAPSIM modelling system, distinguishes 50+ agricultural production activities and output commodities, 30+ agricultural inputs, and 20+ processed commodities. This representation of the agricultural sector determines the upper limit for the level of disaggregation in the target AgroSAMs, as more detailed datasets with the same country-coverage were not available for this study. Apart from this, the CAPRI database is constantly maintained and updated, and the underlying expert knowledge ensures the quality of the included information.

The dominant role of GTAP in the context of policy analyses on global scale gives raise to the consideration that a set of AgroSAMs for EU27 should be compatible with the GTAP classification scheme, such that the creation of datasets usable in the GTAP framework, namely symmetric IOT at basic prices, is possible without extensive additional work. Therefore, the GTAP classification scheme, which distinguishes 12 raw agricultural products and 8 processed food commodities, was set as the lower limit for the level of detail for the agricultural sector. This includes the requirement that the target classification of the AgroSAMs can be mapped into the GTAP classification by simple summation (many-to-one mapping). As the latter requirement is not necessarily fulfilled by the standard classification schemes in which the crucial SUT are provided, the formulation of a "Modified Agro-industrial Classification" (MAC) is pursued, which follows in general the commodity classification of the "Combined Nomenclature" (CN) and the "Commodities Produced by Activities" (CPA) used by EuroStat, but lies within the bounds given by CAPRI and GTAP. The target classification and the correspondence with other models are documented Appendix 1.

2.1 Commodities and Activities

The structure of the target classification MAC is largely determined by the ESA classifications "Nomenclature for Economic Activities" (NACE) and CPA at three-digit level, in which the SUT are provided. Agriculture and food-industry are the exceptions, since a more detailed representation of these two sectors is aimed at. The highest possible level of detail that can be achieved is the one provided by the CAPRI database and the lowest should be provided by GTAP, in order to allow for a correspondence to its classification. A further desirable property of the target classification is a correspondence of the activities with the most refined NACE classification level (5 digits).

1 https://www.gtap.agecon.purdue.edu/
2 URL: http://www.capri-model.org/

2.2 Primary Factors

The requirement that AgroSAMs should be transferable into the GTAP classification also includes the representation of primary factors. GTAP distinguishes here between land, skilled and unskilled labour, capital, and natural resources. However, the available use tables do not provide the detail of information as used within GTAP, but allow only to distinguish between "compensation of employees", "consumption of fixed capital" and "net operating surplus". The representation of factor markets is one of the major assets of CGE models, and an underlying database should consequently support modelling activities in this area. Unfortunately, additional information to separate compensation of skilled and unskilled labour, or payments for "natural resources" was not available at this stage of the AgroSAM project. Since the CAPRI model is a partial equilibrium model and its database is not designed for a detailed analysis of agricultural factor markets it was also not possible to obtain the necessary details from this source.

2.3 Taxes and Institutions

Monetary flows between domestic and foreign institutions are a crucial feature of SAM-representations of national economies. These vary from direct transfers from public budgets to enterprises or households, to the current account of the balance of payments and the savings-investment accounts. Again, the taxes and institutions in the GTAP framework constitute the lowest level of detail for the AgroSAMs. This means that at least one SAM account for each of the domestic institutions (national government, aggregate household and aggregate enterprises) should be included in the SAM. The savings-investment account should be split into a "fixed capital formation" and a "stock changes" account. External trade accounts should allow for a distinction between trade partners. The latter is not supported by the SUT which distinguish only intra-EU and extra-EU trade. Furthermore, the CAPRI database considered only features net trade, with no distinction of origin or destination. Due to this lack of data, a preliminary solution was chosen and external trade represented by a single account. In future versions of the AgroSAMs, it will be desirable to include bilateral trade matrices.

Compatibility with the GTAP modelling system requires that at least taxes or subsidies on products and production, on imports and exports, and value-added type taxes are listed separately. Direct taxes and transfers are needed to complete the SAM accounting system with respect to the monetary flows between institutions, which is supported by the availability of the respective datasets from EuroStat.

3. Datasets

The datasets used for the compilation of highly aggregated macroeconomic SAMs (MacroSAMs) are described in this section. Main sources were EuroStat and the International Monetary Fund (IMF). The final AgroSAMs are mainly based on the SUT (NAIO in the following), the Annual Sector Accounts (NASA), and the Macro Aggregate Indicators (NAMA), but it appeared that there were substantial deviations across the various datasets, such that a comparison and consolidation of macro-totals appeared to be necessary. Table 1 gives an overview on the datasets used and their latest download. The subsequent sections provide a more detailed discussion on the contents and location of the files used.[3]

Table 1. Downloaded Files

Source	Code	File name(s)	Last download	Comments
EuroStat	NAIO	naio_cp15.tsv	18.09.2008	SUT at NACE/CPA 3-digit level (59 sectors)
EuroStat	NAIO	naio_cp16.tsv	18.09.2008	
EuroStat	NAMA	nama_aux_pem.tsv	10.09.2008	Auxiliary indicators
EuroStat	NAMA	nama_exi_c.tsv	10.09.2008	External trade relations
EuroStat	NAMA	nama_fcs_c.tsv	10.09.2008	Final consumption aggregates
EuroStat	NAMA	nama_gdp_k.tsv	10.09.2008	GDP and main components, constant prices
EuroStat	NAMA	nama_gdp_c.tsv	10.09.2008	GDP and main components, current prices
EuroStat	NAMA	nama_inc_c.tsv	10.09.2008	Income, saving, and net lending
EuroStat	NAMA	nama_nace31_c.tsv	25.06.2008	National accounts by 31 branches
EuroStat	GOV_A	gov_a_tax_ag.tsv	18.09.2008	Main national accounts, tax aggregates
EuroStat	NASA	nasa_simplif.tsv	11.09.2008	Non-financial annual sector accounts
IMF	WEO	WEOApr2008all.xls	28.08.2008	World Economic Outlook Indicators

3.1 Annual National Accounts from EuroStat

The annual national accounts (NAMA) provide a comprehensive amount of macroeconomic indicators related to GDP composition calculated from the expenditure and income sides, plus a large amount of auxiliary indicators like employment and population statistics. Furthermore, it is possible to obtain supply-side information for 31 sectors (e.g. compensation of employees, intermediate demand, and gross output). Unfortunately, the information is spread across various files as indicated in Figure 2.

3 Please note that it is necessary to create a user account on the EuroStat homepage to retrieve the bulk datasets and to have the information appearing as presented in the figures.

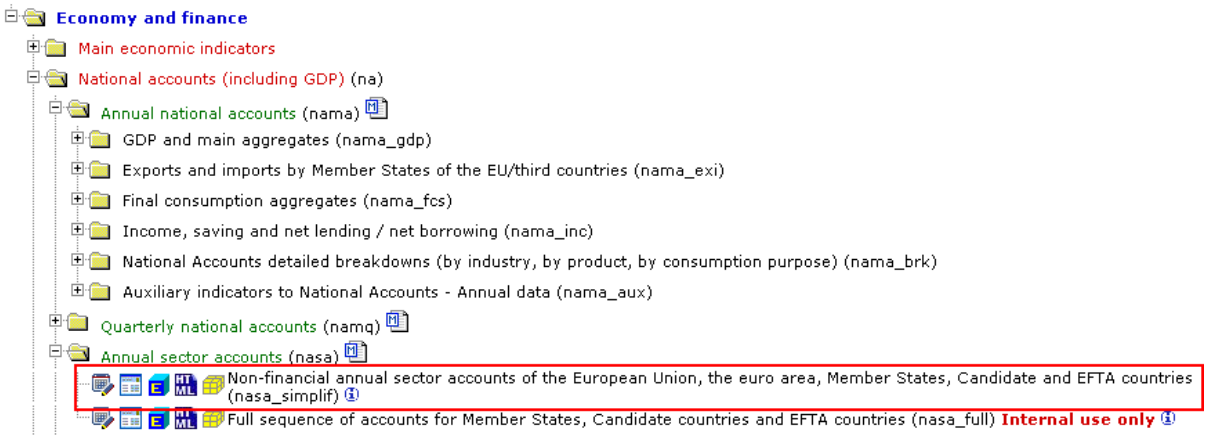
Figure 2. Annual National Accounts: NAMA

Source: Screenshot from the EuroStat homepage:
URL: http://epp.eurostat.ec.europa.eu/portal/page?_pageid=0,1136173,0_45570701&_dad=portal&_schema=PORTAL

3.2 Annual Sector Accounts from EuroStat

The annual sector accounts (NASA) contain in principle the same information as the annual national accounts, but provide additional data on flows between sectors, domestic institutions, and the 'Rest of the World'. In this respect, NASA is the only source for e.g. factor incomes from abroad, transfers received by households and direct taxes paid by enterprises and households. Furthermore, all information is available in one file (nasa_simplif), which contains a simplified version of the full sequence of monetary flows, which is sufficient for this project.

Figure 3. Annual Sector Accounts: NASA



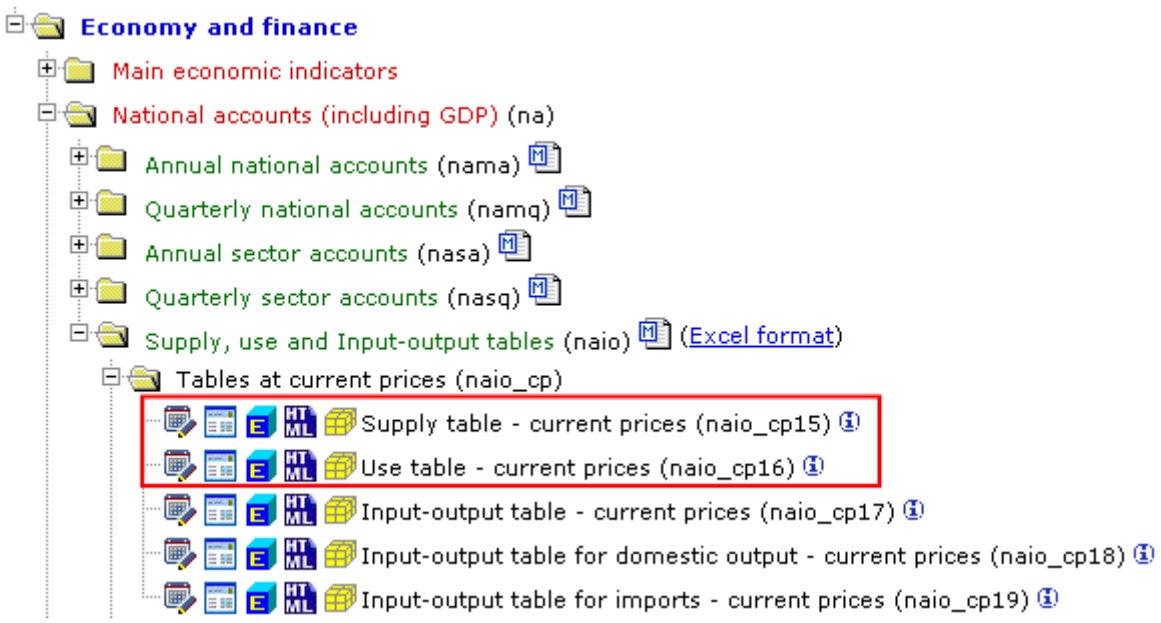
Source: Screenshot from EuroStat homepage:
URL: http://epp.eurostat.ec.europa.eu/portal/page?_pageid=0,1136173,0_45570701&_dad=portal&_schema=PORTAL

3.3 Supply, Use, and Input-Output Tables from EuroStat

SUT are the most relevant database in the context of SAM construction as they represent the full flow of goods and services within an economy and provide also information on trade margins and sector-specific taxes and subsidies. EuroStat provides SUT in two different file-formats:

- bulk download in 'tsv' format
- country-wise downloadable MS-Excel files

Figure 4. Supply, Use and Input-Output Tables: NAIO



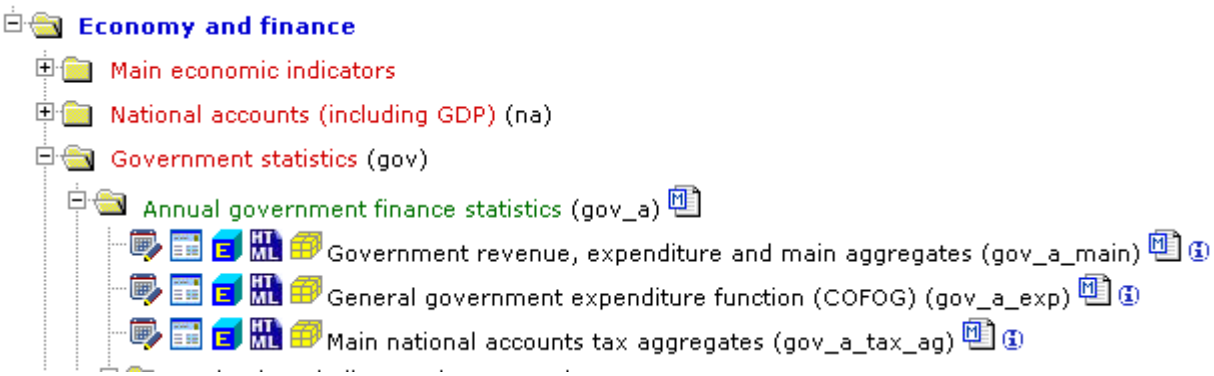
Source: Screenshot from EuroStat homepage:
URL: http://epp.eurostat.ec.europa.eu/portal/page?_pageid=0,1136173,0_45570701&_dad=portal&_schema=PORTAL

Symmetric IOT in basic prices are also provided by the same sources as the SUT and in the same formats and coverage. Although IOT are often used as input for the compilation of SAMs, they do not provide the same amount of information as SUT, particularly since the transformation of basic prices into purchaser prices (e.g. trade margins and taxes on commodities) is missing. In general, SUT are preferable for SAM construction. The NAMA datasets have a wider coverage of the EU-27 Member States and also include main indicators like total intermediate demand, compensation of employees, gross value added and net taxes on production, but only for 31 sectors and not the 59 which are featured in the SUT framework (NACE, CPA 3-digit level). However, a detailed breakdown of commodity uses (final consumption, exports, imports, investment demand) is only available from the SUT.

3.4 Annual Government Finance Statistics from EuroStat

While the aforementioned datasets contain already some information on (net) taxation of commodities, a detailed breakdown into types of taxes (VAT, excise taxes, import tariffs) is not always available. Therefore, the governmental finance statistics are used as supplementary source for this type of information.

Figure 5. Annual Government Finance Statistics: GOV_A

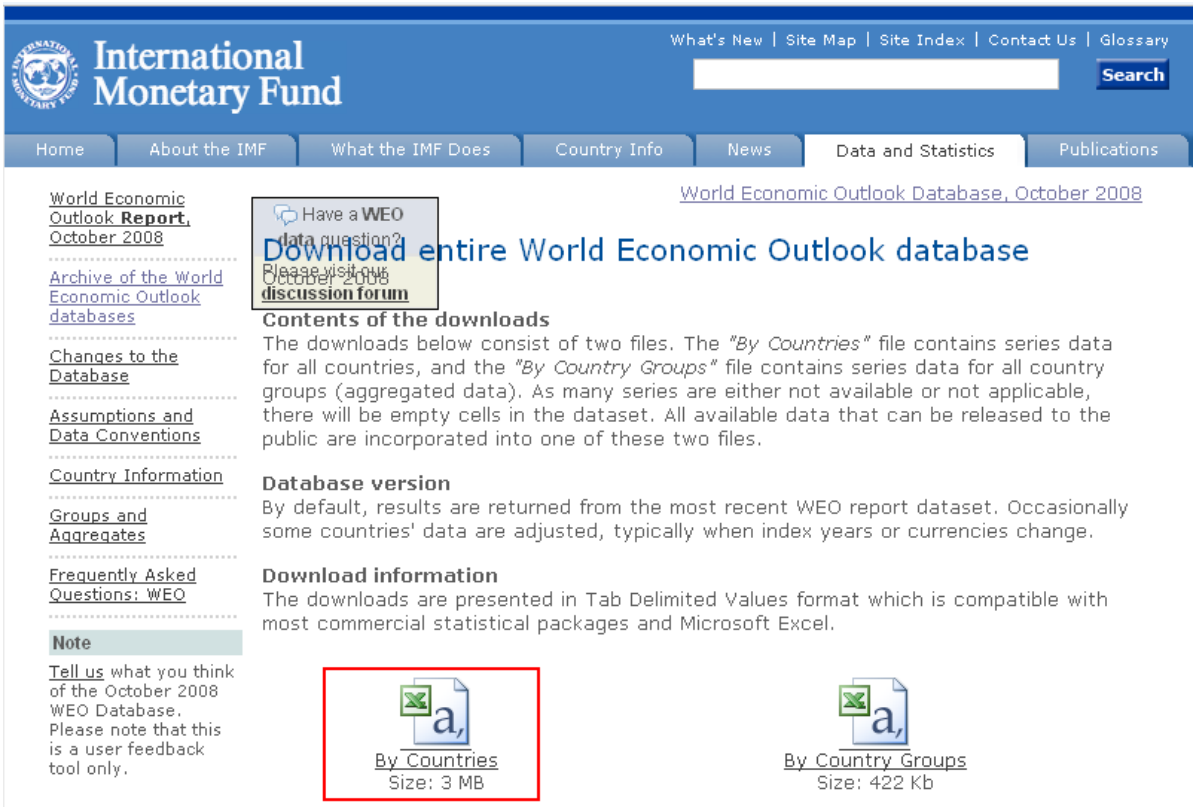


Source: Screenshot from EuroStat homepage:
URL: http://epp.eurostat.ec.europa.eu/portal/page?_pageid=0,1136173,0_45570701&_dad=portal&_schema=PORTAL

3.5 World Economic Outlook from IMF

Although the NAMA and NASA datasets have a wide coverage in terms of years and countries included, it was necessary to rely in some cases on additional datasets (e.g. for calculating GDP at market prices for Romania from 1995 to 1997). With this purpose, the World Economic Outlook (WEO) data were used as reference for the current account balances in case NAMA and NASA data were missing or ambiguous.

Figure 6. World Economic Outlook: WEO



Source: Screenshot from IMF homepage: URL: http://www.imf.org/external/pubs/ft/weo/2008/02/weodata/download.aspx

3.6 Tax and Tariff Rates

Information about tax rates was collected from various sources: VAT rates by commodity groups were obtained from the "European Commission's Directorate General for Taxation and Customs Union" (DG TAXUD 2005) and EuroStat (2008). Import tariff rates by commodity group were obtained from WTO (2007). The considered tax rates are explained in more detail in section 6.2.

3.7 CAPRI Database

The agricultural sector models CAPRI and CAPSIM are both based on a common database (CAPRI) which was developed at the University of Bonn and is the successor of the formerly used SPEL database. Both models and the database are currently available at IPTS (AGRILIFE Unit) and provide a comprehensive picture of the agricultural sector for the EU-27 Member States plus the Balkans. The main data sources for the construction of CAPRI are presented in the following table.

For the purposes of the AgroSAM project, CAPRI is too detailed and includes several elements which are conceptually challenging concerning its transformation into a SAM format (e.g. data on manure production/use, fertilizer consumption, set-aside, milk quotas, activity and product premiums). This has to do with the fact that (1) CAPRI does not strictly follow the "activity from/to commodity" book-keeping structure of ESA (see section 5, "compilation of priors") and (2) it does not consider other sectors of the

Table 2. Data Items and their Main Sources in CAPRI

Data items	Source
Activity levels	Land use statistics, herd size statistics, slaughtering statistics, statistics on import and export of live animals
Production	Farm and market balance statistics, crop production statistics, slaughtering statistics, statistics on import and export of live animals
Farm and market balance positions	Farm and market balance statistics
Sectoral revenues and costs	Economic Accounts for Agriculture (EAA)
Prices	Derived from production and EAA
Output coefficients	Derived from production and activity levels, engineering knowledge
Input coefficients	Different types of estimators, engineering functions
Activity specific income indicators	Derived from input and output coefficients and prices
Policy data	Various sources (Official Journal of the EU)

Source: EuroStat (http://epp.eurostat.cec.eu.int), several bio-physical econometric studies and European Commission (http://publications.eu.int/general/oj_en.html).

economy (e.g. processing of agricultural products like dairies are presented as end-of-pipe products, with no corresponding industrial activities to pay for). Moreover, the CAPRI database includes algorithms for data consistency and completeness, which are key issues to pick up in the AgroSAM project (see Britz and Witzke (2008), pp.15-30).

The combination of the SUT and CAPRI is in fact the major challenge of the AgroSAM project. Eventually, other estimation modules of CAPRI might be picked up in later stages of the project (e.g. estimation of labour and energy inputs, barriers to trade between the EU and the Rest of the World, land prices and quota rents for sugar/milk, etc.). The correspondence between the activity and commodity classifications in CAPRI and the modified agro-industrial classification (MAC) are shown in Appendix 1, Table 15 and Table 16.

3.8 Other Datasets with specific Focus on Agriculture

The "Economic Accounts for Agriculture" (EAA) is a rather extensive dataset for the agricultural sector of the EU-27 and the main input for the CAPRI database. The "Agricultural Information System" (AGR_IS) is only used in the analysis when no information from the EAA is available (e.g. gross trade of agricultural commodities).

FAOSTAT owns the most comprehensive database on trade of agricultural commodities and inputs. This information is important to determine the import/export sub-matrices of the SAM. Since it is already used by the CAPRI model in its market module, the product definitions are consistent with the ones found in the CAPRI database (e.g. trade of wheat measured in terms of 'raw equivalents' found in processed products like beer or pasta). Moreover, the "Agricultural Market Access Database" (AMAD) has very detailed information on market policy instruments (e.g. import tariffs or tariff rate quotas), the OECD provides information on consumer/supply support equivalents (CSE/PSE) for different world aggregates and the World Bank periodically publishes population statistics.

4. A General Methodological Note

The core problem addressed in this study is to compile a (i,j)-dimensional square matrix **S** that includes all monetary transactions between productive sectors, commodity and factor markets, and institutions within a national or regional economy for a given period. The revenues of each account are represented row-wise and the expenditures column-wise. Depending of the structure of the targeted matrix, two basic settings may be distinguished (Round, 2003): first, rows and columns of **S** add up to known, and potentially different, row- and column-totals (denoted $\mathbf{r^0}$ and $\mathbf{c^0}$):

(1) $$\sum_i S_{i,j} = c_j^0; \ \sum_j S_{i,j} = r_i^0$$

Alternatively, if row- and column-sums are unknown, but it is required that they are equal, then the problem may be described formalised as:

(2) $$\sum_i S_{i,j} = \sum_i S_{j,i}$$

Furthermore, the targeted matrix **S** has to be consistent with control-totals Γ (e.g. macro-economic indicators like GDP at market prices, or a more aggregated but balanced SAM), which correspond to certain sub-totals of **S**.

(3) $$\sum_j \left[\sum_i G_{k,i} S_{i,j} \right] G_{j,l} = \Gamma_{k,l}$$

Where **G** is an aggregator matrix mapping the elements of **S** into the corresponding control totals.

Round (2003) also states in this context that **S** has to be non-negative. This assumption is crucial for some of the proposed SAM balancing techniques discussed in the subsequent sections (e.g. RAS). While IOT are structured in a way that negative entries do not occur, it seems that non-negativity of all entries is not necessarily fulfilled in the SAM context. Depending on the chosen structure of the SAM, it may be practical to enter subsidies on activities or commodities as rows with negative entries instead of swapping them into a strictly positive column or vice-versa. However, in some cases entries within one row or column may be positive or negative. Operating surplus of productive sectors, stock changes on commodity markets, or negative savings are the most prominent examples. Swapping single negative cells into positive entries in the opposite cells is often not pragmatic if control totals are available for certain sub-matrices of the SAMs as taking out some entries will change the respective sub-totals. For this reason, in the following no assumption about strict non-negativity of the matrices to be estimated will be made.

Based on available information, it may now be possible to create a prior matrix $\mathbf{S^0}$, which is not balanced because of inconsistencies in the used datasets or because of measurement errors and, therefore, does not comply with the available control totals:

(4) $$\sum_i S_{i,j}^0 \neq \sum_i S_{j,i}^0 \quad \text{and} \quad \sum_j \left[\sum_i G_{k,i} S_{i,j}^0 \right] G_{j,l} \neq \Gamma_{k,l}$$

The researcher cannot observe the true values of **S** but the rather distorted ones of $\mathbf{S^0}$ as a result of e.g. editing or reporting errors. Reporting errors in the statistical raw data sets underlying a SAM may result from many causes: household surveys may lack representativeness; business statistics may not cover companies which are exempt from accounting or reporting obligations; product and sector definitions may not be harmonized; or data sets from different points in time may be combined.

Unfortunately, the nature, magnitude and distribution of measurement errors are typically unknown. For application purposes, a balanced SAM is required, and in order to recover the data, $E\left[\mathbf{S^o}\right] = \mathbf{S}$ must be assumed. In many cases, SAMs are constructed for one point in time, only, based on different national and international data sets. A systematic analysis of the error variance of the elements is hence typically not possible. Therefore, robust estimators must be used to balance the SAM. In general, **S** should be as close as possible to $\mathbf{S^0}$, but should fulfil at least condition (2) and also condition (3) if control totals are available. It is consequently necessary to specify a statistical criterion that allows estimating the true **S** which is as close as possible to $\mathbf{S^0}$ subject to these conditions.

The implications of different assumptions about the measurement error for **S** and the most commonly applied estimation approaches are included in the subsequent sections. Moreover, the methodological approaches are grouped into three estimation categories: RAS, minimization of squared differences and entropy, neglecting different variants of absolute errors minimization approaches.

4.1 Measurement Errors

Due to the fact that different parts of the compiled SAM $\mathbf{S^0}$ may originate from various sources (national accounts, household or firm surveys, trade datasets, ...) the associated measurement errors are likely to differ in nature and magnitude. In principle, it can be distinguished between the 'classical' additive and the multiplicative error term (Carroll *et al* (2006), Hwang (1986), Biewen *et al* (2008)).

The additive measurement error is usually assumed to be normally distributed with zero mean and constant variance:

(5) $$\mathbf{S^o} = \mathbf{S} + \mathbf{U} \sim \mathbf{N}\left(\mathbf{0}, \boldsymbol{\sigma}\right)$$

Alternatively, multiplicative measurement errors have been discussed by Carroll *et al.* (2006), for which some evidence could be found in empirical data (Hwang (1986)).

(6) $$\mathbf{S^o} = \mathbf{SU}, \textit{ with } \mathbf{E(U)} = \mathbf{1} \textit{ and } \mathbf{V(U)} = \boldsymbol{\sigma}^2$$

Assumptions about the distribution of **U** are not as well established as for the additive case. In principle, taking logs of both sides of (6) will create again an additive problem similar to (5), with the logarithms of the error term normally distributed with zero mean and constant variance. However, it can be shown that the expected value of zero of the logarithms does not translate into an expected value of one in the original scale. The essential properties $\mathbf{E(U)=1}$ and $\mathbf{V(U)=\sigma^2}$ can be expressed by normal or log-normal distributions, which yields two variants of (6) (see Greene (2003) for the derivation of equation (8)):

(7) $$\mathbf{S^0 = SU \sim N(1,\sigma)}$$

(8) $$\mathbf{S^0 = SU \sim LN(1,\sigma) \Rightarrow n(U) \sim N\left(-\frac{1}{2}\ln\left(1+\sigma^2\right), \sqrt{\ln\left(1+\sigma^2\right)}\right)}$$

As indicated above, it is usually unknown (except in the case of Hwang (1986)) which model for the measurement error actually applies. It might nonetheless be possible to derive some properties from the observed values $\mathbf{S^0}$. A well established assumption is that a zero entry in the observed SAM $\mathbf{S^0}$ can never originate from a non-zero true entry $\mathbf{S}$, meaning that zero entries in the observations will be preserved in the estimate. A more complicated question is whether the sign of an entry should be preserved. Even if national accounts indicate for instance that the operating surplus of an industry was below zero for one year, this can be because of over-valued estimates of elements on the expenditure side, for instance because of the use of outdated input-coefficients rather than "real" negative values. Also, if stock changes on commodity markets are calculated as residuals, the recorded entries may change their sign quite easily when different assumptions about domestic consumption or net-trade are made. However, national accounting items like the current account balance are usually calculated in a reliable manner, and a change of sign in the balanced SAM $\mathbf{S}$ cannot be justified. As a researcher in general has not many possibilities to evaluate the data generating process underlying national accounting tables issued by the responsible national authorities, and probably has no better information, it will be difficult to justify any change in signs for any element in the final SAM. Sign preservation is consequently a desirable feature of an estimation procedure. Although it would be possible to achieve this by simply adding lower and upper bounds on the SAM entries, a formulation of the error term that already accounts for this would be more elegant.

Based on these considerations, it is possible to deduct some properties of the distribution measurement error $\boldsymbol{\varphi(U)}$ for the multiplicative and additive cases:

Additive case:

(9) $$\boldsymbol{\varphi(U)} = \max\left[\mathbf{0}, \mathbf{N(0,\sigma) - S^0}\right]$$

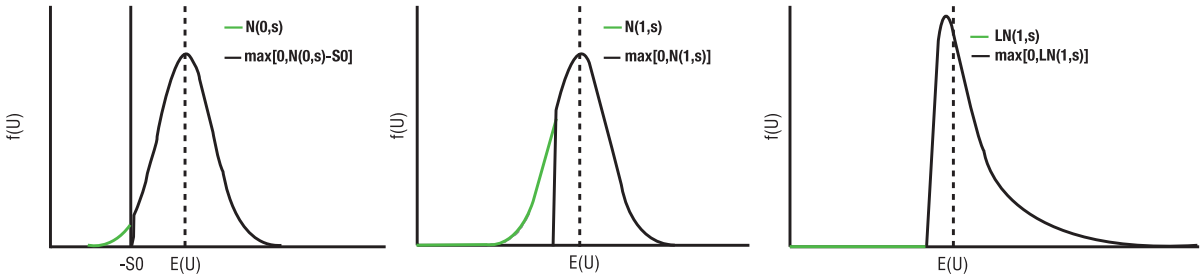
Multiplicative cases:

(10) $$\boldsymbol{\varphi(U)} = \max\left[\mathbf{0}, \mathbf{N(1,\sigma)}\right]$$

(11) $$\boldsymbol{\varphi(U)} = \max\left[\mathbf{0}, \mathbf{LN(1,\sigma)}\right] = \mathbf{LN(1,\sigma)}$$

The distributions of $\mathbf{U}$ are also depicted in Figure 7. These possible distributions of the measurement error of $\mathbf{S}$ and the implications for the performance of the most prominent SAM estimation procedures will be investigated in more detail within this project. For this purpose, a review on the approaches to SAM estimation most frequently applied is presented.

Figure 7. Distributions of Error Term U



4.2 Balancing Methods

4.2.1 RAS Family of Estimators and Minimal Informational Loss

The RAS approach (see e.g. Bacharach 1970) applies iteratively correction factors to the cells in a SAM in order to ensure that given column and row sums are recovered.[4]

(12) $$\mathbf{S} = \hat{\mathbf{R}}\mathbf{S}^{0}\hat{\mathbf{C}}$$

Bacharach (1970) showed that the RAS solution is equivalent to minimizing 'information gain' (see Theil 1967) expressed in terms of observed $\mathbf{S}^0$ and estimated $\mathbf{S}$:

(13) $$\min_{S_{i,j}} Z = \sum_{i,j} S_{i,j} \ln\left(\frac{S_{i,j}}{S^0_{i,j}}\right)$$

The original link to a 'informational loss' permits solving RAS and its modification as a constrained optimization problem by using numerical solution techniques.

(14) $$\begin{aligned} &\min_{S_{i,j}} Z = \sum_{i,j} S_{i,j} \ln\left(\frac{S_{i,j}}{S^0_{i,j}}\right) \\ &s.t. \sum_i S_{i,j} = c^0_j;\ \sum_j S_{i,j} = r^0_i \end{aligned}$$

4.2.2 Minimizing Quadratic Differences

Stone (1968) developed an alternative balancing method based on the minimization of squared deviations that allows incorporating control totals on sub-matrices of the SAM. As before, let $\mathbf{S}^0$ be an initial estimate of a SAM, or a part of a SAM. Assume a set of linear constraints on $\mathbf{S}$, here denoted γ which may represent restrictions on sums of subsets of elements (e.g. the sum of sectoral value added equals total GDP) or restrictions on ratios of elements (e.g. fixed savings ratios).

4 The reader will note that we use $\mathbf{S}$ as nomenclature for the SAM, instead of the original $\mathbf{A}$, what makes the acronym 'RAS'.

(15) $\mathbf{Gs} = \boldsymbol{\gamma}$

Now, let $\mathbf{V}$ be a variance-covariance matrix associated with the vector $\mathbf{s}^0$ (or, equivalently, a matrix of reliability or tolerance estimates of the SAM), the least squares solution gives

(16) $$\mathbf{s} = \mathbf{s}^0 - \mathbf{VG}'\left(\mathbf{GVG}'\right)^{-1}\left(\mathbf{Gs}^0 - \boldsymbol{\gamma}\right)$$

Byron (1978) prefers an interpretation with a quadratic loss-function over Stone's statistical interpretation based on the variance-covariance matrix.

Whereas Stone and Byron apply a matrix inversion linked to the first-order conditions of the minimization model and putting bounds on the size of matrices which can be handled, more recent authors solve this class of problems within a constrained optimization framework. Several approaches are documented minimizing (normalized) quadratic differences between the estimates $\mathbf{S}$ and the observations $\mathbf{S}^0$.

Following Byron's interpretation of a loss function, older versions of the SAM Balancing subroutines in the standard CGE model of the "International Food Policy Research Institute" (IFPRI) used the following penalty function (Lofgren *et al,* GAMS code before (2001)):

(17) $$\min_{S_{i,j}} Z = \sum_{i,j} \left(\frac{S_{i,j} - S_{i,j}^0}{S_{i,j}^0} \right)^2$$

This approach is discussed by Round (2003) and applied by Nakamura (1998), and in fact also by Müller and Wehrheim (2004) for the estimation of a SAM for Russia.

Essentially, this is the same approach as Stone-Byron for the special case where the coefficients of variation of all elements are equal; that is, where the initial estimates are judged to be of equal relative reliability. Usually, our prior judgement about the relative reliability of different data sources will allow us to do better than this and therefore, in general, Stone-Byron would be preferred to the quadratic minimand above (Round (2003)), whereas Round (2003) and Canning and Wang (2005) allow for cell specific weights:

(18) $$\min_{S_{i,j}} Z = \sum_{i,j} \frac{\left(S_{i,j} - S_{i,j}^0\right)^2}{w_{i,j}}$$

4.2.3 Entropy Approaches

Already before the publication of the seminal textbook by Golan, Judge, and Miller on Entropy Econometrics in 1996, entropy based estimation procedures were introduced by Golan, Judge, and Robinson in 1994 (Golan *et al* (1994)). Here, two variants can be distinguished: a probabilistic (Golan *et al* (1994)) and an information-theoretic approach (Robinson *et al* (2001)), both based on the column-coefficients $\mathbf{A}$ (equation 19) rather than on the actual values $\mathbf{S}$ in the SAM.

(19) $$A_{i,j} = \frac{S_{i,j}}{c_j},\ A^0_{i,j} = \frac{S^0_{i,j}}{c^0_j}$$

The probabilistic approach attaches a discrete probability distribution to each element of **A**, defined by a set of **s** discrete points **BA** and associated probabilities **WA**.

(20) $$\min_{W^A_{s,i,j}} Z = \sum_{s,i,j} W^A_{s,i,j} \ln\left(\frac{W^A_{s,i,j}}{W^{A0}_{s,i,j}}\right)$$

subject to

(21) $$A_{i,j} = \sum_s B^A_{i,j,s} W^A_{i,j,s}$$ Coefficient definition

(22) $$\sum_i A_{i,j} = 1,\ 0 \le A_{i,j} \le 1$$ Coefficient constraint

(23) $$\sum_s W^A_{i,j,s} = 1,\ 0 \le W^A_{i,j,s} \le 1$$ Probability constraint

(24) $$r^0_i = \sum_j A_{i,j} c^0_j$$ Balancing constraint

Some more practical applications use variations of the original model, for instance Breisinger *et al* (2007) in the case of a SAM for Ghana, where original column coefficients are associated with an error term:

(25) $$A_{i,j} = A^0_{i,j} + e^A_{i,j};\ e^A_{i,j} = \sum_s B^A_{i,j,s} W^A_{i,j,s}$$

In this case an additional error term is placed on the control-totals:

(26) $$\sum_j \left[\sum_i G_{k,i} S_{i,j}\right] G_{j,l} = \Gamma_{k,l} + e_{k,l};\quad e_{k,l} = \sum_s B^\Gamma_{s,k,l} W^\Gamma_{s,k,l}$$

The resulting objective function (27) minimizes the joint entropy of errors on coefficients and control totals, subject to the SAM balancing constraint (24), probability constraints for **WA** and **WΓ** similar to (23), coefficient constraint (22) and the control totals (26), while the coefficient definition (21) is replaced by (25):

(27) $$\min_{W^A_{s,i,j}, W_{j,s}} Z = \sum_{s,i,j} W^A_{s,i,j} \ln\left(\frac{W^A_{s,i,j}}{W^{A0}_{s,i,j}}\right) + \sum_{s,k,l} W^\Gamma_{s,k,l} \ln\left(\frac{W^\Gamma_{s,k,l}}{W^{0\Gamma}_{s,k,l}}\right)$$

Alternatively the estimation procedure may be based on the actual values rather than the column-coefficients (e.g. Müller (2006) in the case of a SAM for Uzbekistan). Also here it is possible to account for the fact that some datasets used for the compilation of $\mathbf{S^0}$ are more reliable than others by choosing support points **BS** to allow for larger or smaller deviations from the original SAM entry $\mathbf{S^0}$. The objective function (29) is here minimized subject to the SAM-entry definition (28), a balancing constraint (2) and a control-total constraint (3).

(28) $$S_{i,j} = \sum_{s} B^{S}_{s,i,j} W^{S}_{s,i,j}$$

(29) $$\min_{W_{s,i,j}} Z = \sum_{s,i,j} W^{S}_{s,i,j} \ln\left(\frac{W^{S}_{s,i,j}}{W^{0S}_{s,i,j}}\right)$$

The support points **BS** can be derived based on the assumption of a normal distributed measurement error associated with **S**, for which the variances are known or that there is at least a guess that they might be larger for some entries than for others:

(30) $$S_{i,j} \sim N\left(S^{0}_{i,j}, \sigma_{i,j}\right)$$

Following the three-sigma rule, it is possible to define a number of support points within an interval of plus/minus three times the (known or assumed) standard deviation:

(31) $$B^{S}_{s,i,j} = \left[S^{0}_{i,j} - 3\sigma_{i,j}, \ldots, S^{0}_{i,j} + 3\sigma_{i,j}\right]$$

The approach by Müller (2006) is essentially similar to a hybrid-approach based on the RAS and GCE methods as suggested by Round (2003), since it combines a cross entropy minimand based on transactions (i.e. flows) instead of coefficients with the requirement of row- and column equality and additional (linear and/or nonlinear) constraints.

Golan, Judge, and Robinson (1994) also discuss an information-theoretic approach in which the entropy distance between observed column-coefficients and the balanced estimate is minimized instead of the deviations between prior and estimated probabilities:

(32) $$\min_{A_{i,j}} Z = \sum_{i,j} A_{i,j} \ln\left(\frac{A_{i,j}}{A^{0}_{i,j}}\right)$$

The minimization problem in (32) is subject to the balancing constraints (24) and (22). This approach was extended by Robinson, Cattaneo, and El-Said in 2001 (Robinson *et al* (2001)) by including an error term for column sums, such that the objective function minimizes the joint entropies of column-coefficients and error on observed column totals. The error term is expressed by a discrete probability distribution $\mathbf{Wc^0}$ associated with a set of support points $\mathbf{B^c}$.

(33) $$\min_{A_{i,j},W^c_{j,s}} Z = \sum_{i,j} A_{i,j} \ln\left(\frac{A_{i,j}}{A^0_{i,j}}\right) + \sum_{j,s} W^c_{j,s} \ln\left(\frac{W^c_{j,s}}{W^{c0}_{j,s}}\right)$$

(34) s.t.

(35) $$\sum_i A_{i,j} = 1;\ 0 \le A_{i,j} \le 1$$

(36) $$\sum_s W^c_{j,s} = 1;\ 0 \le W^c_{j,s} \le 1$$

(37) $$r_i = c^0_i + e_i;\ e_i = \sum_s B^c_{i,s} W^c_{i,s}$$

(38) $$r_i = \sum_j A_{i,j}\left(c^0_j + e_j\right)$$

(39) $$\sum_j \left[\sum_i G_{k,i} S_{i,j}\right] G_{j,l} = \Gamma_{k,l} \quad \text{with} \quad S_{i,j} = A_{i,j}\left(c^0_j + e_j\right)$$

This approach was widely applied in SAM compilation projects at the "International Food Policy Research Institute" (IFPRI). The disadvantage of this approach is that it cannot accommodate negative entries in the coefficient matrix **A** and that it does not permit expressing information about the reliability of data-sources as in the case of Golan *et al* (1994), Stone-Byron or, more generally, any weighted least squares method.

4.3 Chosen Approach

Given the considerations above on negative entries, sign preservation and the need to express confidence in the different types of used datasets, the variant of the Golan *et al*. (1994) approach as suggested by Round (2003) and also implemented by Müller (2006) (equations (2), (3), (28), and (29)) appears to be the most suitable approach for the general problem to compile SAMs based on the datasets described in section 3. This approach is throughout the following compilation steps combined with the assumption of a multiplicative disturbance term associated with the observed values $\mathbf{S^0}$. These disturbances are incorporated in the estimation models used in the subsequent sections by defining a correction coefficient kappa ($\boldsymbol{\kappa}$):

(40) $$S_{i,j} = S^0_{i,j} \cdot \kappa_{i,j}$$

Kappa should have the following properties:

(41) $$E[\kappa] = 1$$
$$0 < \kappa \le \infty$$

The expected value should be 1 (in which case the balanced value equals the prior), and it should not assume negative values in order to avoid the change of the sign of any prior entry. Furthermore, it should not be equal to zero as it is assumed that once there is a prior entry, there should also be a non-zero entry in the balanced dataset. The assumed reliability of the prior data should also influence the possible outcomes for kappa. Kappa is here expressed as an exponential function of **s** support points (**b**) and the associated weights (**W**).

(42) $$\kappa = \exp\left[\sum_s W_s \cdot b_s\right]$$

These weights have to add up to unity and should be as close as possible to a set of pre-defined prior weights.

The support points are arbitrarily defined according to the 3-sigma rule (in the case of two support points):

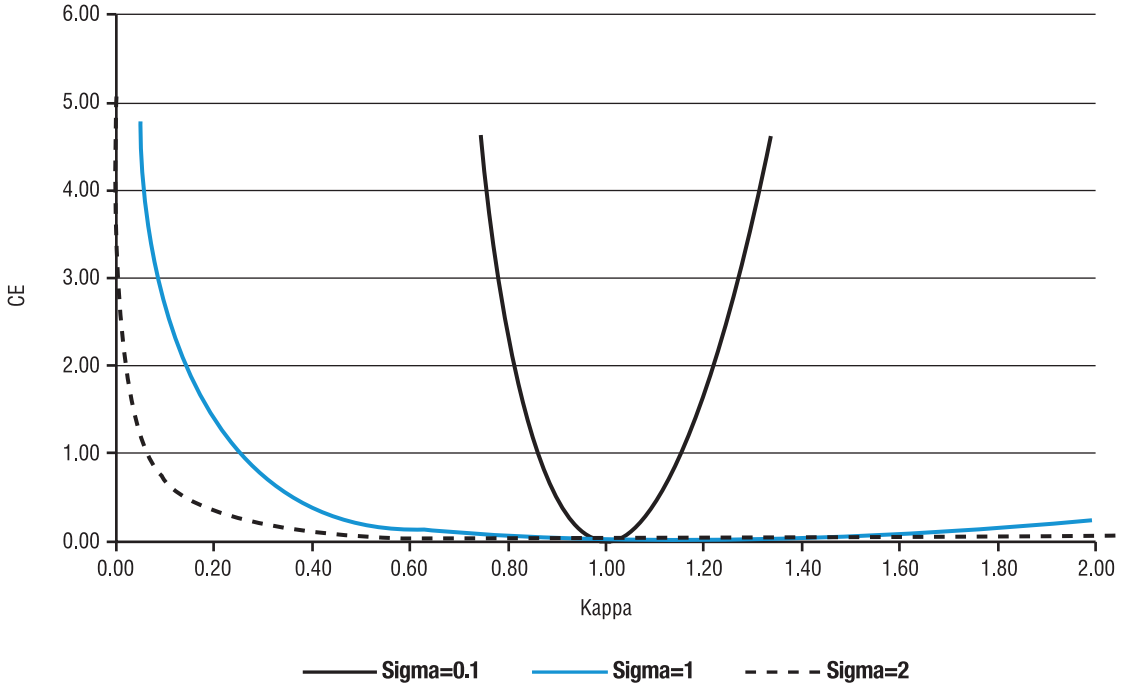
(43) $$\begin{aligned} b_s &= [-3,3]\cdot\sigma \\ \bar{W}_s &= \left[\tfrac{1}{2},\tfrac{1}{2}\right] \end{aligned}$$

Where **σ** is a variance parameter the cross-entropy minimization model can be summarized as follows:

(44) $$\begin{aligned} & CE = \min \sum_s W_s \cdot \ln\left[W_s / \bar{W}_s\right] \\ & s.t. \\ & \kappa = \exp\left[\sum_s W_s \cdot b_s\right] \\ & \sum_s W_s = 1 \\ & S_{i,j} = S^0_{i,j} \cdot \kappa_{i,j} \\ & \textit{accounting identities for } S \end{aligned}$$

When solving the problem above for different values for kappa and σ, a plot of the objective function is obtained (i.e. by neglecting the accounting identities for **S**), as shown in Figure 8:

Figure 8. Cross-Entropy Function of Kappa



Source: own calculations

The values for σ (0.1, 1, and 2), as used in Figure 8, were chosen to express the trust attached to the different prior data at hand. It has to be noted here that the decision on which value to choose for σ is a qualitative judgement and cannot be supported by a systematic quantitative assessment of potential variances of the prior data[5]. Instead of deriving any other quantitative indicator like variance over time or MS, domestic production and trade of cereals, oilseeds, and dairy products were considered as comparatively well monitored and, therefore, given a high trust. "Fodder crops", "other crops", or "other animals", since they are derived as residuals or according to rigid assumptions about input coefficients in the raw dataset, were given a lower trust.

5 Knowledge about the variances obtained in the original CAPRI database estimation procedure could improve the quality of this decision, but they were not available for this project.

5. Compilation of Macro Totals

For the creation of a consolidated set of macroeconomic indicators (mid) and associated MacroSAMs, the indicators listed in Table 3 were considered. At this stage, it appeared to be more pragmatic not to distinguish between households, enterprises, and national government but merge them into an aggregate 'domestic institutions'. Likewise no distinctions between taxes and subsidies on production (or on products) were made. By doing this it was assured that none of the 22 entries in the MacroSAMs (Table 4) could plausibly assume a value of zero, which substantially facilitated further compilation procedures.

As can be seen from Table 3, there were many cases in which the considered macro indicators could be drawn from multiple datasets. In general, the most complete dataset was the preferred option, but it had to be decided case by case which dataset was chosen as base value for the following estimation procedure.

5.1 Target Structure of the Macroeconomic SAMs

A set MacroSAMs and related macroeconomic indicators was generated to ensure consistency with macro-totals such e.g. Gross Domestic Product at Market Prices (GDPm) or total domestic savings (Sd), which are not directly elements of the SAMs as certain items in the SAMs have to be added up to generate these values,. This step proved to be problematic, since the macro indicators provided by EuroStat were neither complete nor consistent across different sources (e.g. NASA, NAIO, or NAMA), see for instance in Figure 9 the deviation in gross domestic capital formation (ESA code: p5) in the cases of Austria and Malta. When mapping the macro indicators in the format of a MacroSAM as shown in Table 4, results were in some cases satisfying (Table 5), whereas in other cases substantial differences in the available information rendered the MacroSAMs unbalanced (Table 6).

The chosen structure of the MacroSAMs as shown in Table 4 represents a compromise between desirable detail and available information. It consists of 22 entries, for which it can be assumed that they have to be different from zero. Apart from this, most of the chosen data points could be obtained from all three main sources, and the comparatively small number of values for each year and Member State facilitates the detection of unrealistic or false entries. The process at this stage could not be fully automatised and a significant amount of manual operations was necessary to account for all occurring exceptions. Table 5 illustrates a case where domestic final consumption (ESA code p3, row: COM, column: DIN) has different values when originating from NAMA or NASA. Here, NAMA was selected, but the final choice had to be made for each Member State and indicator separately.

Table 3. Selected Macro Indicators from ESA95

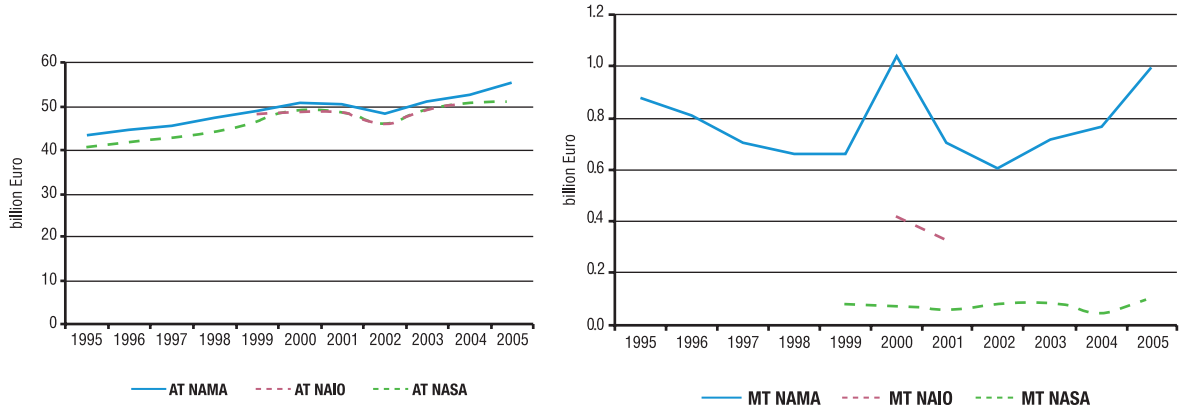
ESA Code	Description	Sources
p7	Imports of goods and services[2)]	NAMA, NASA, NAIO
p6	Exports of goods and services[2)]	NAMA, NASA, NAIO
p5	Gross capital formation[2)]	NAMA, NASA, NAIO
p3	Final consumption expenditure[2)]	NAMA, NASA, NAIO
p2	Intermediate consumption[2)]	NAMA, NASA, NAIO
p118_r	Trade and transport margins, received by commodity markets[1) 2)]	NAIO
p118_p	Trade and transport margins, paid by commodity markets[1) 2)]	NAIO
p1	Output at basic prices[2)]	NAMA, NASA, NAIO
d4_d8_s2_p	Direct taxes and transfers, paid by ROW[1) 2)]	NASA
d4_d8_s1_p	Direct taxes and transfers, paid by DINS[1) 2)]	NASA
d29_m_d39	Other net taxes on production[2)]	NASA
d21_m_d31	Taxes less subsidies on products[2)]	NASA
d2_d8_s2_r	Taxes and transfers, received by ROW[1) 2)]	NASA
d2_d8_s1_r	Taxes and transfers, received by DINS[1) 2)]	NASA
d1_s2_r	Compensation of employees abroad[1) 2)]	NASA
d1_s2_p	Compensation of employees from abroad[1) 2)]	NASA
d1_s1_r	Compensation of domestic employees[1) 2)]	NASA
d1	Compensation of employees[2)]	NAMA, NASA, NAIO
b8g	Gross saving[2)]	NAMA, NASA, IMF[3)]
b2g_b3g	Gross operating surplus and gross mixed income[2)]	NAMA, NASA, NAIO
b1g	Value added at basic prices	NAMA, NASA, NAIO
b12	Current external balance[2)]	NAMA, NASA, IMF[3)]
p33	Final consumption expenditure of resident households in the rest of the world - total	NAIO, NAMA
p34	Final consumption expenditure of non-resident households on the economic territory - total	NAIO, NAMA
b1gm	Gross domestic product at market prices	NAMA, NASA, NAIO, IMF
Dp7a	Cif/ fob adjustments on imports [4)]	NAIO
Dp6a	Cif/ fob adjustments on exports [4)]	NAIO
EUX	Euro-exchange rate (NAC/EURO) [1)]	Derived
USX	US dollar exchange rate (NAC/USD) [1)]	Derived
CPI	Consumer price index[1)]	Derived
PPI	Producer price index[1)]	Derived
p1_p_p7	Domestic absorption[1)]	Derived
b1gm_2000	Gross domestic product at constant 2000 market prices	Derived
i2000	Index, 2000=100	NAMA
POP	Total population	NAMA
SAL	Employees	NAMA

Notes:
1) Identifiers are not used as displayed within ESA95. E.g. trade and transport margins (p118) are recorded as positive (paid) and negative (received) entries in the NAIO datasets, which sum up to zero. As it is important within the SAM framework to distinguish between paying and receiving sectors, this difference is identified by adding a suffix '_p' and '_r'. The same applies for direct taxes and transfers paid and received by domestic institutions and the 'rest of the world'
2) Element of MacroSAMs
3) Data available as share in GDP at market prices
4) Cost, insurance and freight (cif) and free on board (fob)

Table 4. Target Structure of the MacroSAMs and Correspondence to ESA95

		ACT	COM	LAB	CAP	TRD	TNF	DIN	ROW	SIA	
Activities	ACT		p1								Revenues
Commodities	COM	p2				p118_r		p3	p6	p5	
Labour	LAB	d1							d1_s2_p		
Physical capital	CAP	b2g_b3g									
Trade and transport margins	TRD		p118_p								
Direct and indirect taxes, transfers	TNF	d29_m_d39	d21_m_d31					d4_d8_s1_p	d4_d8_s2_p		
Domestic institutions	DIN			d1_s1_r	b2g_b3g		d2_d8_s1_r				
Rest of the world	ROW		p7	d1_s2_r			d2_d8_s2_r				
Saving-Investment account	SIA							b8g	b12		
		Expenditures									

Figure 9. Deviating Information between Macroeconomic Indicators: Gross Capital Formation in Malta and Austria



Source: EuroStat, own representation
Note: MT: Malta, AT: Austria

Table 5. Matching Information: MacroSAM for Germany in 2000, billion Euro, current

		ACT	COM	LAB	CAP	TRD	TNF	DIN	ROW	SIA	
Activities	ACT		3681								3681
Commodities	COM	1824				317		1606	688	449	4885
Labour	LAB	1102							4		1106
Physical capital	CAP	744									744
Trade and transport margins	TRD		317								317
Direct and indirect taxes, transfers	TNF	11	206					2250	118		2585
Domestic institutions	DIN			1100	744		2428				4272
Rest of the world	ROW		681	6			157				844
Saving-Investment account	SIA							416	34		449
	Total	3681	4885	1106	744	317	2585	4272	844	449	

Legend: Colours for data sources NAMA NASA NAIO

Table 6. Ambiguous Information: MacroSAM for Luxemburg in 2000, billion Euro, current

		ACT	COM	LAB	CAP	TRD	TNF	DIN	ROW	SIA	
Activities	ACT		53								53
Commodities	COM	33				2		12 *3*	33	5 *1*	86
Labour	LAB	10							1		11
Physical capital	CAP	9									9
Trade and transport margins	TRD		2								2
Direct and indirect taxes, transfers	TNF	0	2					4	60		67
Domestic institutions	DIN			8	9		9				26
Rest of the world	ROW		28	3			61				92
Saving-Investment account	SIA							3	-3	*0*	0
	Total	53	86	11	9	2	70	19	91	5	

Legend: Colours for data sources NAMA | NASA | *NASA not selected* | NAIO | WEO

Although the choice of the sources for each entry in the MacroSAMs was made such that they were as close as possible to being balanced, the result in most cases was similar to the situation depicted in Table 6. As a result, and because of the relevance of a consistent set of macro-totals for all subsequent steps, it was necessary to employ a balancing procedure already at this stage. GDPm was chosen as indicator for normalising all other indicators, since it was the only one with time-series fully available for all 27 EU MS from 1995 to 2005.

(45) $$msr\_obs_{ms27,t,i} = \frac{mid\_obs_{ms27,t,i}}{GDPm_{ms27,t}}$$

Where:

mid_obs Observed macro-economic indicator

msr_obs Observed share of each macro-indicator in GDPm

For those shares, country-specific averages and standard deviations were computed in case there was at least some observation available:

(46) $$msr\_avr_{ms27,i} = \frac{\sum_t msr\_obs_{ms27,t,i}}{\sum_t 1} \forall t : msr\_obs_{ms27,t,i} \neq 0$$

(47) $$msr\_std_{ms27,i} = \sqrt{\frac{\sum_t \left(msr\_avr_{ms27,i} - msr\_obs_{ms27,t,i}\right)^2}{\left(\sum_t 1\right) - 1}} \forall t : msr\_obs_{ms27,t,i} \neq 0$$

In case time series did not exist for one country (e.g. trade margins in those countries for which no SUT were available), the EU average and standard deviation shares were used for the respective:

$$(48) \qquad msr\_avr\_ms_i = \frac{\sum_{t,ms27} msr\_obs_{ms27,t,i}}{\sum_{t,ms27} 1} \forall t, ms27 : msr\_obs_{ms27,t,i} \neq 0$$

$$(49) \qquad msr\_std\_ms_i = \sqrt{\frac{\sum_{t,ms27} \left(msr\_avr\_ms_i - msr\_obs_{ms27,t,i}\right)^2}{\left(\sum_{t,ms27} 1\right) - 1}} \forall t, ms27 : msr\_obs_{ms27,t,i} \neq 0$$

For each indicator the relation between prior and balanced data was expressed via a correction coefficient kappa (κ) (see also equation (41)):

$$(50) \qquad mid\_fin_i = mid\_bse_i \cdot \exp(\kappa)_i$$

Where:

mid_fin	Final macro indicator
mid_bse	Prior value for macro indicator
$\exp(\kappa)$	Correction coefficient

This formulation of a correction coefficient had the following desirable properties:

- Final values and base values have the same sign
- Non-zero values remain non-zero, but can assume very small values
- Feasibility of the system is guaranteed as potentially large deviations are possible

The base values that entered the estimation procedure were chosen according to the information available (i.e. observation taken if existing). In case no information was available for a number of years but there were observations from other years, the average GDP-shares for the country in question were used. Finally, in case no observation was existing for a country, the average of all existing observation is chosen:

$$(51) \qquad mid\_bse = \begin{cases} mid\_obs \ \forall mid\_obs \neq 0 \\ msr\_avr \cdot GDPm \ \forall mid\_obs = 0 \wedge msr\_avr \neq 0 \\ msr\_avr\_ms \cdot GDPm \ \forall mid\_obs = 0 \wedge msr\_avr = 0 \end{cases}$$

The standard deviations were chosen in a similar manner: in case observations exist, a deviation of only 0.1 % from the original value was permitted. In case an observation is missing in a number of years, but there were sufficient data-points from other years available for a given country, the country-specific average standard deviation was taken. Else, the total average was used:

(52) $$mid\_std = \begin{cases} 0.001 \cdot mid\_obs \ \forall mid\_obs \neq 0 \\ msr\_std \cdot GDPm \ \forall mid\_obs = 0 \wedge msr\_avr \neq 0 \\ msr\_std\_ms \cdot GDPm \ \forall mid\_obs = 0 \wedge msr\_avr = 0 \end{cases}$$

Additionally, and in order to implement a cross-entropy estimator, a set of support points (b) and associated weights (W) was defined, such that

(53) $$\sum_s \bar{W}_s b_s = E[\kappa]$$

Where:

(54) $$\begin{aligned} b_s &= [E[\kappa] - 3STD[\kappa], E[\kappa] + 3STD[\kappa]] \\ \bar{W}_s &= \left[\tfrac{1}{2}, \tfrac{1}{2}\right] \end{aligned}$$

The estimation model is written in form of a penalising function (i.e. minimisation of deviations between estimated and base values) and can be summarised as follows[6]:

(55) $$\begin{aligned} & CE = \min \sum_s W_s \cdot \ln\left[W_s / \bar{W}_s\right] \\ & s.t. \\ & \kappa = \sum_s W_s \cdot b_s, \quad \sum_s W_s = 1 \\ & mid\_fin = mid\_bse \cdot \exp[\kappa] \\ & accounting\ identities\ for\ mid\_fin \end{aligned}$$

Once the estimation routine was solved, the MacroSAMs were completed by mapping the relevant entries into the SAM framework. The thus obtained MacroSAMs would serve as control-totals in the following step to create a set of institutional SAMs (InstSAMs)

5.2 Institutional SAMs

The MacroSAMs obtained in the previous step are still too coarse to serve as starting point for the structure of the final SAMs targeted in this project, which should include a detailed representation of productive activities and commodity markets. Moreover, and in order to allow for their potential further use within General Equilibrium Models and the creation of IOT, the distinction of different tax types and domestic institutions is required. Consequently the next step is to disaggregate tax accounts and domestic institutions and to estimate the flows between those. The structure of the InstSAMs compiled in the following step will be similar to the targeted structure of the final AgroSAMs, apart from the disaggregation of activities and commodities.

6 This procedure was implemented in GAMS and solved with a CONOPT3 solver. For 27 MS and 11 years it required 260 Mb of memory space and 63 minutes a solving time on a XEON dual core processor with 2.80GHz.

The compilation of base values for the InstSAMs was carried out in a manner similar to the compilation of macro indicators. The main dataset used here was NASA as this provided the needed information on flows (taxes and transfers) between domestic and foreign institutions. The main difference in this step with respect to the previous one is that the available data was not expressed as shares in Gross Domestic Product at Market Prices (GDPm) but as shares of the corresponding total in the MacroSAM. In principle, this procedure is based on the information-theoretic approach by Golan at al (1994), as outlined in section 4.2, but uses the shares within sub-totals rather than column coefficients. The general idea is nevertheless similar as in both cases deviations from structural information are penalised by the objective function.

An example may illustrate this procedure. Consider the case of direct taxes and transfers paid by domestic institutions in Germany (d4_d8_s1_p, or cell [TNF,DIN] in the MacroSAM, see Table 5). The value 2250 billion Euro has to be split here into a maximum of five types of taxes and transfers (d4, d5, d6, d7, d8) times three types of domestic institutions (enterprises s11_s12, government s13, and households s14_s15). For each potential entry, a share matrix A was calculated based on the available information (see Figure 10). As these shares have to add up to unity and range between zero and one, they can directly enter a cross-entropy framework.

Practically, this concept is put to work by first, calculating an expected value for the share matrices ($\bar{A}$) and second, by expanding the balanced MacroSAMs into the framework of the institutional SAMs ($\overline{InstSAM}$).

Figure 10. Example of Compilation of an Institutional SAM from a Macroeconomic SAM (Direct Taxes and Transfers Paid by Domestic Institutions in Germany)

InstSAM

	I_s11_s12	I_s13	I_s14_s15
T_d4	683	65	73
T_d5	35	0	217
T_d6	34	380	427
T_d7	68	186	67
T_d8	15		

=

A

	I_s11_s12	I_s13	I_s14_s15
T_d4	0.30	0.03	0.03
T_d5	0.02	0.00	0.10
T_d6	0.02	0.17	0.19
T_d7	0.03	0.08	0.03
T_d8	0.01	0.00	0.00

·

$\overline{InstSAM}$

	I_s11_s12	I_s13	I_s14_s15
T_d4	2250	2250	2250
T_d5	2250	2250	2250
T_d6	2250	2250	2250
T_d7	2250	2250	2250
T_d8	2250	2250	2250

The model is summarized in equation (56) an example for an institutional SAM is given in Table 7.

$$
\begin{aligned}
& CE = \min \sum_{AC,AC'} A_{AC,AC'} \cdot \ln\left[A_{AC,AC'} / \bar{A}_{AC,AC'}\right] \\
& s.t. \\
& InstSAM_{AC,AC'} = \overline{InstSAM}_{AC,AC'} \cdot A_{AC,AC'} \\
(56) \quad & \sum_{AC} InstSAM_{AC,AC'} = \sum_{AC'} InstSAM_{AC,AC'} \\
& \sum_{AC'} \left[\sum_{AC} G_{MKR,AC} \cdot A_{AC,AC'} \right] \cdot G_{MKR',AC'} = 1, \quad 0 \le A_{AC,AC'} \le 1 \\
& \sum_{AC'} \left[\sum_{AC} G_{MKR,AC} \cdot InstSAM_{AC,AC'} \right] \cdot G_{MKR',AC'} = MacroSAM_{MKR,MKR'}
\end{aligned}
$$

Table 7. Balanced Institutional SAM for Cyprus in 2000, billion Euro, current

Description	Code	A_Total	C_Total	F_d1	F_b2g_b3g	T_p118	T_d29	T_d39	T_d211	T_d212	T_d214	T_d31	T_d4	T_d5	T_d6	T_d7	T_d8	I_s11_s12	I_s13	I_s14_s15	I_s2	I_p5	Total
Activities	A_Total		19.2																				19.2
Commodities	C_Total	10.1				1.8													1.7	6.3	5.5	1.8	27.1
Compensation of employees	F_d1	4.2																			0.0		4.2
Gross operating surplus	F_b2g_b3g	4.6																					4.6
Trade and transport margins	T_p118		1.8																				1.8
Other taxes on production	T_d29	0.3																					0.3
Other subsidies on production	T_d39	0.0																					0.0
Value added type taxes (VAT)	T_d211		0.6																				0.6
Taxes and duties on imports excl. VAT	T_d212		0.3																				0.3
Other taxes on products	T_d214		0.3																				0.3
Subsidies on products	T_d31		-0.4																				-0.4
Property income	T_d4																	0.7	0.1	0.1	0.0		0.9
Current taxes on income, wealth, etc	T_d5																	0.1	0.0	0.2	0.0		0.2
Social contributions and benefits	T_d6																	0.0	0.4	0.2	0.0		0.6
Other current transfers	T_d7																	0.1	0.2	0.1	0.0		0.3
Adjustment for change in net equity	T_d8																	0.0					0.0
Corporations	I_s11_s12				3.0								0.1		0.0	0.0							3.2
General government	I_s13				0.1		0.3	0.0	0.6	0.2	0.3	0.0	0.0	0.2	0.3	0.1							2.1
Households	I_s14_s15			4.2	1.5								0.1		0.3	0.1	0.0						6.2
Rest of the world	I_s2		5.4	0.0					0.0	0.0	0.0	-0.4	0.6	0.0	0.0	0.1							5.8
Savings-Investment account	I_p5																	2.4	-0.2	-0.6	0.3		1.8
	Total	19.2	27.1	4.2	4.6	1.8	0.3	0.0	0.6	0.3	0.3	-0.4	0.9	0.2	0.6	0.3	0.0	3.2	2.1	6.2	5.8	1.8	

Note: The dark blue areas indicate entries which were not estimated during step 1, but for which control-totals were estimated

Where:

InstSAM:	Balanced InstSAM
$\overline{InstSAM}$:	Expanded MacroSAM
A:	Share of InstSAM entry in the corresponding macro totals
$\overline{A}$:	Expected value of A
CE:	Cross-entropy minimand
AC,AC':	Index Accounts of the InstSAM
MKR,MKR':	Index for Macro Totals
G:	Aggregator matrix between MKR and AC

5.3 Estimation Results

The estimation of 33 macro indicators covered a period of 11 years and 27 Member States, times two currencies (National Currency Units and Euro), resulting in a total of 19602 entries and a corresponding number of correction coefficients kappa (κ). The frequencies of exp[k] are depicted in Figure 11 for the full sample. Despite the sometimes large permitted deviation from unity (as expressed by the standard deviation of each indicator), the mass of the final values for the correction coefficient ranges quite narrowly around 1. However, there were also several cases where exp[k] assumed values larger than 5, e.g. 'indirect taxes on production' (d29) in Romania 2000 deviates almost by the factor 14 from its base value (Figure 12). Although this seems to be a problematic result on first sight, a more detailed investigation reveals that this particular outcome is in fact in line with the general idea underlying the presented estimation procedure: values for d29 are not available before 2001 from the official statistics, so the base value is derived based on the average GDP share of d29 in GDP at market prices. However, 'net taxes on production' (d29_m_d39) are available from 1998 onwards. The final value for d29 is therefore corrected such that it complies with the available entry for d29_d39.

Figure 11. Frequency of the Correction Term

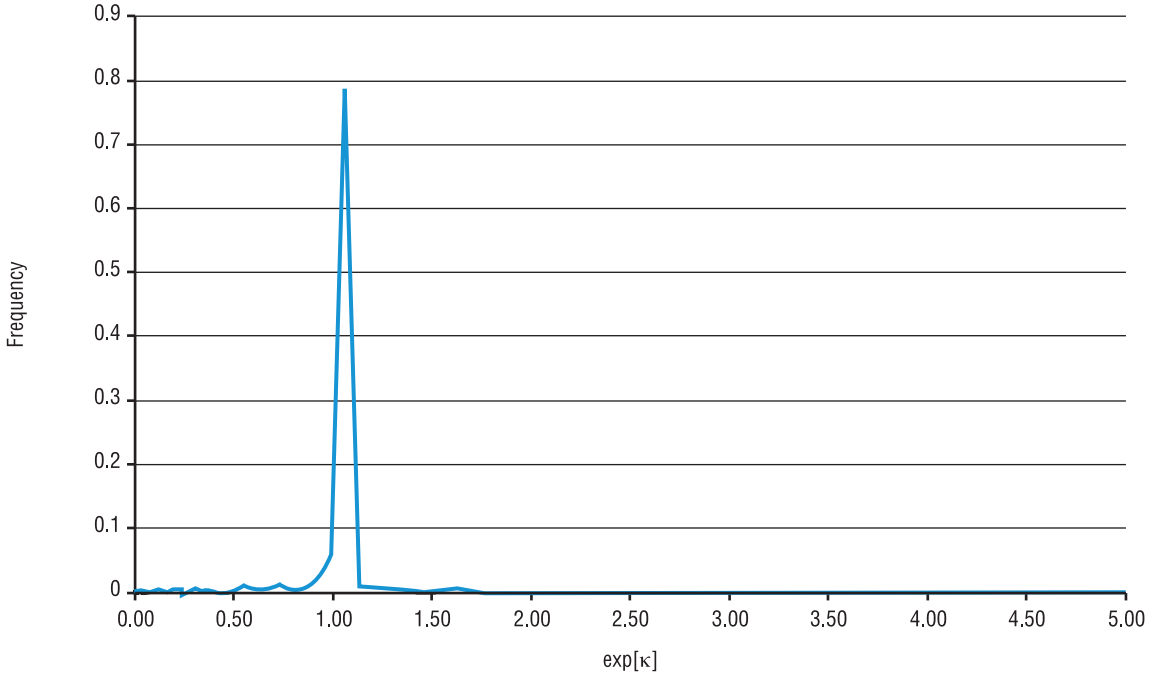


Figure 12. Cases for Correction Terms Larger 5

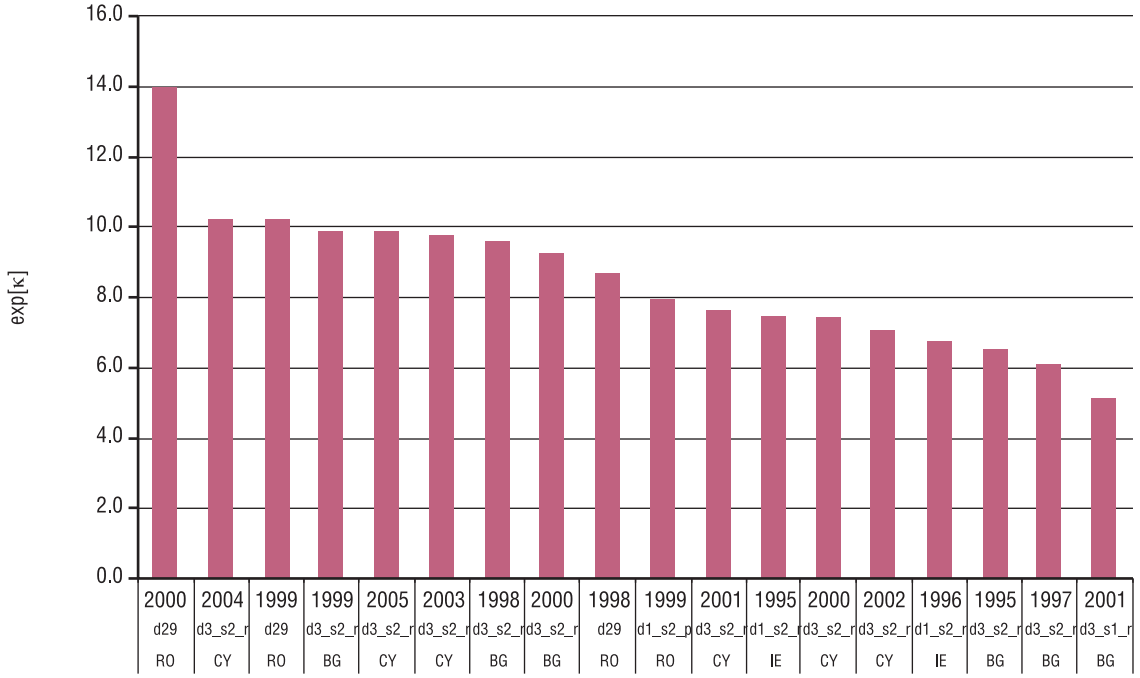
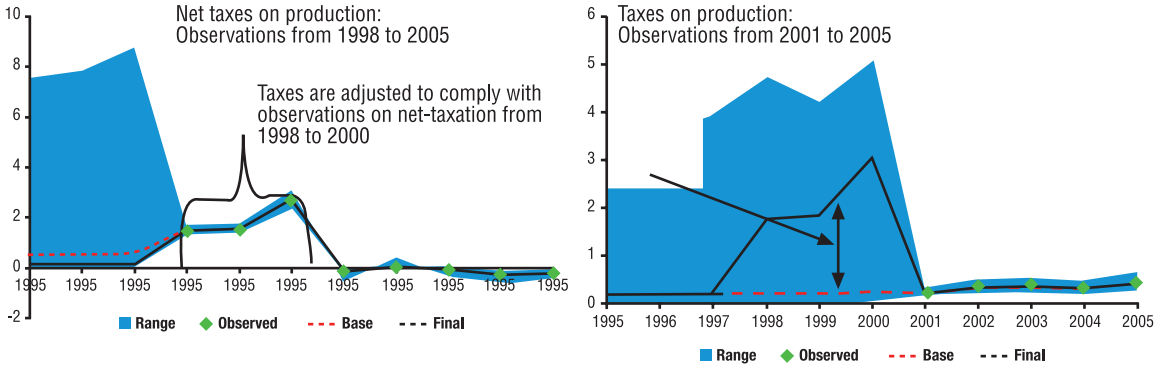


Figure 13. Indirect Taxes in Romania



Note: Blue shaded area indicates the allowed range for deviations from the base value

6. Construction of SAMs in ESA95 Format

The next step for the construction of the AgroSAM database is the compilation of a comprehensive set of SAMs according to the ESA95 classification used by EuroStat. These SAMs distinguish 59 productive sectors and commodities and will be noted as ESASAM in the following. The ESA95 classification scheme can be found in the correspondence sheets in Appendix 1 (Table 15). The stylized structure of the ESASAM is mainly shaped by the structure of the main input datasets, namely the SUT (NAIO datasets) and the institutional accounts taken from the InstSAMs from the previous steps. A full SAM in this format can be readily compiled based on these two sets of data. Table 8 gives an overview for which of the EU-27 Member States both of these datasets exist.

Table 8. Availability of NASA and NAIO Datasets for EU-27 Member States

Country	Code	1995	1996	1997	1998	1999	2000	2001	2002	2003	2004
Belgium	BE	1		1		1	1	1	1		
Bulgaria	BG										
Czech Republic	CZ	1	1	1	1	1	1	1	1	1	1
Denmark	DK	1	1	1	1	1	1				
Germany	DE	1		1	1	1	1	1	1	1	
Estonia	EE			1			1	1	1		
Ireland	IE										
Greece	GR	1	1	1	1	1					
Spain	ES	1	1	1	1		1	1			
France	FR	1		1		1	1	1			
Italy	IT	1	1	1	1	1	1	1	1	1	
Cyprus	CY										
Latvia	LV		1		1						
Lithuania	LT						1	1	1	1	
Luxembourg	LU	1	1	1	1	1	1	1	1	1	
Hungary	HU										
Malta	MT						1	1			
Netherlands	NL	1	1	1	1	1	1	1		1	
Austria	AT	1		1		1	1	1	1	1	
Poland	PL	1	1	1	1	1	1				
Portugal	PT	1	1	1	1	1	1	1			
Romania	RO										
Slovenia	SI										
Slovakia	SK	1	1	1	1	1	1		1		
Finland	FI	1	1	1	1	1	1	1	1	1	1
Sweden	SE	1	1	1	1	1	1	1			
United Kingdom	UK	1	1	1	1	1	1	1	1	1	

For all years and Member States indicated with a 1, a 59-sector SAM was compiled directly based on the available data. Nevertheless, not all of the used datasets were complete and consistent. For instance, the SUT for Italy do not provide sectoral operating surplus and compensation of employees, the intermediate demand tables for Luxemburg have missing entries, and the institutional accounts for the Czech Republic are not balanced. For this reason, a balancing procedure of the thus obtained ESASAMs is also necessary.

6.1 Compilation of ESASAMs based on SUT and Institutional SAMs

This section gives an overview on the compilation of SAMs on the basis of NAIO and flows between the institutions. The description is structured according to the stylized SAMs in Table 9.

Table 9. Target Structure of ESASAMs and Correspondence to NAIO datasets

		NACE3	CPA3	FCTR	TRD	TNF	DIN	ROW	SIA	
Activities	NACE3		p1							Revenues
Commodities	CPA3	p2			p118_r		p3	p6	p5	
Primary factors	FCTR	d1		InstSAM						
Trade and transport margins	TRD		p118_p							
Direct and indirect taxes, transfers	TNF	d29_m_d39	d211, d212, d214, d31							
Domestic institutions	DIN									
Rest of the world	ROW		p7							
Saving-Investment account	SIA									
		Expenditures								

Domestic intermediate demand and domestic production by activity (D) are computed as:

$$
\begin{aligned}
ESASAM_{C,A} &= naio\_cp16_{CPA3,NACE3} \\
ESASAM_{A,C} &= naio\_cp15_{CPA3,NACE3}
\end{aligned}
\tag{57}
$$

Where:

ESASAM:	SAM based on ESA data
naio_cp16:	ESA95 use table in current prices
naio_cp15:	ESA95 supply table
CPA3:	CPA commodity accounts (at 3-digit level)
NACE3:	NACE activity accounts (at 3-digit level)
C:	ESASAM commodity account (corresponds with CPA3)
A:	ESASAM activity account (corresponds with NACE3)

Payments of domestic activities to primary factors are provided by the use tables, from which only "compensation of employees" (d1 in ESA notation) and "operating surplus, gross" (b2g_b3g) are used for the ESASAMs. "Consumption of fixed capital" (k1) and "net operating surplus" (b2n_b3n) are not distinguished at this stage, mainly due to the limited information available for the disaggregated agricultural sub-sectors.

(58) $$ESASAM_{FCTR,NACE3} = naio\_cp16_{FCTR,NACE3}$$

Where:

FCTR:	Index for primary factors:
d1:	Compensation of employees
b2g_b3g:	Operating surplus, gross

Taxes and subsidies on production are only available as net-values in NAIO:

(59) $$ESASAM_{\text{"T\_d29\_m\_d39"},NACE3} = naio\_cp16_{\text{"d29\_m\_d39"},NACE3}$$

Where:

T_d29_m_d39:	Other net taxes on production (in ESASAM format)
d29_m_d39:	Other net taxes on production (in SUT format)

Trade margins on commodities are computed in a similar manner, while taxes on commodities are taken from the estimation outlined in section 5.

(60) $$ESASAM_{\text{"T\_p118"},C} = naio\_cp15_{CPA3,\text{"p118"}}$$
$$ESASAM_{\text{"T\_d21\_m\_d31"},C} = naio\_cp15_{CPA3,\text{"d21\_m\_d31"}}$$

Where:

p118:	Trade and transport margins (in ESA format)
d21_m_d31:	Taxes less subsidies on products (in ESA format)
T_p118:	Trade and transport margins (in ESASAM format)
T_d21_m_d31:	Taxes less subsidies on products (in ESASAM format)

Imports and exports in NAIO are distinguished by the direction of trade to and from either MS or third countries. This distinction is maintained here, as it can serve as a benchmark for the trade-balances at a later stage.

(61) $$ESASAM_{XINS,C} = naio\_cp15_{CPA3,XINS}$$
$$ESASAM_{C,XINS} = naio\_cp16_{CPA3,XINS}$$

Where:

M:	Imports
E:	Exports
XINS:	Index for external institutions (trade partners):
I_s21	European Union
I_s22	Third countries and international organisations
I_s2	Rest of the world

Domestic demand for commodities is derived similarly:

(62) $$ESASAM_{C,DINS} = naio\_cp16_{CPA3,DINS}$$

Where:

DINS:	Index for domestic institutions (in ESASAM format):
I_s11_s12:	Financial and non-financial corporations
I_s13:	General government
I_s14_s15:	Households
I_s15:	Non-profit institutions serving households
I_s14_S15:	Households; non-profit institutions serving households
I_p51:	Gross fixed capital formation
I_p52_p53:	Changes in inventories
I_p5:	Gross capital formation

With the computation of domestic consumption, the accounts for activities and commodities are closed. It remains to derive the monetary flows between institutions, like taxes, transfer payments, distribution of factor incomes, and so on. For this, the NASA dataset mentioned in section 3.3 is used.

Taxes and transactions paid and received from and by the institutions are the following:

(63) $$ESASAM_{TRNS,INST} = InstSAM_{TRNS,INST}$$
$$ESASAM_{INST,TRNS} = InstSAM_{INST,TRNS}$$

Where:

INST:	Index for institutions (XINS $\cup$ DINS)
TRNS:	Index for taxes and transactions:
T_p118:	Trade and transport margins
T_d4:	Property income
T_d5:	Current taxes on income, wealth, etc.
T_d6:	Social contributions and benefis
T_d7:	Other current transfers
T_d8:	Adjustment for change in net equity

The distribution of factor income and the factor payments from abroad are computed as follows:

(64) $$ESASAM_{FCTR,XINS} = InstSAM_{XINS,FCTR}$$
$$ESASAM_{INST,FCTR} = InstSAM_{INST,FCTR}$$

And finally, the system is closed by introducing the savings of domestic and non-domestic institutions from the InstSAMs:

(65) $$ESASAM_{"I\_p5",INST} = InstSAM_{"I\_p5",INST}$$

6.2 Indirect Taxes on Commodities and Activities

Commodity taxes and subsidies are relevant political instruments and information about the applicable rates and actual payments for the different commodity groups is essential for model-based analysis of alternative tax regimes. For instance, simulations of income-neutral shifts from commodity-based subsidies to direct transfers to farming enterprises (decoupling), as currently ongoing in the context of the CAP reform, require information on the amount of subsidies paid on commodity markets, either for a given year or for a reference period. Although this data is potentially available from various sources, including CAPRI, it is not necessarily embedded in a SAM structure and thus potentially inconsistent with other data relevant in this context, like other balance items for the agricultural sector. NAIO contains a vector of net-taxes on activities and commodities (d29_m_d39, d21_m_d31) as auxiliary component for the transformation of producer into basic prices and of basic prices into purchaser's prices. Other datasets (NASA, NAMA) include totals for the different tax types, without identifying the sources for those payments. These four types of taxes and subsidies on commodities (d2211, d212, d214, d31, see Appendix 2) and two on activities (d29, d39) had to be split across the commodity groups, based on the net-taxation vectors in NAIO and additional information on applicable value added tax rates and import tariffs (DG–TAXUD (2005), WTO (2007)). The obtained average tax rates (txravr) were used to populate the respective sub-matrices, for which only row-totals (from the macro-totals, see section 5) and column-totals (the net-taxes) are known. The underlying market transactions (Z) for which the taxes are paid were identified based on the definitions in Appendix 2. VAT (or d211 in ESA nomenclature) for instance is not only levied on final consumption (position p3 in the ESA nomenclature, see also section 5), but also paid by firms that are not entitled to deduct VAT on their intermediate inputs (p2) or investments (p5). Consequently, the market transaction for VAT is the sum of all three positions (p2+p3+p5). Import duties are levied on imports (p7), while subsidies are paid for total domestic production (p1), and so on. For indirect taxes on activities, Z was simply defined as the total activity output. For each commodity C and activity A and each tax type (TXCM for taxes on commodities, TCAC for taxes on activities), the prior value for the tax payments in the ESASAM was computed as follows:

$$
\begin{aligned}
ESASAM_{TXCM,C} &= txr_{AVR,TXCM} Z_C \\
ESASAM_{TXAC,A} &= txr_{AVR,TXAC} Z_A
\end{aligned}
\tag{66}
$$

6.3 Balancing the ESASAMs

The datasets used at this stage are in general consistent and the ESASAMs are in most cases balanced. However, there are few cases where small deviations between row- and column-sum of the ESASAMs could be observed. In order to get a set of balanced SAMs as control-totals for the target AgroSAMs, the balancing of the ESASAMs has to be ensured before entering the next stage of the compilation procedure. For this, the cross-entropy procedure described in section 4 is used, which balances the SAMs by employing a multiplicative error term with an expected value of 1 and a range sufficiently large to accommodate possibly high deviations between row- and column-sums of the ESASAMs. The error term is defined by a set of **s** support points and associated weights. The support points are arbitrarily defined as shown below for the case of five support points:

(67)
$$b_s = [-3,3]$$
$$\bar{W}_s = 0.5$$

The weights have to add up to unity and should be as close as possible to a set of pre-defined prior weights, for which a uniform distribution is assumed.

Information on net-indirect taxes on activities and commodities (NetTAX) was incorporated in the estimation procedure as a stochastic control total on the respective sub-matrices. Stochastic control totals means in this context that a distortion term was associated with the control-vectors.

The objective function of the balancing model is to minimize the cross-entropy between prior weights and final weights. The minimization is subject to the constraint that the weights range between 0 and 1, add up to unity and that the final ESASAMs are as close as possible to the prior SAMs derived from the NASA and NAIO datasets, but has equal row- and column-sums. The balancing model is summarized in equation (68)[7].

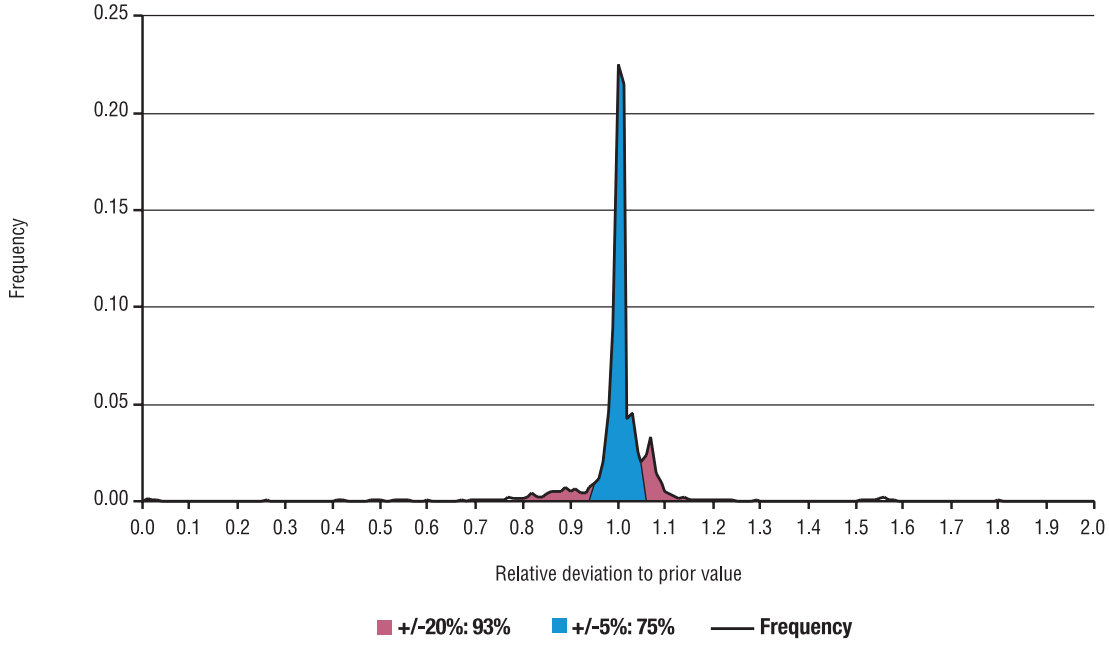
(68)
$$CE = \min \sum_{AC,AC'} \sum_{s} W_{AC,AC',s} \cdot \ln\left[W_{AC,AC',s} / \bar{W}_s\right] + \sum_{AC} \sum_{s} W^{NTX}_{AC,s} \cdot \ln\left[W^{NTX}_{AC,s} / \bar{W}^{NTX}_{AC,s}\right]$$
$$s.t.$$
$$ESASAM_{AC,AC'} = \overline{ESASAM}_{AC,AC'} \cdot \exp\left[\sum_{s} W_{AC,AC',s} \cdot b_s\right]$$
$$\sum_{s} W_{AC,AC',s} = 1, \quad 0 \le W_{AC,AC',s} \le 1$$
$$\sum_{AC} ESASAM_{AC,AC'} = \sum_{AC'} ESASAM_{AC,AC'}$$
$$\sum_{AC'} \left[\sum_{AC} G_{MKR,AC} \cdot ESASAM_{AC,AC'}\right] \cdot G_{MKR',AC'} = InstSAM_{MKR,MKR'}$$
$$\sum_{TXIN} ESASAM_{TXIN,AC} = NetTAX_{AC} \cdot \exp\left[\sum_{s} W^{NTX}_{AC,s} \cdot b^{NTX}_s\right]$$

Where:

ESASAM:	Balanced ESASAM
$\overline{ESASAM}$:	Prior ESASAM derived from SUT and NASA datasets
W:	Weights of error support points on ESASAM entries
NetTAX:	Control vectors on net-indirect taxes on commodities and activities (d21_m_d31 and d29_m_d39)
W^{NTX}:	Weights of distortion term support points on net indirect taxes
b, b^{NTX}:	Support points for error weights
G:	Aggregator matrix
InstSAM:	Institutional SAM, see section 5.2
CE:	Cross-entropy minimand
s:	Index for support points

7 Indices for time and state have been omitted in order to improve the readability of the equations

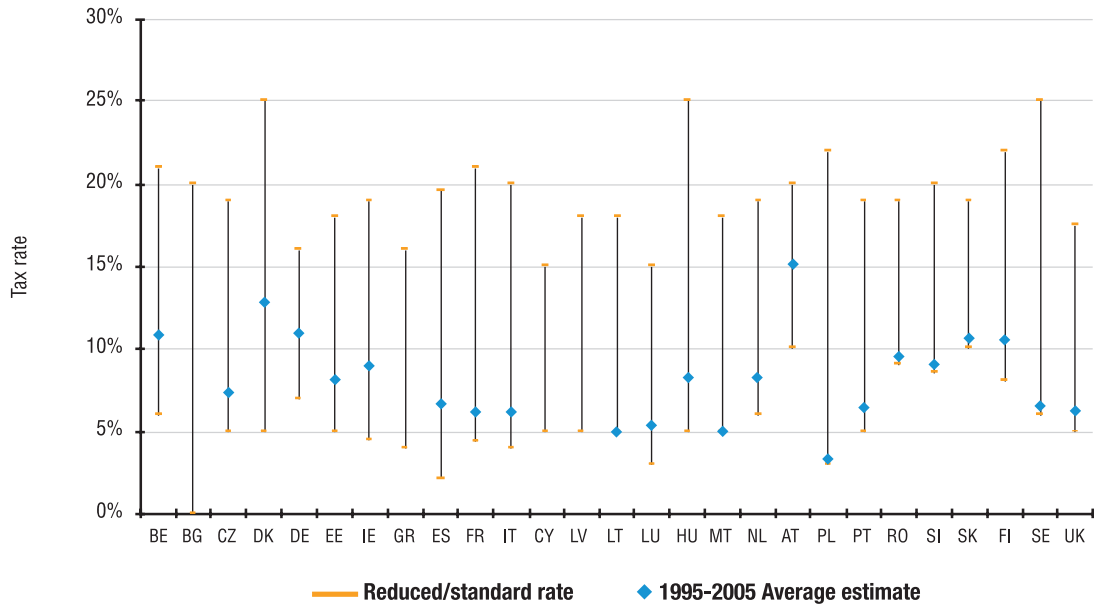
Figure 14. Relative Deviations between Prior and Final ESASAM



The frequency of relative deviations between compiled prior and finally estimated ESASAM are depicted in Figure 14. In 75% of cases, the final values deviate by plus or minus 5% from the prior data, while 93% deviate by 20%. In general, this result was considered as satisfying, particularly when considering the sometimes severe discrepancies between the combined datasets, the need to fill blanks with share coefficients from other years or even other countries, and the lack of information on indirect taxes.

With respect to the indirect tax rates, Figure 15 shows the reported average value-added tax rates for the EU-27 Member States.

Figure 15. Value-Added Tax Rates: Reported Rates and Estimated Averages



Notes: 1) Reduced rate for Denmark (DK) was set to 5%

7. Compilation of Priors for AgroSAMs (PriorSAM)

The objective of estimating a reliable, balanced social accounting matrix with disaggregated agricultural and food industry sectors depends largely on the reliability of the a priori information drawn from the various sources. The compilation of the prior SAM should hence be carried out in a careful and transparent manner. Particularly challenging is to perform in a transparent way the necessary re-arrangement of entries in the parent datasets, in order to achieve the required compatibility of formats and contents. GAMS code was developed and adjusted whenever new challenges occurred in the process of including more countries, years or datasets. In general, a four-step procedure was followed:

1. Re-arrange the CAPRI data into the SAM format (agricultural accounting matrix AAM)
2. Merge ESASAM and AAM into an unbalanced PriorSAM
3. Balance activity and commodity account totals
4. Balance the PriorSAM

Table 10 provides an overview on the sources used to obtain a priori information for the AgroSAM. Since the food-industry sector is not covered exhaustively either in ESA or in CAPRI[8], here it is necessary to incorporate other sources of information as well, like the database on Products of the European Community (PRODCOM).

Table 10. Sources of Prior Information for the Agricultural and Food Processing Industries in the AgroSAMs

Description	Code	Preferred source	Second best source
Domestic output by sectors	**D**	CAPRI	EAA/AGRI_IS/PRODCOM
Intermediate demand	**I**	CAPRI	Qualitative Prior/Estimate
Domestic final consumption	**C**	CAPRI	EAA/AGRI_IS
Exports	**E**	CAPRI	EAA/AGRI_IS/TRADEX
Domestic factor payments (value added)	**Fd**	ESA/CAPRI	
Factor revenues from abroad	**Fe**	ESA	
Trade margins	**H**	ESA	
Taxes and subsidies on production	**Ta**	CAPRI	Estimate
Taxes and subsidies on products	**Tc**	ESA	Estimate
Direct taxes paid by institutions	**Ti**	ESA	
Distribution of factor income across institutions	**F**	ESA	
Distribution of taxes and transfers across institutions	**T**	ESA	
Imports	**M**	CAPRI	EAA/AGR_IS/TRADEX
Savings of institutions	**S**	ESA	
Total domestic production value	**VX**	CAPRI/ESA	AGR_IS
Total domestic absorption	**VQ**	CAPRI/ESA	AGR_IS

8 This might change in the near future, since the CAPRI Model is currently improving processing functions for dairies, oilseeds and biofuels in its market module, information which will most likely be fed into the base year database and used in the construction of a future PriorSAM.

7.1 Compilation of an Agricultural Accounting Matrix based on CAPRI Data

The CAPRI database builds upon the meta-database of the NewCRONOS domain manager of EUROSTAT (sub-domains: ZPA1, COSA, PRAG). Although the raw data is processed to meet the demand for completeness and consistency (Britz and Witzke (2008)), it still follows the general accounting principles of the EAA. This "data massaging" property creates serious difficulties when attempting to combine the data with data in ESA95 format (e.g. SUT) as the distinction between an agricultural commodity and a processed commodity is not done in the same way. For instance, wine is considered as an agricultural commodity in EAA but as a processed output of the "beverage industry" in the ESA95 framework (EuroStat (1997)). Furthermore, it is unclear how "feed cereals" should be mapped into the AgroSAM format: either as non-marketed on-farm consumption or as demand of the agricultural sector for products of the "animal feed" industry, which in turn would demand cereals as intermediate input. The mapping of farm and market balances is also not straightforward. For instance, seed use, internal use and losses on farm are not considered in the ESA framework.

The first step in utilizing the CAPRI dataset was to transform it into an agricultural accounting matrix (AAM) to facilitate the mapping of CAPRI and MAC accounts at a later stage. The AAM distinguishes strictly between activity and commodity accounts and agricultural, processing, and other industrial activities. As a consequence, it was necessary to introduce activity accounts not included in the CAPRI database. The commodity 'beef' for instance is produced by the cattle sectors, which is not consistent with the concept of the ESA accounts. In there, the transformation of live cattle into beef ready for human consumption or further processing is an activity within the food and beverage industry complex (ESA code da15) rather than belonging to agriculture. The same applies for pork, poultry meat, and wine.

An important feature in this context is that basically two agricultural accounting matrices were created: One in value terms (AAMV) and one in quantity terms (AAMQ). AAMQ is basically a balance sheet for CAPRI commodities, arranged in SAM format, but with empty accounts for activity expenditures and consequently only with balanced commodity accounts. AAMV is the corresponding matrix with filled activity accounts and quantities on the commodity markets measured at basic prices obtained from CAPRI (Unit Value at Basic Prices, UVAB). This treatment of the available data allows controlling the estimates for prices and quantities at a later stage and prevents the creation of un-plausible values, which can occur when using only value-data for the SAM estimation. An outline of the operations to obtain the AAM from the CAPRI dataset is also displayed in Table 11.

7.1.1 Activity Accounts of the Agricultural Sector

For the agricultural sector, the procedure of re-arranging the CAPRI data is in general straightforward. In the following, the CAPRI notation is used whenever possible to allow the comparison of the computations with the CAPRI documentation (Britz and Witzke (2008)). Starting with the activity accounts, the first step is the derivation of an aggregate output value of each agricultural activity:

Table 11. Summary of Operations from CAPRI Database to AAM

			Activities		$\sum$	Commodities		$\sum$		Institutions				
			A_AGR	A_NAG						I_GVT	I_HHD	I_STC	I_FCF	I_RoW
Activities	Agriculture	A_AGR	$\sum_C[LOSF_C\text{-}SEDF_C\text{-}INTF_C] \cdot ASHR_{C,A}$			$LEVL_A \cdot IO_{A,C}$ - $LOSF_C$ - $SEDF_C$ - $INTF_C$		$LEVL_A \cdot TOOU_A$						
	Non-agriculture	A_NAG					$MAPR_C \cdot ASHR_{C,A}$							
	Total	$\sum$				$NETF_C$, MAPR								
Commodities	Agriculture	C_AGR	$LEVL_A \cdot IO_{C,A}$,	$NETF_C$ / $ASHR_{C,A}$	$INDM_C + PRCM_C + FEDM_C + SEDM_C$						$HCOM_C$	$STCM_C$		$EXPT_C$
	Non-agriculture (including processed commodities)	C_NAG	$LEVL_A \cdot IO_{C,A}$											
	Total	$\sum$	$LEVL_A \cdot TOIN_A$											
Factors	Labour	F_LAB	$LEVL_A \cdot WAGE_A$											
	Capital	F_CAP												
	Total	$\sum$	$LEVL_A$ * $GVAP_A$											
	Trade	T_TRD				-$LOSM_C$ – $SADM_C$								
	Taxes	T_TAX	$LEVL_A$ * $PRME_A$											
…	…	…												
	Rest of the world	I_RoW				$IMPT_C$								
		Total	$LEVL_A$ * $TOOU_A$											

(69) $$AAMV_{'Total',A} = ALV_A = TOOU_A \cdot LEVL_A \qquad \forall A \in Agriculture$$

Where:

AAMV: Agricultural accounting matrix in value terms based on CAPRI data
ALV: Total value of activity level
TOOU: Total output value per activity level at producer prices (CAPRI)
LEVL: Activity level (CAPRI)
A: Index for activities (here only agriculture)

Taxes paid (or received as negative taxes i.e. subsidies) by each activity equal the CAP premiums per activity as indicated by CAPRI times the activity level:

(70) $$AAMV_{'T\_PRME',A} = TXA_A = -PRME_A \cdot LEVL_A \qquad \forall A \in Agriculture$$

Where:
TXA: Value of tax or subsidy received or paid by activity
PRME: CAP premium effectively paid (CAPRI)

The rate for activity-related taxes is here computed as the share of taxes paid (or subsidies received) in the total output value of the activities, $ta_A = \frac{TXA_A}{ALV_A} \quad \forall A \in Agriculture.$

Value added at basic prices can also be taken directly from CAPRI:

(71) $$\sum_F AAMV_{F,A} = VAD_A = GVAB_A \cdot LEVL_A \qquad \forall A \in Agriculture$$

Where:

VAD: Value-added per activity
GVAB: Gross value-added per activity level at basic prices (CAPRI)
F: Fixed factors (here: labour and capital)

A wage indicator is also provided in the CAPRI database, but the exact unit in which they are measured is not explained in the documentation (Britz and Witzke (2008)). Furthermore, it is not clear whether this information was processed by the consistency algorithm of CAPRI. However, in the absence of other data, WAGE was used as an instrument for the distribution of the corresponding entry in the ESASAM.

(72) $$AAMV_{'F\_LAB',A} = LAB_A = \frac{WAGE_A \cdot LEVL_A}{\sum_A WAGE_A \cdot LEVL_A} \cdot LAB^{ESA}_{Agriculture} \qquad \forall A \in Agriculture$$

Aggregate input demand from agricultural activities is expressed as input demand per unit of activity level times the activity level.

(73) $$\sum_{C} AAMV_{C,A} = IDA_A = TOIN_A \cdot LEVL_A \qquad \forall A \in Agriculture$$

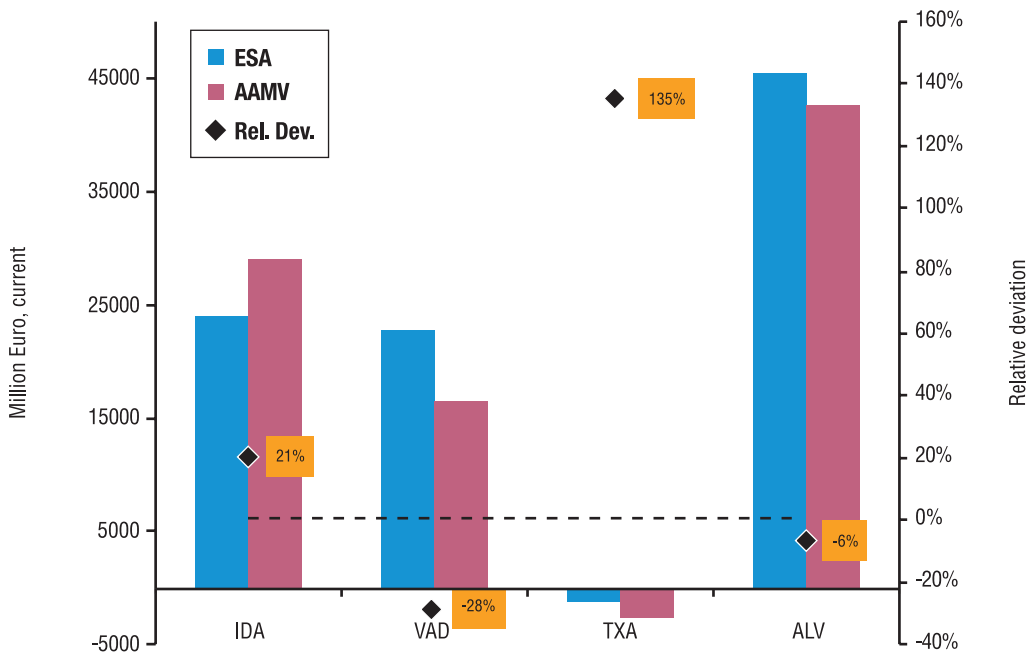
Where:

IDA: Vector of aggregate input demand per activity (in million Euros)
TOIN: Total intermediate input at producer prices (CAPRI)

The results for the agricultural sector are displayed against the corresponding ESA totals in Figure 16. It appears that intermediate demand of the agricultural sector as obtained from the CAPRI database is 21% larger than the corresponding figure from the ESA accounts. The reason behind this could be that CAPRI provides values also for non-marketed inputs like pastures and manure. The higher total output value indicated by ESA may originate in the fact that agricultural output encompasses a wider range of products as are considered by CAPRI. Taxes on activities ("Other net taxes on production", d29_m_d39, in ESA notation) indicated by ESA are considerably lower (in absolute terms) than the aggregate CAP Premiums from CAPRI. Again, the reason for this observation is not clear, since details on the composition of the figures in question are not provided by either source. It seems anyway that some components of the CAP Premiums are booked as direct subsidies to agricultural holdings in the ESA framework rather than as activity-related payments in the CAPRI database.

Although the two databases present substantial differences in the definition and coverage of featured items, without clear information on the exact nature of those deviations, a multitude of components of the CAPRI database can be considered as reliable information (e.g. produced and trade quantities of

Figure 16. Comparison between ESA and AAMV Totals, Agricultural Sector, Germany 2000, in million Euro, current



Source: EuroStat, CAPRI, own calculations

agricultural and some processed commodities, activity levels, output- and input-coefficients, and basic prices). Both databases can be harmonized by incorporating the qualitative information about the potential sources of the deviations in the finally chosen estimation method.

Having derived IDA, VAD, TXA, and ALV, the minimum necessary set of items in the activity accounts was obtained to connect it to the corresponding commodity accounts.

7.1.2 Commodity Accounts

The CAPRI database provides information on quantities of produced and trade commodities as well as the related prices. This information is deemed to be of significant use for the final estimation of the monetary flows within the target AgroSAMs, since the usage of quantities and plausible bounds on price estimates can be used to curb the possible variation of the final estimate and hence avoid severely distorted results.

Starting point is the transformation of the quantity-related data of the CAPRI database into SAM format, which will be called AAMQ (Agricultural Accounting Matrix in quantity terms) in the following. Again, CAPRI notation is used whenever possible to allow the comparison of the computations with the CAPRI documentation (Britz and Witzke (2008)).

Domestic marketed production quantities QX are computed by:

(74) $$\sum_{A} AAMQ_{A,C} = QX_C = \begin{cases} NETF_C \ \forall C \in Agriculture \\ MAPR_C \ \forall C \in FoodIndustry \end{cases}$$

Where:

AAMQ	Agricultural accounting matrix in quantity terms based on CAPRI data
QX	Domestic marketed production (quantities)
NETF	Net trade on farm (CAPRI)
MAPR	Marketed production (CAPRI)
C:	Index for commodities

Imports and exports are derived in a similar way:

(75) $$\begin{aligned} AAMQ_{C,'ROW'} &= QE_C = EXPT_C \ \forall C \in Agriculture, FoodIndustry \\ AAMQ_{'ROW',C} &= QM_C = IMPT_C \ \forall C \in Agriculture, FoodIndustry \end{aligned}$$

Where:

QE	Exports of commodities (quantities)
QM	Imports of commodities (quantities)
EXPT	Exports total (CAPRI)
IMPT	Imports total (CAPRI)

Total domestic supply QDS is composed of domestic production QX plus imports QM minus exports QE. On the demand side, the items IDC (domestic intermediate demand for commodities; note the difference to IDA which is the intermediate demand for commodities by activities), GVT (governmental consumption), H (final consumption by households), STC (stock changes), FCF (fixed capital formation), and LOS (losses on markets) can only partially be derived from the CAPRI. So is investment demand for agricultural commodities treated as "on-farm usage" of investment commodities like young animals and live plants (e.g. trees for orchards), but not as consumption on markets. Domestic demand in the AAMQ as derived from CAPRI data is consequently represented by the following entries:

(76) $$QDD_C = QX_C + QM_C - QE_C = QIDC_C + QH_C + QSTC_C + QLOS_C$$
$$\forall C \in Agriculture, Foodindustry$$

With the following correspondence to CAPRI data:

Intermediate consumption:

(77) $$\sum_A AAMQ_{C,A} = QIDC_C = INDM_C + PRCM_C + FEDM_C + SEDM_C$$

Household consumption:

(78) $$AAMQ_{C,'I\_HHLD'} = QH_C = HCOM_C$$

Stock changes:

(79) $$AAMQ_{C,'I\_STCH'} = QSTC_C = STCM_C$$

Losses:

(80) $$AAMQ_{C,'T\_TRD'} = QLOS_C = LOSM_C + SADM_C$$

Where:

QDD	Domestic absorption
QIDC	Intermediate demand per commodity
QH	Household final consumption per commodity
QSTC	Stock changes
QLOS	Losses on markets
INDM	Industrial use market (CAPRI)
PRCM	Processing to derived products market (CAPRI)
FEDM	Feed use on market (CAPRI)
SEDM	Seed use on market (CAPRI)
HCOM	Human consumption market (CAPRI)
STCM	Stock changes on market (CAPRI)
LOSM	Losses on market (CAPRI)
SADM	Statistical adjustment on market (CAPRI)

Losses on markets are here booked in the account for transaction costs and will serve at a later stage as proxy for the estimation of commodity-specific trade margins in the AgroSAM.

7.1.3 Input and Output Matrices

Accounts for activities and commodities are linked via two sub-matrices, the intermediate input use and the output-by-activity tables (fields I and D, respectively, in Table 13). CAPRI provides information about the intermediate demand of the agricultural sector in value terms (e.g. pharmaceutical inputs or energy in constant 1995 Euro/ha) and in quantity terms (fertilizer in kg/ha), while outputs (or yields) of each agricultural activity are recorded as quantities (kg/ha).

The sub-matrix for domestic output can be therefore fully derived by multiplying the output coefficient with the activity levels:

(81) $$AAMQ_{A,C} = QD_{A,C} = OUTP_{C,A} \cdot LEVL_A$$

Where:

QD: Domestic production quantity by activity
OUTP: Output coefficient (CAPRI)

The input matrix on the other hand has two representations:

(82) $$\begin{aligned} AAMQ_{C,A} &= QI_{C,A} = INPT_{C,A} \cdot LEVL_A \quad \forall INPT_{C,A} \text{ measured in quantities} \\ AAMV_{C,A} &= VI_{C,A} = INPT_{C,A} \cdot LEVL_A \quad \forall INPT_{C,A} \text{ measured in values} \end{aligned}$$

Where:

QI: Domestic intermediate demand quantity by activity
VI: Domestic intermediate demand value by activity
INPT: Input coefficient

7.1.4 Splitting Agriculture and Food Industry

One of the main challenges when attempting to harmonize the CAPRI database with the SUT in ESA format is the fundamental difference in the treatment of processed agricultural commodities. These are part of the agricultural sector in the EAA (and consequently CAPRI) framework, but belong to the food processing industries in the ESA framework.

For this reason, a new set of auxiliary activities was introduced in the SAM while processing the CAPRI data. These correspond with the agricultural outputs in CAPRI considered as processed commodities in the AgroSAM framework, particularly beef, pork, sheep and goat meat, and wine. The domestically produced quantities are here mapped to the corresponding industrial activities:

(83) $$AAMQ_{A,C} = QD_{A,C} = QX_C \cdot PRCOUT_{A,C} \qquad \forall A,C \in FoodIndustry$$

Where:

PRCOUT: Binary aggregator matrix (1 if activity A produces commodity C, else 0)

Since these products are not anymore considered as outputs of the agricultural sector, the corresponding entries under agricultural activities have to be set to 0:

(84) $$AAMQ_{A,C} = QD_{A,C} = 0 \qquad \forall A \in Agriculture, C \in FoodIndustry$$

A similar approach is chosen for the input demand. The new activity "beef meat" (A_BEEF) demands slaughtered animals from the agricultural sector, the activity "Rice milled" demands paddy rice, and so on:

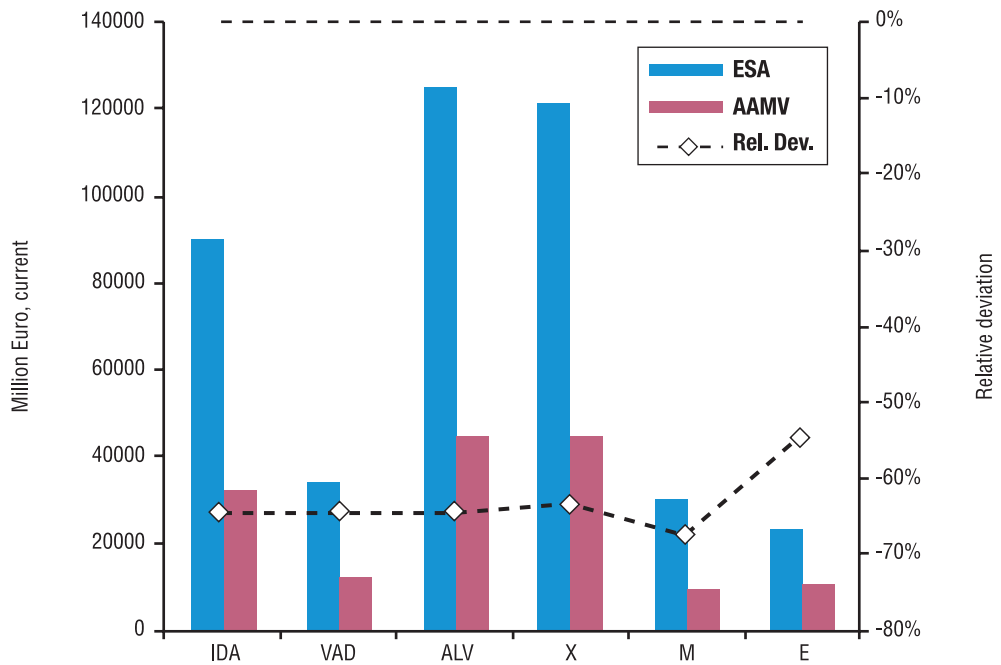
(85) $$AAMQ_{C,A} = QI_{C,A} = QX_C \cdot PRCINP_{C,A} \qquad \forall A \in FoodIndustry, C \in Agriculture$$

Where:

PRCINP Binary aggregator matrix (1 if activity A demands commodity C, else 0)

The thus derived values are compared with the ESA totals in Figure 17. It appears that, in contrast to the agricultural sector, the food industry sector is only represented to a limited extent in the CAPRI database and consequently in the agricultural accounting matrix. This issue will be addressed when compiling a prior SAM in section 7.2.

Figure 17. Comparison between ESA and AAMV Totals, Food Industry Sector, Germany 2000, in million Euro, current



Source: EuroStat, CAPRI, own calculations

Table 12. New Activities and Corresponding CAPRI Commodities

CAPRI Commodities	Code	New processing activities	
Rice milled	C_RICE	A_RICE	Rice milled
Molasse	C_MOLA	A_SUGA	Processed sugar
Starch	C_STAR	A_STAR	Starch
Processed sugar	C_SUGA	A_SUGA	Processed sugar
Rape seed oil	C_RAPO	A_RAPO	Rape seed oil
Sunflower seed oil	C_SUNO	A_SUNO	Sunflower seed oil
Soya oil	C_SOYO	A_SOYO	Soya oil
Olive oil	C_OLIO	A_OLIO	Olive oil
Other oil	C_OTHO	A_OTHO	Other oil
Rape seed cake	C_RAPC	A_RAPO	Rape seed oil
Sunflowe seed cake	C_SUNC	A_SUNO	Sunflower seed oil
Soya cake	C_SOYC	A_SOYO	Soya oil
Olive cake	C_OLIC	A_OLIO	Olive oil
Other cake	C_OTHC	A_OTHO	Other oil
Raw milk at dairy	C_MILK	A_MILK	Raw milk at dairy
Butter	C_BUTT	A_MILK	Raw milk at dairy
Skimmed milk powder	C_SMIP	A_MILK	Raw milk at dairy
Cheese	C_CHES	A_MILK	Raw milk at dairy
Fresh milk products	C_FRMI	A_MILK	Raw milk at dairy
Cream	C_CREM	A_MILK	Raw milk at dairy
Concentrated milk	C_COCM	A_MILK	Raw milk at dairy
Whole milk powder	C_WMIP	A_MILK	Raw milk at dairy
Beef meat	C_BEEF	A_BEEF	Beef meat
Pork meat	C_PORK	A_PORK	Pork meat
Sheep and goat meat	C_SGMT	A_SGMT	Sheep and goat meat
Poultry meat	C_POUM	A_POUM	Poultry meat
Fodder rich protein	C_FPRO	A_ANFD	Animal feed
Fodder rich energy	C_FENE	A_ANFD	Animal feed
Fodder other	C_FOTH	A_ANFD	Animal feed

7.1.5 Prices and Values

The ESA SUT distinguishes between basic prices and purchaser's prices. However, as the CAPRI data do not provide a conversion from basic prices to purchaser's prices (as done in ESA SUT with vectors for trade margins (p118) and net-taxes on commodities (d21_m_d39), only unit values at basic prices (UVAB) were used at this stage to determine starting values for domestic, import, and export prices. In case they were not available from CAPRI for certain processed commodities (e.g. oilcakes or molasses), it was necessary to rely on other sources, among which FAOSTAT appeared to be the most appropriate for the commodity groups distinguished in CAPRI. It should be emphasized at this stage that the prices entering the following computations are best first guesses, which will be altered in the subsequent balancing steps.

The starting values for basic prices are:

(86) $$PB_C = \begin{cases} UVAB_C \ \forall C \in Agriculture, FoodIndustry \\ FAOPRIC_C \ if \ not \ UVAB_C \end{cases}$$

Where:

PB	Starting values for basic commodity prices
UVAB	Unit value at basic prices (CAPRI)
FAOPRIC	Prices for processed commodities from FAOSTAT

With this price vector at hand, it is now possible to connect the two agricultural accounting matrices into a common format (AAM):

(87) $$AAM_{AC,AC'} = \begin{cases} AAMV_{AC,AC'} \ if \ AAMV_{AC,AC'} \\ AAMQ_{AC,AC'} \cdot PB_{AC} \ if \ not \ AAMV_{AC,AC'} \end{cases}$$

Where:

AAM	Agricultural accounting matrix in basic prices

7.2 Prior AgroSAM

The next step is to combine ESASAM and AAM into a prior AgroSAM, which will enter the balancing procedure later on. First, the AAM were aggregated into the target classification. It has to be noted again, that the AAM is measured at basic prices while the ESASAM are a mixture of basic and purchaser's prices. It is therefore not possible to merge directly all accounts of the agricultural and food-industry sectors, but only those which are given in basic prices within the ESASAM. These accounts are mainly domestic production and imports of commodities, but also total activity output and total intermediate demand. Exports and domestic use accounts are measured in purchaser's prices and the AAM data cannot be used directly. Instead, the row-shares of each commodity account were used to disaggregate the respective account total. However, the population of the agricultural and food-commodity accounts based on AAM data comes at the end of a rather lengthy compilation procedure, which is outlined in the subsequent section.

7.2.1 Deriving Row- and Column Totals

Deriving prior values for row- and column totals of the agricultural and food commodity and activity accounts is a first and crucial step in the compilation of the AgroSAM prior as these values represent the boundary conditions for all transactions within the respective accounts. To ensure consistency with the AAM, these tables were aggregated into the target classification (AAM2), and the row-totals were calculated. For readability purposes, from now on activity and commodity accounts of the AgroSAM will be denoted A and C, respectively, whereas the accounts of AAM will be indicated with C_COC and A_COC respectively.

(88) $$AAM2_{AC,AC'} = \sum_{AC \in AC\_COC} \sum_{AC' \in AC\_COC'} AAM_{AC\_COC,AC\_COC'} \quad \forall AC \in Agriculture, FoodIndustry$$

Table 13. Target Structure of the Agricultural Accounting Matrices

				Activities	Σ	Commodities	Σ		Institutions				
	Activities	Agriculture	A_AGR			D	ALV						
		Industry	A_IND										
		Services	A_SER										
		Total	Σ			X							
	Commodities	Agriculture	C_AGR	I	IDC				GVT	H	STC	FCF	E
		Industry	C_IND										
		Services	C_SER										
		Total	Σ	IDA									
	Factors	Labour	F_LAB	LAB									
		Capital	F_CAP	CAP									
		Total	Σ	VAD									
		Trade	T_TRD			hm							
		Taxes	T_TAX	ta									
	…	…	…										
		Rest of the world	I_RoW			M							
			Total	ALV									

For the agricultural sector, it is assumed that it is completely covered by the data used in the CAPRI model, whereas the food industry sector only to some extent, as "other food" and "beverage" industries are not part of the AAM. To generate a complete set of row- and column totals, additional information on these sub-sectors is used from PRODCOM datasets.

Because of the substantial deviation between ESA and AAM values in the food processing sector and because of the fact that some accounts are not covered completely by the CAPRI database ('other food products', 'beverages'), it was necessary to include additional information from PRODCOM and COMEXT datasets. The account totals were derived as follows:

Domestic production:

(89) $$VX_C = \begin{cases} \sum_A AAM2_{A,C} & if\ AAM2_{A,C} > 0 \\ PRODCOM_C & if\ AAM2_{A,C} = 0 \end{cases}$$

Imports:

(90) $$VM_C = \begin{cases} AAM2_{'I\_ROWD',C} & if\ AAM2_{'I\_ROWD',C} > 0 \\ COMEXT_C & if\ AAM2_{'I\_ROWD',C} = 0 \end{cases}$$

Where:

I_ROWD: Index for trade partners (Rest of World)

Domestic production (VX) and imports (VM) were then added up to derive total commodity supply, which serves as prior for the column-totals (AgroCOLABS) of the AgroSAM:

(91) $$AgroCOL_C^{ABS} = VX_C + VM_C$$

The total activity output was derived as either the sum over all commodities produced by the activities in question, if available from AAM. For those accounts not included in AAM, the activity totals were derived by multiplying domestic production (VX) with a binary matrix PRCOUT that maps the produced commodities to the respective activities:

(92) $$AgroCOL_A^{ABS} = \begin{cases} \sum_C AAM2_{A,C} & if\ AAM2_{A,C} > 0 \\ \sum_C VX_C \cdot PRCOUT_{A,C} & if\ AAM2_{A,C} = 0 \end{cases}$$

Next, the share of each activity and commodity (AgroCOL^{SHR}) within the corresponding account in ESA format (AC_ESA) was derived:

(93) $$AgroCOL_{AC}^{SHR} = \frac{AgroCOL_{AC}^{ABS}}{\sum_{AC\_ESA} \left[\sum_{AC} AgroCOL_{AC}^{ABS} \cdot G_{AC,AC\_ESA} \right] \cdot G_{AC,AC\_ESA}}$$

Where:

G: Aggregator matrices between AgroSAM and ESASAM accounts
AC_ESA: Account in ESA format

These shares served to compute a first, default prior for the AgroSAM by expanding the ESASAM into target classification and multiplying it row- and column-wise with the derived share-vectors, as discussed in the following section.

7.2.2 Combining ESASAM and AAM

Having determined column-totals and the corresponding share vectors, it is now possible to expand the ESASAM in the following manner:

$$\overline{AgroSAM}_{AC,AC'} = AgroCOL_{AC}^{SHR} \cdot AgroCOL_{AC'}^{SHR} \cdot \left[\sum_{AC\_ESA'} \left[\sum_{AC\_ESA} G_{AC,AC\_ESA} \cdot ESASAM_{AC\_ESA,AC\_ESA'} \right] \cdot G_{AC',AC\_ESA'} \right] \quad (94)$$

Where:

$\overline{AgroSAM}$: Prior AgroSAM
G: Aggregator matrices between AgroSAM and ESASAM accounts

This represents the default setting, in the sense that in the absence of additional information, the total-shares are used to populate the unbalanced prior AgroSAM. The disadvantage of this approach becomes evident when assuming that e.g. wheat production has the highest share of all agricultural production activities and dairy commodities the highest share within food commodities. As a result, the combined shares will lead to a high value of intermediate demand of the wheat activity for dairy commodities in the intermediate demand sub-matrix (commodity C demanded by activity A), which is clearly implausible and unrealistic. However, in the absence of additional information on e.g. final consumption of food commodities, it is not implausible to assume that the commodity with the highest share in domestic production also has a high share in consumption as the aggregate output has to be consumed one way or another. The same applies for the production of agricultural commodities from e.g. the ferrous industry, for which non-zero entries may occur in the ESASAM. The reason for such entries is the fact that firms may generate more than 50% of their annual income from their main activity, thus being recorded as belonging to a certain economic branch, but having side-activities as well. In such a case, it is as likely as any other assumption that the agricultural output of these activities is composed similar to the economy-wide agricultural output-patterns.

Additional information on agriculture and food industry accounts is available from the AAM, for instance the intermediate demand of agricultural activities for food commodities (which, in reference to the example above, is always zero for intermediate demand for dairy commodities from crop producing activities). For those sub-matrices that are measured in basic prices (e.g. domestic production by activities and imports), the AAM entries can be used directly:

(95)
$$\overline{AgroSAM}_{A,C} = AAM2_{A,C} \qquad \forall C \in Agriculture, FoodIndustry$$
$$\overline{AgroSAM}_{'I\_ROWD',C} = AAM2_{'I\_ROWD',C} \forall C \in Agriculture, FoodIndustry$$

Final and intermediate consumption on the other hand are measured in basic prices in AAM and cannot be introduced directly into the prior AgroSAM as the commodity accounts are row-wise measured in purchaser's prices. Instead, the row-wise share of the commodity accounts for agriculture and food industries were multiplied with the corresponding row-totals:

(96)
$$\overline{AgroSAM}_{C,AC} = \frac{AAM2_{C,AC}}{AAM2_{C,'Total'}} \cdot \overline{AgroSAM}_{C,'Total'} \forall C \in Agriculture, FoodIndustry$$

The operations above ensure that the entries of the commodity accounts reflect the data from AAM either directly as values (commodity-columns) or at least according to the consumption shares (commodity-rows) for final and intermediate demand. For agricultural activities, subsidies on activities and column-totals were also taken directly from AAM:

(97)
$$\overline{AgroSAM}_{'T\_SBAC',A} = AAM2_{'T\_SBAC',A}; \ \overline{AgroSAM}_{'Total',A} = AAM2_{'Total',A} \forall A \in Agriculture$$

It has to be noted again that the sets 'Agriculture' and 'FoodIndustry' refer to those accounts available from AAM. This means that they exclude accounts which are part of the agricultural and food industry sectors in ESA95 classification, but are not part of the CAPRI model ('beverages' and 'other food'). For those accounts, the default prior (Equation (94)) applies.

7.2.3 Control Totals for Agriculture and Food Industries

In addition to information on specific entries in the targeted AgroSAM (like domestic outputs by activities or import values of commodities), the CAPRI dataset includes information on sub-totals of the AgroSAM like total intermediate demand (TOIN) of agricultural activities or total marketed production of agricultural and food commodities (MAPR). Furthermore, as it is foreseen to expand the trade account of the AgroSAMs with respect to additional trade partners (at least intra- and extra-EU trade) a control total on imports was also introduced. Altogether, the following control totals were considered:

(98)
$$CAPCTR_{'TOOU',A} = TOOU_A; \ LEVL_A \qquad \forall A \in Agriculture$$
$$CAPCTR_{'TOIN',A} = TOIN_A \cdot LEVL_A \qquad \forall A \in Agriculture$$
$$CAPCTR_{'MAPR',C} = MAPR_C \qquad \forall C \in Agriculture, FoodIndustry$$
$$CAPCTR_{'IMPT',C} = IMPT_C \qquad \forall C \in Agriculture, FoodIndustry$$

8. Balancing the AgroSAMs

Having constructed a set of unbalanced (a priori) AgroSAMs ($\overline{AgroSAM}$) based on the available information as described in section 7.2, the next step is to apply an estimation procedure to create a balanced set of AgroSAMs, which is as close as possible to the prior data. In principle, a procedure similar to section 6.3 was applied, where the ESASAMs were balanced with respect to macro totals, only that now the AgroSAMs are forced to add up to the corresponding values of the ESASAMs. In addition to this deterministic control totals, stochastic control totals were also included, which were derived from the CAPRI database (CAPCTR). At this point it is important to recall that the accounting schemes for ESA95 and EAA, on which the CAPRI database is mainly built, deviate to some extent. This means that is not possible to strictly enforce control totals from both datasets at the same time. As the main input and framework for the AgroSAMs are the national SUT in ESA95 format, it is necessary to associate the control totals derived from the CAPRI datasets with an error term, which enters the objective function in the same way as the error terms in the previous steps.

$$
\begin{aligned}
CE = \min & \sum_{AC,AC'} \sum_{s} W_{AC,AC',s} \cdot \ln\left[W_{AC,AC',s} / \bar{W}_{s}\right] \\
& + \sum_{CAX,AC'} \sum_{s} W^{CAPCTR}_{CAX,AC',s} \cdot \ln\left[W^{CAPCTR}_{CAX,AC',s} / \bar{W}^{CAPCTR}_{CAX,AC',s}\right] \\
s.t. & \\
& AgroSAM_{AC,AC'} = \overline{AgroSAM}_{AC,AC'} \cdot \exp\left[\sum_{s} W_{AC,AC',s} \cdot b_{AC,AC',s}\right] \\
& \sum_{s} W_{AC,AC',s} = 1, \quad 0 \le W_{AC,AC',s} \le 1 \\
& \sum_{s} W^{CAPCTR}_{CAX,AC',s} = 1, \quad 0 \le W^{CAPCTR}_{CAX,AC',s} \le 1 \\
& \sum_{AC} AgroSAM_{AC,AC'} = \sum_{AC'} AgroSAM_{AC,AC'} \\
& \sum_{AC'} \left[\sum_{AC} G_{ESA,AC} \cdot AgroSAM_{AC,AC'}\right] \cdot G_{ESA',AC'} = ESASAM_{ESA,ESA'} \\
& \sum_{AC} G_{CAX,AC} \cdot AgroSAM_{AC,AC'} = CAPCTR_{CAX,AC'} \cdot \exp\left[\sum_{s} W^{CAPCTR}_{CAX,AC',s} \cdot b^{CAPCTR}_{CAX,AC',s}\right]
\end{aligned}
\tag{99}
$$

Where:

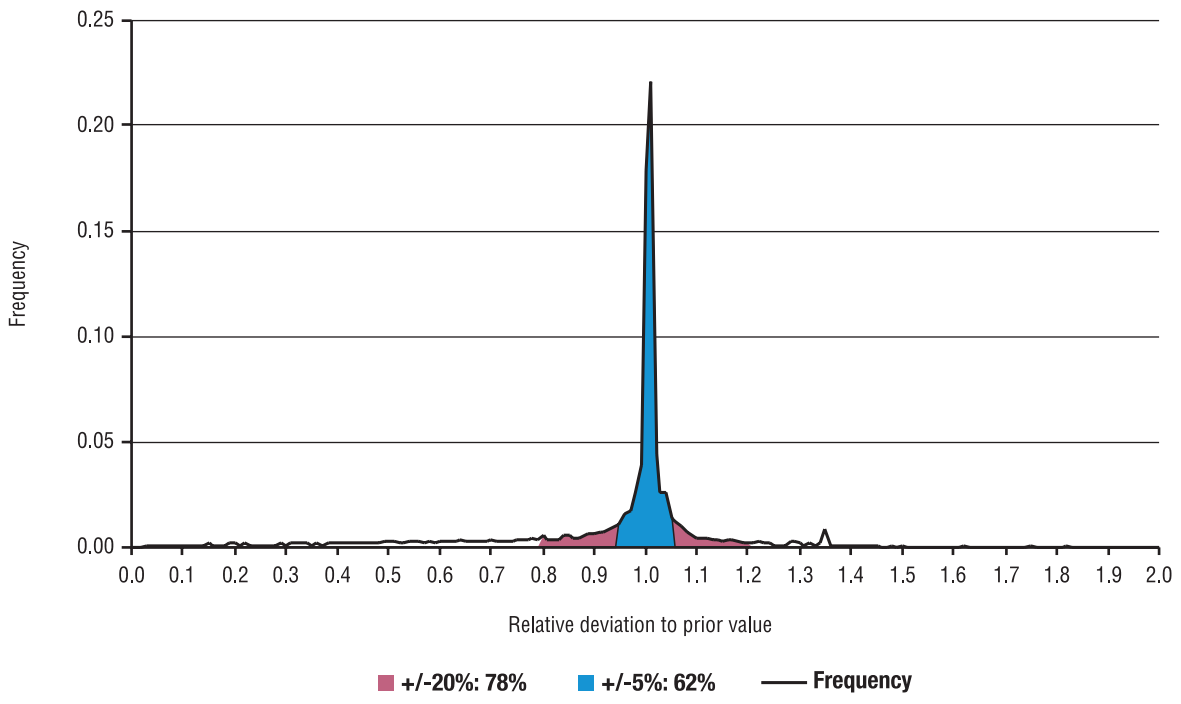
AgroSAM:	Balanced AgroSAM
$\overline{AgroSAM}$:	Prior AgroSAM derived from steps before
W:	Weights of correction-term support points
b:	Support points of correction-term
CAPCTR:	Control totals derived from CAPRI model database
WCAPCTR:	Weights of correction-term support points for stochastic control totals

bCAPCTR:	Support points of correction-term for stochastic control totals
CE:	Cross-entropy minimand
s:	Index for support points
AC:	AgroSAM accounts
ESA:	ESASAM accounts
CAX:	Rows and columns of the stochastic control totals derived from CAPRI
G:	Aggregator matrices between AgroSAM and ESASAM or CAPRI control accounts

A crucial issue was the determination of allowed deviations from the CAPRI control totals. Originally, it was intended to force the final values to replicate the available entries within the agriculture and food industries accounts, and ensuring consistency with ESASAM via the residual 'other crops' and 'other animals' accounts. However, it became clear that the decreased variability rendered the problem infeasible and it was decided to distinguish two types of accounts within the accounts to be disaggregated: Core accounts that have to be close to the control totals and buffer-accounts that could deviate to a higher extent (see Table 14, bold entries denote core accounts with a low permitted deviation). In addition, it was necessary to increase the allowed variation for individual accounts in certain Member States if the default setting rendered certain accounts to be infeasible. This could for instance be the case when the total demand form the tobacco industry for agricultural commodities required a substantial increase of the allowed variation of 'other crops' or the comparatively high value of private household consumption of agricultural commodities required additional variance for 'fruit and vegetable' accounts.

The results for the correction term for the disaggregated sectors (exp[Wb]) are displayed in Figure 18. It appears that 62% of the relative deviations between prior and estimated data are within a narrow range around 1 (0.95 and 1.05), implying that 62% of the final AgroSAM entries are considerably close to the prior entries derived from the CAPRI database and other sources. Furthermore, 78% of the final AgroSAM

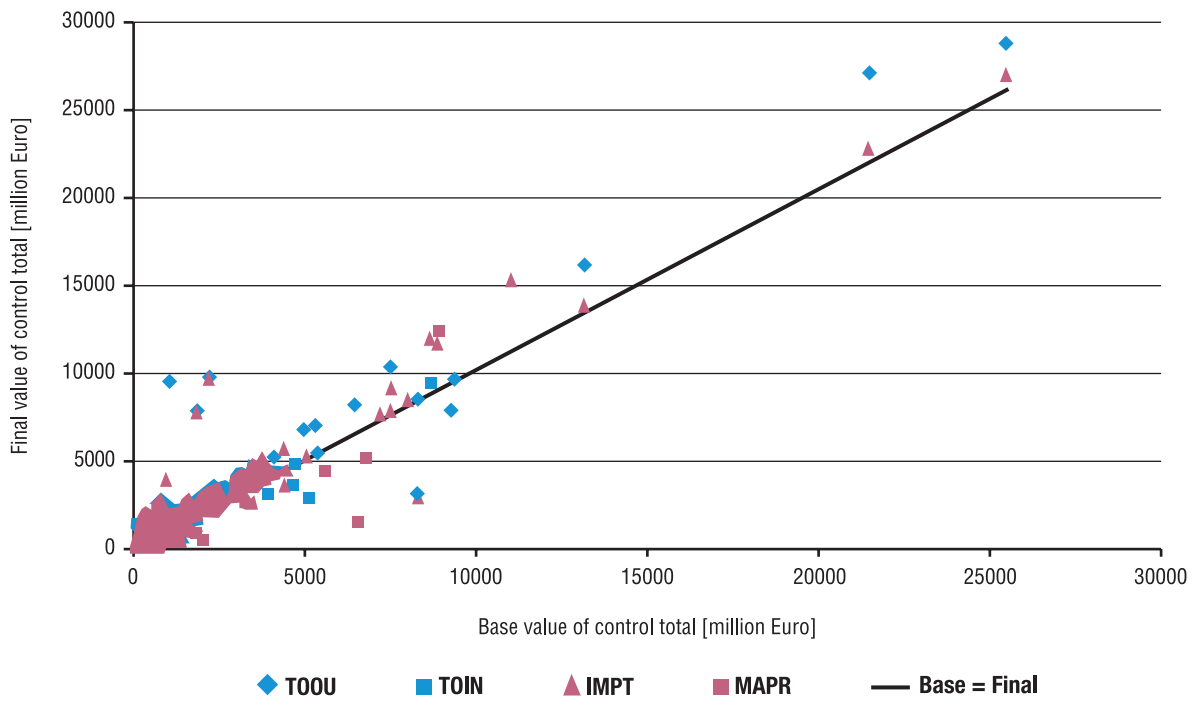
Figure 18. Frequency of Deviations between Prior an Final AgroSAM



entries deviate by a maximum of 20% from the prior entries. As implied by the chosen variation parameter (σ) for the disaggregated accounts, the largest deviations can be found within accounts like "other animals", "other crops", "other cereals", and "animal feed", which are likely to be less reliable than entries in the accounts for main marketed agricultural commodities like wheat, oilseeds, or milk from cows. The fact that the majority of the final entries are close to the prior data indicate that there is no systematic deviation between prior AgroSAM and the imposed control-totals in the form of the ESASAM. When comparing the final results with the control totals derived from the CAPRI model's database (Figure 19), it can be seen that the majority of estimated values lie within a comparatively narrow range around the imposed corresponding control values. Notable exceptions are highlighted in Figure 19 with the blue and purple circles. The blue circle refers to large positive deviations (estimated is much larger than observed), while the purple circle to large negative deviations (observed is much larger than estimated). The data points within the blue circle refer to 'animal feed' and 'poultry meat' in Italy, and to 'other crops' in France. Large negative deviations can be found for 'fruit and vegetable' production in Italy and Belgium. The main reason for these deviations is the fact that these activities and commodity groupings were not considered as core accounts (see Table 14 for a list of core accounts) of the AgroSAM estimation procedure and the variance of the stochastic control total was therefore larger than in the case of the core accounts (bold entries in Table 14).

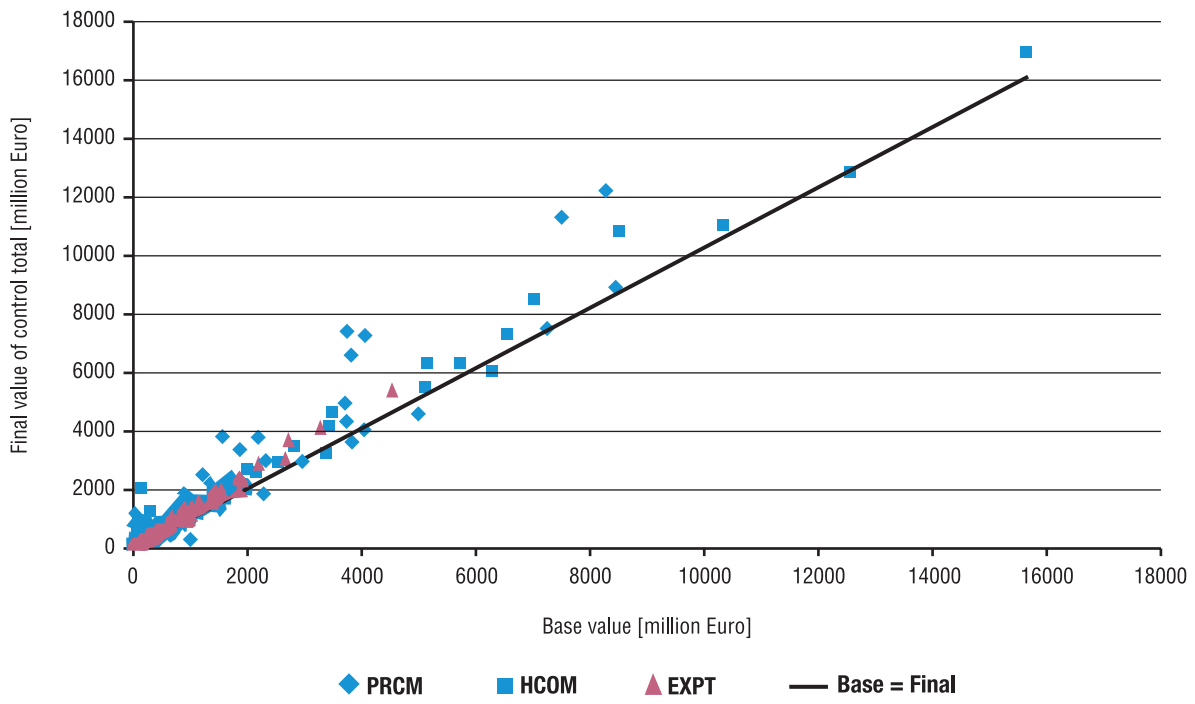
Particularly the large deviations for the 'fruit and vegetable' sector in Italy will have to be addressed in further stages of the AgroSAM project as this sector plays an important role within the national agricultural sector. A potential solution will be to impose the stochastic control totals not with equal variances for all Member States as done here, but according to the relative importance of the respective sub-sector.

Figure 19. Control Totals and Final Estimates for all Commodities and Activities



Notes: TOOU: Total output at basic prices
TOIN: Total input at basic prices
MAPR: Marketed production at basic prices
IMPT: Imports at basic prices

Figure 20. Uses of Core Commodities



Notes: PRCM: Processing demand
HCOM: Human consumption
EXPT: Exports

The estimation results for demand-side items are in general less satisfying than for the supply-side, but show no deviations in the order of magnitude for the core commodities. Approaches to estimate the demand-side with a higher accuracy will require more detailed information on the domestic transformation from basic- to purchaser's prices for the commodities in question. Particularly trade margins and export subsidies will be the information to be compiled for each commodity group. In general, the deviations are highest for 'processing demand'. As the CAPRI database does not distinguish between the demanding industries, the distribution of the total 'processing demand' has to be distributed based on shares and plausibility considerations. In some cases, for instance in the case of 'raw milk', it is evident that the largest share is processed by the dairy industry, but 'other food industries' may also demand a smaller amount. In most Member States, the total supply of raw milk is large enough to be distributed across numerous demanding industries while maintaining the relevance of dairy as the main consumer. Likewise, raw tobacco has to be mainly processed by the tobacco industry, but total tobacco supply is in some cases smaller than the intermediate demand of the tobacco industry for agricultural commodities as indicated by the EuroStat SUT. These figures sometimes even exceed the total supply of the aggregate 'other crop products', of which 'raw tobacco' is a part in the MAC scheme. As a consequence, the corresponding total supply values have to be adjusted such that compliance with the EuroStat SUT is achieved.

Table 14. Agriculture and Food Industry Sub-sectors, and Core Accounts

Description	Activity	Commodity
Agriculture		
Other wheat	**A_OWHE**	**C_OWHE**
Durum wheat	**A_DWHE**	**C_DWHE**
Barley	**A_BARL**	**C_BARL**
Grain maize	**A_MAIZ**	**C_MAIZ**
Other cereals	A_OCER	C_OCER
Paddy rice	A_PARI	C_PARI
Rape seed	**A_RAPE**	**C_RAPE**
Sunflower seed	**A_SUNF**	**C_SUNF**
Soya seed	**A_SOYA**	**C_SOYA**
Other oil plants	A_OOIL	C_OOIL
Other starch and protein plants	A_STPR	C_STPR
Potatoes	A_POTA	C_POTA
Sugar beet	**A_SUGB**	**C_SUGB**
Fibre plants	A_FIBR	C_FIBR
Other crop products	A_OTCR	C_OTCR
Grapes	A_GRPS	C_GRPS
Fresh vegetables, fruit, and nuts	A_FVEG	C_FVEG
Live plants	A_LPLT	C_LPLT
Fodder crops	A_FODD	C_FODD
Set-aside	A_SETA	
Raw milk from bovine cattle	**A_COMI**	**C_COMI**
Bovine cattle, slaughtered	A_LCAT	C_LCAT
Swine, slaughtered	**A_PIGF**	**C_PIGF**
Raw milk from sheep and goats	A_SGMI	C_SGMI
Sheep, goats, horses, asses, mules and hinnies, slaughtered	A_LSGE	C_LSGE
Eggs	A_EGGS	C_EGGS
Poultry, slaughtered	A_PLTR	C_PLTR
Other animals, live, and their products	A_OANM	C_OANM
Food Industry		
Rice, milled or husked	**A_RICE**	**C_RICE**
Processed sugar	**A_SUGA**	**C_SUGA**
Vegetable oils and fats, crude and refined; oil-cake and other solid residues, of vegetable fats or oils	**A_VOIL**	**C_VOIL**
Dairy products	**A_DAIR**	**C_DAIR**
Meat of bovine animals, fresh, chilled, or frozen	**A_BFVL**	**C_BFVL**
Meat of swine, fresh, chilled, or frozen	**A_PORK**	**C_PORK**
Meat of sheep, goats, and equines, fresh, chilled, or frozen	A_SGMT	C_SGMT
Meat and edible offal of poultry, fresh, chilled, or frozen	A_POUM	C_POUM
Prepared animal feeds	A_ANFD	C_ANFD

Note: Bold entries denote core commodities and activities

9. Summary and Conclusion

This technical report describes the steps of the compilation of AgroSAMs for 27 EU Member States. It starts with the creation of a complete set of macroeconomic indicators, which served as control totals for the compilation of SAMs at later stages. During this initial exercise, it occurred that the processed datasets were not always consistent or complete, and revealed often substantial deviations in key indicators. As a consequence, estimation procedures had to be applied already at an early stage to fill or adjust the entries in question (missing or inconsistent). Based on the thus derived macro-indicators, a set of highly aggregated SAMs were derived, which already featured the final level of detail with regard to institutions and monetary flows among them (InstSAM, see section 5.2). Due to the fact that Supply- and Use tables only included net taxes on commodities and activities, it became necessary to estimate the different types of taxes and subsidies paid or received on the commodity markets, particularly import duties and VAT (section 6.2). Inter-institutional flows from the InstSAMs were also included to obtain a set of not necessarily balanced (prior) versions of SAMs in ESA format (ESASAM). These ESASAMs entered a Cross-Entropy procedure that ensured their balance and consistency with the derived macro-totals (section 6.3).

Agricultural sector data from the CAPRI model was then processed into an Agricultural Accounting Matrix, both in values at basic prices and physical quantities (balance sheets). The chosen format permitted the direct mapping with the target AgroSAMs.

The described task to combine the database of an agricultural sector model with Supply- and Use tables from EuroStat resulted in a number of methodological and data-handling challenges. Although information from both databases were used in a most exhaustive manner, at some stages (e.g. in the case of accounts of the food industry) additional sources had to be consulted. Although the chosen procedure was tailored to available data and respected the main requirements, there is still huge potential for improvement. The main challenges for future work are:

1) Although considerable effort was devoted to the construction of an Agricultural Accounting Matrix based on data from the CAPRI model, it was not possible to eliminate the sometimes substantial deviations from the corresponding entries in the ESA95 matrices. This caused a need to distribute the deviations across the accounts to be disaggregated, such that sometimes large deviations from the original CAPRI data could not be avoided. The main reason for this is essentially the fact that the Agricultural Accounting Matrix compilation is merely a re-structuring of the EAA/CAPRI data in SAM format, in which only the introduction of additional non-agricultural processing activities respects the structure of the ESA95 framework. Consequently would a revision of the compilation procedure with respect to the structural deviations between EAA and ESA95 accounting schemes help to generate a prior that is closer to the corresponding ESA95 totals and improve the performance of the balancing procedure.

2) The original objective was to create a database which can be mapped (many-to-one) into the format required by GTAP. This task could not be fulfilled totally. Although it was possible to represent the agricultural and food-industry sectors in a way that is compatible with GTAP, the oil and gas sector, the ferrous and non-ferrous metal sector, and a few others could not be split with the available data sources (see Appendix 1, Table 17). Using the GTAP database itself was not considered as an

option because of the general paradigm of the AgroSAM project to focus on publicly available data from EuroStat. The decision, which dataset to use for the split of the respective sectors, is left to the respective user, in case he intents to run GTAP on the AgroSAM database.

3) Commodity market taxes like VAT and import duties had to be estimated, as the Supply- and Use tables only provided information on net-taxes, while the used macroeconomic datasets contained only the total amount of taxes paid without a distinction of the market transactions on which the taxes were levied. Although the applied approach takes information on applicable tax rates within the EU into account (namely for import duties and value-added type taxes), it would be desirable to use data directly obtained from the national statistical departments.

4) The representation of factor markets in the AgroSAM is comparatively coarse as the available datasets did not support a distinction between labour inputs by skill-type or the contribution of land and natural resources to the sectoral value-added.

One of the major paradigms of the project summarized here was that Supply- and Use tables from EuroStat determine the boundary condition for the final AgroSAM, such that the aggregated AgroSAM replicate exactly the Supply- and Use tables. This strict requirement means that deviations from the agricultural sector data from the CAPRI model had to be permitted. An alternative would have been to treat the CAPRI data as given and re-arrange the Supply- and Use tables accordingly. This alternative would have guaranteed a better representation of the agricultural and food industry sectors at the cost of the information from the economy-wide datasets. However, as the main purpose of the AgroSAM project was to create a database for general equilibrium models that allow analysing the linkages between agriculture and other sectors of the economy, it was decided to maintain the inter-sectoral structure as represented in the Supply- and Use tables.

It also has to be emphasized that the AgroSAM mainly serve as a database from which a model dataset may be derived: the existence of empty accounts and the wide range of entry-values (very small to very large) in the AgroSAM creates a need for aggregation (e.g. into GTAP format) and elimination of small entries before a CGE model is calibrated. Again, an alternative would have been to perform these steps as part of the AgroSAM project, which would have come at the cost of its versatility.

Although the structural deviations between the combined datasets created some difficulties for the project, it is still a major achievement that a full set of Social Accounting Matrices in ESA95 format could be compiled that is consistent with the respective sets of macro-economic control totals. These ESASAM can be created flexibly from EuroStat data for any desired year between 1995 and 2005 and may serve as control-totals for further disaggregation. The current state of the approach allows a fully flexible incorporation of additional data, which is intended to continue upon data availability. In any case the compiled AgroSAM constitute a valuable resource for modellers in the fields of general and partial equilibrium models.

10. References

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

- Biewen, E., S. Nolte, and M. Rosemann (2008): Multiplicative Measurement Error and the Simulation Extrapolation Method. IAW Discussion Paper 39

- Breisinger, C.; J. Thurlow; and D. Magnus (2007): A 2005 Social Accounting Matrix (SAM) for Ghana. Ghana Statistical Services (GSS); International Food Policy Research Institute

- Britz, W. and P. Witzke (2008): CAPRI Modelling System Documentation. URL: http://www.ilr1.uni-bonn.de/agpo/rsrch/capri/capri-documentation.pdf, University of Bonn.

- Byron, R. P. (1978): The estimation of large social account matrices. In: Journal of the Royal Statistical Society, Series A(141): 359–367.

- Canning, P. and Z. Wang (2005): A Flexible Mathematical Programming Model to Estimate Interregional Input-Output Accounts. In: Journal of Regional Sciences, 45(3): 539-563.

- Carroll, R. J., D. Ruppert, L. A. Stefanski (2006): Measurement Error in Nonlinear Models: A Modern Perspective. 2. Edition, CRC Press.

- DG-TAXUD (2005): VAT Rates Applied in the Member States of the European Community.

- EuroStat (1997): Manual on the Economic Accounts for Agriculture and Forestry EAA/EAF 97 (Rev. 1.1).

- EuroStat (2008): Taxation trends in the European Union. Data for the EU Member States and Norway.

- Golan, A., G. Judge, and S. Robinson (1994): Recovering Information from Incomplete or Partial Multisectoral Economic Data. In: Review of Economics and Statistics, 76(3): 541–49.

- Golan, A., G. Judge, and D. Miller (1996): Maximum Entropy Econometrics: Robust Estimation with Limited Data, New York.

- Greene, W. H. (2003): Econometric Analysis, fifth edition, New York.

- Hwang, J. T. (1986): Multiplicative Errors-in-Variables Models with Applications to Recent Data Released by the U.S. Department of Energy. In: Journal of the American Statistical Association, 81:680-688.

- Lofgren, H., R. Harris, and S. Robinson (2001): A Standard Computable General Equilibrium (CGE) Model in GAMS. International Food Policy Research Institute (IFPRI), Trade and Macroeconomics Divison (TMD), Discussion Paper No. 75, Washington, D.C.

- McDonald, S., K. Thierfelder, and S. Robinson (2007), Globe: A SAM Based Global CGE Model using GTAP Data, Departmental Working Paper No. 14, Department of Economics, United States Naval Academy, URL: http://www.usna.edu/EconDept/RePEc/usn/up/usnawp14.pdf.

- Müller, M. (2006): A General Equilibrium approach to Modelling Water and Land Use Reforms in Uzbekistan. Dissertation University of Bonn.

- Müller, M., P. Wehrheim (2004): Russia's cereal markets: current trends, changes in net-trade position, and policy implications. In: Agrarwirtschaft, 53:4.

- Nakamura, Y. (1998): Investment and Saving in Russian Macroeconomy: Compilation and Analyses of an Aggregated SAM for Russia, 1995. CERT Working Paper No. DP98/09.

- Pyatt, G. and J. Round (1985): Social accounting matrices: A basis for planning. Washington, D.C.: World Bank.

- Robinson, S., A. Cattaneo, and M. El-Said (2001): Updating and estimating a Social Accounting Matrix using cross entropy methods. In: Economic Systems Research, 13(1): 47-64.

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

- Stone R. (1968), Input-output projection: consistent prices and quantity structures. L'Industria, No.2 , 212-224.

- Theil, H. (1967): Economics and Information Theory. Amsterdam: North Holland.

- WTO (2007): Trade Policy Review: European Communities. Report by the Secretariat WT/TPR/S/177.

Appendix 1: Classifications and Correspondence Sheets

NOTE: The codes for the MAC are here presented in the format used for the compilation of the SAMs, which means that they are supplemented with the following pre-fixes:

A_:	*Activities*
C_:	*Commodities*
A/C_:	*Activities and commodities*
F_:	*Primary factors*
T_:	*Taxes and transactions*
I_:	*Institutions*
X_:	*Auxiliary accounts*

The rest of the codes are similar to the original classifications. In case identifiers are used in different classifications (like a01 (Agriculture) appearing in NACE and CPA, or DWHE (Durum Wheat) appearing in CAPRI ROWS and COLS) an additional identifier is added (e.g. NACE: or COLS:)

Table 15. Modified Agricultural Classification and Correspondence to European System of National Accounts

Modified agricultural classification (MAC)		European System of National Accounts (ESA95)	
Description	**Code**	**Code**	**Description**
Production of other wheat	A_OWHE	NACE: a01	Agriculture, hunting and related service activities
Production of durum wheat	A_DWHE		
Production of barley	A_BARL		
Production of grain maize	A_MAIZ		
Production of other cereals	A_OCER		
Production of paddy rice	A_PARI		
Production of rape seed	A_RAPE		
Production of sunflower seed	A_SUNF		
Production of soya seed	A_SOYA		
Production of other oil plants	A_OOIL		
Production of other starch and protein plants	A_STPR		
Production of potatoes	A_POTA		
Production of sugar beet	A_SUGB		
Production of fibre plants	A_FIBR		
Production of grapes	A_GRPS		
Production of fresh vegetables, fruit, and nuts	A_FVEG		
Production of live plants	A_LPLT		
Other crop production activities	A_OTCR		
Production of fodder crops	A_FODD		
Set aside	A_SETA		
Production of raw milk from bovine cattle	A_COMI		
Production of bovine cattle, live	A_LCAT		
Production of swine, live	A_PIGF		
Production of raw milk from sheep and goats	A_SGMI		
Production of sheep, goats, horses, asses, mules and hinnies, live	A_LSGE		
Production of eggs	A_EGGS		
Production of poultry, live	A_PLTR		
Production of wool and animal hair; silk-worm CAPRIons suitable for reeling	A_ANHR		
Production of other animals, live, and their products	A_OANM		
Agricultural service activities	A_AGSV		
Forestry, logging and related service activities	A_FORE	NACE: a02	Forestry, logging and related service activities
Fishing, operation of fish hatcheries and fish farms; service activities incidental to fishing	A_FISH	NACE: b05	Fishing, operating of fish hatcheries and fish farms; service activities incidental to fishing
Mining of coal and lignite; extraction of peat	A_COAL	NACE: ca10	Mining of coal and lignite; extraction of peat
Extraction of crude petroleum; service activities incidental to oil extraction excluding surveying	A_COIL	NACE: ca11	Extraction of crude petroleum and natural gas; service activities incidental to oil and gas extraction excluding surveying
Mining of uranium and thorium ores	A_URAN	NACE: ca12	Mining of uranium and thorium ores
Mining of metal ores	A_MEOR	NACE: cb13	Mining of metal ores
Other mining and quarrying	A_OMIN	NACE: cb14	Other mining and quarrying
Processing of rice, milled or husked	A_RICE	NACE: da15	Manufacture of food products and beverages
Production of other food	A_OFOD		
Processing of sugar	A_SUGA		
Production of vegetable oils and fats, crude and refined; oil-cake and other solid residues, of vegetable fats or oils	A_VOIL		
Dairy	A_DAIR		
Production of meat of bovine animals, fresh, chilled, or frozen	A_BFVL		
Production of meat of swine, fresh, chilled, or frozen	A_PORK		
Production of meat of sheep, goats, and equines, fresh, chilled, or frozen	A_SGMT		

Modified agricultural classification (MAC)		European System of National Accounts (ESA95)	
Description	**Code**	**Code**	**Description**
Meat and edible offal of poultry, fresh, chilled, or frozen	A_POUM	NACE: da15	Manufacture of food products and beverages
Production of beverages	A_BEVR		
Production of prepared animal feeds	A_ANFD		
Manufacture of tobacco products	A_TOBA	NACE: da16	Manufacture of tobacco products
Manufacture of textiles	A_TEXT	NACE: db17	Manufacture of textiles
Manufacture of wearing apparel; dressing and dyeing of fur	A_GARM	NACE: db18	Manufacture of wearing apparel; dressing and dyeing of fur
Tanning and dressing of leather; manufacture of luggage, handbags, saddlery, harness and footwear	A_LETH	NACE: dc19	Tanning and dressing of leather; manufacture of luggage, handbags, saddlery, harness and footwear
Manufacture of wood and of products of wood and cork, except furniture; manufacture of articles of straw and plaiting materials	A_WOOD	NACE: dd20	Manufacture of wood and of products of wood and cork, except furniture; manufacture of articles of straw and plaiting materials
Manufacture of pulp, paper and paper products	A_PULP	NACE: de21	Manufacture of pulp, paper and paper products
Publishing, printing and reproduction of recorded media	A_MDIA	NACE: de22	Publishing, printing and reproduction of recorded media
Manufacture of coke, refined petroleum products and nuclear fuels	A_RPET	NACE: df23	Manufacture of coke, refined petroleum products and nuclear fuels
Manufacture of chemicals and chemical products	A_CHEM	NACE: dg24	Manufacture of chemicals and chemical products
Manufacture of rubber and plastic products	A_PLST	NACE: dh25	Manufacture of rubber and plastic products
Manufacture of other non-metallic mineral products	A_NMMP	NACE: di26	Manufacture of other non-metallic mineral products
Manufacture of basic metals	A_FRMT	NACE: dj27	Manufacture of basic metals
Manufacture of fabricated metal products, except machinery and equipment	A_FAME	NACE: dj28	Manufacture of fabricated metal products, except machinery and equipment
Manufacture of machinery and equipment n.e.c.	A_MACH	NACE: dk29	Manufacture of machinery and equipment n.e.c.
Manufacture of office machinery and computers	A_OFMA	NACE: dl30	Manufacture of office machinery and computers
Manufacture of electrical machinery and apparatus n.e.c.	A_ELMA	NACE: dl31	Manufacture of electrical machinery and apparatus n.e.c.
Manufacture of radio, television and communication equipment and apparatus	A_COEQ	NACE: dl32	Manufacture of radio, television and communication equipment and apparatus
Manufacture of medical, precision and optical instruments, watches and clocks	A_MEIN	NACE: dl33	Manufacture of medical, precision and optical instruments, watches and clocks
Manufacture of motor vehicles, trailers and semi-trailers	A_MOTO	NACE: dm34	Manufacture of motor vehicles, trailers and semi-trailers
Manufacture of other transport equipment	A_TREQ	NACE: dm35	Manufacture of other transport equipment
Manufacture of furniture; manufacturing n.e.c.	A_FURN	NACE: dn36	Manufacture of furniture; manufacturing n.e.c.
Recycling	A_RECY	NACE: dn37	Recycling
Electricity, gas, steam and hot water supply	A_EGSW	NACE: e40	Electricity, gas, steam and hot water supply
Collection, purification and distribution of water	A_WATR	NACE: e41	Collection, purification and distribution of water
Construction	A_CONS	NACE: f45	Construction
Sale, maintenance and repair of motor vehicles and motorcycles; retail sale services of automotive fuel	A_SMOT	NACE: g50	Sale, maintenance and repair of motor vehicles and motorcycles; retail sale services of automotive fuel
Wholesale trade and commission trade, except of motor vehicles and motorcycles	A_WTRD	NACE: g51	Wholesale trade and commission trade, except of motor vehicles and motorcycles
Retail trade, except of motor vehicles and motorcycles; repair of personal and household goods	A_RTRD	NACE: g52	Retail trade, except of motor vehicles and motorcycles; repair of personal and household goods
Hotels and restaurants	A_HORE	NACE: h55	Hotels and restaurants
Land transport; transport via pipelines	A_TRLD	NACE: i60	Land transport; transport via pipelines
Water transport	A_TRWA	NACE: i61	Water transport
Air transport	A_TRAR	NACE: i62	Air transport
Supporting and auxiliary transport activities; activities of travel agencies	A_TROT	NACE: i63	Supporting and auxiliary transport activities; activities of travel agencies
Post and telecommunications	A_TECO	NACE: i64	Post and telecommunications
Financial intermediation, except insurance and pension funding	A_FINA	NACE: j65	Financial intermediation, except insurance and pension funding

Modified agricultural classification (MAC)		European System of National Accounts (ESA95)	
Description	**Code**	**Code**	**Description**
Insurance and pension funding, except compulsory social security	A_INSU	NACE: j66	Insurance and pension funding, except compulsory social security
Activities auxiliary to financial intermediation	A_FIAX	NACE: j67	Activities auxiliary to financial intermediation
Real estate activities	A_ESTA	NACE: k70	Real estate activities
Renting of machinery and equipment without operator and of personal and household goods	A_MARE	NACE: k71	Renting of machinery and equipment without operator and of personal and household goods
Computer and related activities	A_COMP	NACE: k72	Computer and related activities
Research and development	A_RESC	NACE: k73	Research and development
Other business activities	A_OTBS	NACE: k74	Other business activities
Public administration and defence; compulsory social security	A_PUAD	NACE: l75	Public administration and defence; compulsory social security
Education	A_EDUC	NACE: m80	Education
Health and social work	A_HESO	NACE: n85	Health and social work
Sewage and refuse disposal, sanitation and similar activities	A_SANI	NACE: o90	Sewage and refuse disposal, sanitation and similar activities
Activities of membership organization n.e.c.	A_ORGA	NACE: o91	Activities of membership organisation n.e.c.
Recreational, cultural and sporting activities	A_CULT	NACE: o92	Recreational, cultural and sporting activities
Other service activities	A_OTSV	NACE: o93	Other service activities
Private households with employed persons	A_PRHH	NACE: p95	Private households with employed persons
Other wheat	C_OWHE	CPA: a01	Products of agriculture, hunting and related services
Durum wheat	C_DWHE		
Barley	C_BARL		
Grain maize	C_MAIZ		
Other cereals	C_OCER		
Paddy rice	C_PARI		
Rape seed	C_RAPE		
Sunflower seed	C_SUNF		
Soya seed	C_SOYA		
Other oil plants	C_OOIL		
Other starch and protein plants	C_STPR		
Potatoes	C_POTA		
Sugar beet	C_SUGB		
Fibre plants	C_FIBR		
Grapes	C_GRPS		
Fresh vegetables, fruit, and nuts	C_FVEG		
Live plants	C_LPLT		
Other crop products	C_OTCR		
Fodder crops	C_FODD		
Raw milk from bovine cattle	C_COMI		
Bovine cattle, slaughtered	C_LCAT		
Swine, slaughtered	C_PIGF		
Raw milk from sheep and goats	C_SGMI		
Sheep, goats, horses, asses, mules and hinnies, slaughtered	C_LSGE		
Eggs	C_EGGS		
Poultry, slaughtered	C_PLTR		
Wool and animal hair; Silk-worm CAPRIons suitable for reeling	C_ANHR		
Other animals, live, and their products	C_OANM		
Agricultural services	C_AGSV		
Products of forestry, logging and related services	C_FORE	CPA: a02	Products of forestry, logging and related services
Fish and other fishing products; services incidental of fishing	C_FISH	CPA: b05	Fish and other fishing products; services incidental of fishing
Coal and lignite; peat	C_COAL	CPA: ca10	Coal and lignite; peat

Modified agricultural classification (MAC)		European System of National Accounts (ESA95)	
Description	**Code**	**Code**	**Description**
Crude petroleum and natural gas; services incidental to oil and gas extraction excluding surveying	C_COIL	CPA: ca11	Crude petroleum and natural gas; services incidental to oil and gas extraction excluding surveying
Uranium and thorium ores	C_URAN	CPA: ca12	Uranium and thorium ores
Metal ores	C_MEOR	CPA: cb13	Metal ores
Other mining and quarrying products	C_OMIN	CPA: cb14	Other mining and quarrying products
Rice, milled or husked	C_RICE	CPA: da15	Food products and beverages
Other food products	C_OFOD		
Processed sugar	C_SUGA		
Vegetable oils and fats, crude and refined; oil-cake and other solid residues, of vegetable fats or oils	C_VOIL		
Dairy products	C_DAIR		
Meat of bovine animals, fresh, chilled, or frozen	C_BFVL		
Meat of swine, fresh, chilled, or frozen	C_PORK		
Meat of sheep, goats, and equines, fresh, chilled, or frozen	C_SGMT		
Meat and edible offal of poultry, fresh, chilled, or frozen	C_POUM		
Beverages	C_BEVR		
Prepared animal feeds	C_ANFD		
Tobacco products	C_TOBA	CPA: da16	Tobacco products
Textiles	C_TEXT	CPA: db17	Textiles
Wearing apparel; furs	C_GARM	CPA: db18	Wearing apparel; furs
Leather and leather products	C_LETH	CPA: dc19	Leather and leather products
Wood and products of wood and cork (except furniture); articles of straw and plaiting materials	C_WOOD	CPA: dd20	Wood and products of wood and cork (except furniture); articles of straw and plaiting materials
Pulp, paper and paper products	C_PULP	CPA: de21	Pulp, paper and paper products
Printed matter and recorded media	C_MDIA	CPA: de22	Printed matter and recorded media
Coke, refined petroleum products and nuclear fuels	C_RPET	CPA: df23	Coke, refined petroleum products and nuclear fuels
Chemicals, chemical products and man-made fibres	C_CHEM	CPA: dg24	Chemicals, chemical products and man-made fibres
Rubber and plastic products	C_PLST	CPA: dh25	Rubber and plastic products
Other non-metallic mineral products	C_NMMP	CPA: di26	Other non-metallic mineral products
Basic metals	C_FRMT	CPA: dj27	Basic metals
Fabricated metal products, except machinery and equipment	C_FAME	CPA: dj28	Fabricated metal products, except machinery and equipment
Machinery and equipment n.e.c.	C_MACH	CPA: dk29	Machinery and equipment n.e.c.
Office machinery and computers	C_OFMA	CPA: dl30	Office machinery and computers
Electrical machinery and apparatus n.e.c.	C_ELMA	CPA: dl31	Electrical machinery and apparatus n.e.c.
Radio, television and communication equipment and apparatus	C_COEQ	CPA: dl32	Radio, television and communication equipment and apparatus
Medical, precision and optical instruments, watches and clocks	C_MEIN	CPA: dl33	Medical, precision and optical instruments, watches and clocks
Motor vehicles, trailers and semi-trailers	C_MOTO	CPA: dm34	Motor vehicles, trailers and semi-trailers
Other transport equipment	C_TREQ	CPA: dm35	Other transport equipment
Furniture; other manufactured goods n.e.c.	C_FURN	CPA: dn36	Furniture; other manufactured goods n.e.c.
Secondary raw materials	C_RECY	CPA: dn37	Secondary raw materials
Electrical energy, gas, steam and hot water	C_EGSW	CPA: e40	Electrical energy, gas, steam and hot water
Collected and purified water, distribution services of water	C_WATR	CPA: e41	Collected and purified water, distribution services of water
Construction work	C_CONS	CPA: f45	Construction work
Trade, maintenance and repair services of motor vehicles and motorcycles; retail sale of automotive fuel	C_SMOT	CPA: g50	Trade, maintenance and repair services of motor vehicles and motorcycles; retail sale of automotive fuel
Wholesale trade and commission trade services, except of motor vehicles and motorcycles	C_WTRD	CPA: g51	Wholesale trade and commission trade services, except of motor vehicles and motorcycles

Modified agricultural classification (MAC)		European System of National Accounts (ESA95)	
Description	**Code**	**Code**	**Description**
Retail trade services, except of motor vehicles and motorcycles; repair services of personal and household goods	C_RTRD	CPA: g52	Retail trade services, except of motor vehicles and motorcycles; repair services of personal and household goods
Hotel and restaurant services	C_HORE	CPA: h55	Hotel and restaurant services
Land transport; transport via pipeline services	C_TRLD	CPA: i60	Land transport; transport via pipeline services
Water transport services	C_TRWA	CPA: i61	Water transport services
Air transport services	C_TRAR	CPA: i62	Air transport services
Supporting and auxiliary transport services; travel agency services	C_TROT	CPA: i63	Supporting and auxiliary transport services; travel agency services
Post and telecommunication services	C_TECO	CPA: i64	Post and telecommunication services
Financial intermediation services, except insurance and pension funding services	C_FINA	CPA: j65	Financial intermediation services, except insurance and pension funding services
Insurance and pension funding services, except compulsory social security services	C_INSU	CPA: j66	Insurance and pension funding services, except compulsory social security services
Services auxiliary to financial intermediation	C_FIAX	CPA: j67	Services auxiliary to financial intermediation
Real estate services	C_ESTA	CPA: k70	Real estate services
Renting services of machinery and equipment without operator and of personal and household goods	C_MARE	CPA: k71	Renting services of machinery and equipment without operator and of personal and household goods
Computer and related services	C_COMP	CPA: k72	Computer and related services
Research and development services	C_RESC	CPA: k73	Research and development services
Other business services	C_OTBS	CPA: k74	Other business services
Public administration and defence services; compulsory social security services	C_PUAD	CPA: l75	Public administration and defence services; compulsory social security services
Education services	C_EDUC	CPA: m80	Education services
Health and social work services	C_HESO	CPA: n85	Health and social work services
Sewage and refuse disposal services, sanitation and similar services	C_SANI	CPA: o90	Sewage and refuse disposal services, sanitation and similar services
Membership organisation services n.e.c.	C_ORGA	CPA: o91	Membership organisation services n.e.c.
Recreational, cultural and sporting services	C_CULT	CPA: o92	Recreational, cultural and sporting services
Other services	C_OTSV	CPA: o93	Other services
Private households with employed persons	C_PRHH	CPA: p95	Private households with employed persons
Compensation of employees	F_LABT	d1	Compensation of employees
Operating surplus, gross	F_CAPT	b2g_b3g	Operating surplus, gross
Trade and transport margins	T_TTRM	p118	Trade and transport margins
Other taxes on production	T_TXAC	d29	Other taxes on production
Other subsidies on production	T_SBAC	d39	Other subsidies on production
Value added type taxes (VAT)	T_TXVA	d211	Value added type taxes (VAT)
Taxes and duties on imports excluding VAT	T_TXIM	d212	Taxes and duties on imports excluding VAT
Taxes on products, except VAT and import taxes	T_TXSL	d214	Taxes on products, except VAT and import taxes
Subsidies on products	T_SBCM	d31	Subsidies on products
Property income	T_PINC	d4	Property income
Current taxes on income, wealth, etc.	T_DTXT	d5	Current taxes on income, wealth, etc
Social contributions and benefits	T_SCCN	d6	Social contributions and benefits
Other current transfers	T_OTCT	d7	Other current transfers
Adjustment for the change in net equity of households in pension funds reserves	T_ADCH	d8	Adjustment for the change in net equity of households in pension funds reserves
Corporations	I_ENTR	s11_s12	Corporations
General government	I_NGOV	s13	General government
Households	I_HHLD	s14_s15	Households, non-profit institutions serving households
Rest of the world	I_ROWD	s2	Rest of the world
Gross fixed capital formation	I_GFCF	p51	Gross fixed capital formation
Changes in inventories	I_STCH	p52_p53	Changes in inventories and valuables
Gross capital formation	I_GCFM	p5	Gross capital formation
Row, Columns, and Sub-totals	Total	Total	

Table 16. CAPRI Classifications and Correspondence to Modified Agricultural Classification

CAPRI Classifications		Modified agricultural classification (MAC)	
Description	**Code**	**Code**	**Description**
Soft wheat production activity	COLS: SWHE	A_OWHE	Production of other wheat
Durum wheat production activity	COLS: DWHE	A_DWHE	Production of durum wheat
Rye and meslin production activity	COLS: RYEM	A_OCER	Production of other cereals
Barley production activity	COLS: BARL	A_BARL	Production of barley
Oats and summer cereal mixes production activity without triticale	COLS: OATS	A_OCER	Production of other cereals
Grain maize production activity	COLS: MAIZ	A_MAIZ	Production of grain maize
Other cereals production activity including triticale	COLS: OCER	A_OCER	Production of other cereals
Paddy rice production activity	COLS: PARI	A_PARI	Production of paddy rice
Rape production activity	COLS: RAPE	A_RAPE	Production of rape seed
Sunflower production activity	COLS: SUNF	A_SUNF	Production of sunflower seed
Soya production activity	COLS: SOYA	A_SOYA	Production of soya seed
Olive production activity for the oil industry	COLS: OLIV	A_OOIL	Production of other oil plants
Other seed production activities for the oil industry	COLS: OOIL		
Pulses production activity	COLS: PULS	A_STPR	Production of other starch and protein plants
Potatoes production activity	COLS: POTA	A_POTA	Production of potatoes
Sugar beet production activity	COLS: SUGB	A_SUGB	Production of sugar beet
Flax and hemp production activity	COLS: TEXT	A_FIBR	Production of fibre plants
Tobacco production activity	COLS: TOBA	A_OTCR	Other crop production activities
Other industrial crops production activity	COLS: OIND		
Tomatoes production activity	COLS: TOMA	A_FVEG	Production of fresh vegetables, fruit, and nuts
Other vegetables production activity	COLS: OVEG		
Apples pears and peaches production activity	COLS: APPL		
Other fruits production activity	COLS: OFRU		
Citrus fruits production activity	COLS: CITR		
Table grapes production activity	COLS: TAGR	A_GRPS	Production of grapes
Table olives production activity	COLS: TABO	A_FVEG	Production of fresh vegetables, fruit, and nuts
Wine production activity	COLS: TWIN	A_BEVR	Production of beverages
Nurseries production activity	COLS: NURS	A_LPLT	Production of live plants
Flowers production activity	COLS: FLOW		
Other crops production activity	COLS: OCRO	A_OTCR	Other crop production activities
Fodder maize production activity	COLS: MAIF	A_FODD	Production of fodder crops
Fodder root crops production activity	COLS: ROOF		
Fodder other on arable land production activity	COLS: OFAR		
Gras and grazings production activity	COLS: GRAS		
Set aside	COLS: SETA	A_SETA	Set aside
Non food production activities on set aside	COLS: NONF		
Fallow land and pastures	COLS: FALL	A_FODD	Production of fodder crops
Dairy cows production activity	COLS: DCOW	A_COMI	Production of raw milk from bovine cattle
Suckler cows production activity	COLS: SCOW	A_LCAT	Production of bovine cattle, live
Male adult fattening activity	COLS: BULF		
Heifers fattening activity	COLS: HEIF		
Heifers raising activity	COLS: HEIR		
Calves male fattening activity	COLS: CAMF		
Calves female fattening activity	COLS: CAFF		
Calves male raising activity	COLS: CAMR		
Calves female raising activity	COLS: CAFR		
Pig fattening activity	COLS: PIGF	A_PIGF	Production of swine, live

CAPRI Classifications		Modified agricultural classification (MAC)	
Description	**Code**	**Code**	**Description**
Sows for piglet production	COLS: SOWS	A_PIGF	Production of swine, live
Sheep and goats activity for milk production	COLS: SHGM	A_SGMI	Production of raw milk from sheep and goats
Sheep and goats activity for fattening	COLS: SHGF	A_LSGE	Production of sheep, goats, horses, asses, mules and hinnies, live
Laying hens production activity	COLS: HENS	A_EGGS	Production of eggs
Poultry fattening activity	COLS: POUF	A_PLTR	Production of poultry, live
Other animals production activity	COLS: OANI	A_OANM	Production of other animals, live, and their products
Rice milled	COLS: RICE	A_RICE	Processing of rice, milled or husked
Starch	COLS: STAR	A_OFOD	Production of other food
Processed sugar	COLS: SUGA	A_SUGA	Processing of sugar
Rape seed oil	COLS: RAPO	A_VOIL	Production of vegetable oils and fats, crude and refined; oil-cake and other solid residues, of vegetable fats or oils
Sunflower seed oil	COLS: SUNO		
Soya oil	COLS: SOYO		
Olive oil	COLS: OLIO		
Other oil	COLS: OTHO		
Raw milk at dairy	COLS: MILK	A_DAIR	Dairy
Animal feed	COLS: ANFD	A_ANFD	Production of prepared animal feeds
Beef meat	COLS: BEEF	A_BFVL	Production of meat of bovine animals, fresh, chilled, or frozen
Pork meat	COLS: PORK	A_PORK	Production of meat of swine, fresh, chilled, or frozen
Sheep and goat meat	COLS: SGMT	A_SGMT	Production of meat of sheep, goats, and equines, fresh, chilled, or frozen
Poultry meat	COLS: POUM	A_POUM	Meat and edible offal of poultry, fresh, chilled, or frozen
Agricultural Services Output	COLS: SERO	A_AGSV	Agricultural service activities
Soft wheat	ROWS: SWHE	C_OWHE	Other wheat
Durum wheat	ROWS: DWHE	C_DWHE	Durum wheat
Rye and meslin	ROWS: RYEM	C_OCER	Other cereals
Barley	ROWS: BARL	C_BARL	Barley
Oats	ROWS: OATS	C_OCER	Other cereals
Grain maize	ROWS: MAIZ	C_MAIZ	Grain maize
Other cereals	ROWS: OCER	C_OCER	Other cereals
Paddy rice	ROWS: PARI	C_PARI	Paddy rice
Rape seed	ROWS: RAPE	C_RAPE	Rape seed
Sunflower seed	ROWS: SUNF	C_SUNF	Sunflower seed
Soya seed	ROWS: SOYA	C_SOYA	Soya seed
Olives for the oil industry	ROWS: OLIV	C_OOIL	Other oil plants
Other seeds for the oil industry	ROWS: OOIL		
Pulses	ROWS: PULS	C_STPR	Other starch and protein plants
Potatoes	ROWS: POTA	C_POTA	Potatoes
Sugar beet	ROWS: SUGB	C_SUGB	Sugar beet
Flax and hemp	ROWS: TEXT	C_FIBR	Fibre plants
Tobacco	ROWS: TOBA	C_OTCR	Other crop products
Other industrial crops	ROWS: OIND		
Tomatoes	ROWS: TOMA	C_FVEG	Fresh vegetables, fruit, and nuts
Other vegetables	ROWS: OVEG		
Apples pears and peaches	ROWS: APPL		
Other fruits	ROWS: OFRU		
Citrus fruits	ROWS: CITR		
Table grapes	ROWS: TAGR	C_GRPS	Grapes

CAPRI Classifications		Modified agricultural classification (MAC)	
Description	**Code**	**Code**	**Description**
Table olives	ROWS: TABO	C_FVEG	Fresh vegetables, fruit, and nuts
Table wine	ROWS: TWIN	C_BEVR	Beverages
Nurseries	ROWS: NURS	C_LPLT	Live plants
Flowers	ROWS: FLOW		
Other crops	ROWS: OCRO	C_OTCR	Other crop products
Fodder maize	ROWS: MAIF	C_FODD	Fodder crops
Fodder root crops	ROWS: ROOF		
Fodder other on arable land	ROWS: OFAR		
Grass	ROWS: GRAS		
Straw	ROWS: STRA	C_OTCR	Other crop products
Other animals output	ROWS: OANI	C_OANM	Other animals, live, and their products
Cow and buffalo milk	ROWS: COMI	C_COMI	Raw milk from bovine cattle
Sheep and goat milk	ROWS: SGMI	C_SGMI	Raw milk from sheep and goats
Eggs	ROWS: EGGS	C_EGGS	Eggs
Milk for feeding	ROWS: COMF	C_COMI	Raw milk from bovine cattle
Sheep and goat milk for feeding	ROWS: SGMF	C_SGMI	Raw milk from sheep and goats
Slaughtered cow output	ROWS: SCOW	C_LCAT	Bovine cattle, slaughtered
Slaughtered bull output	ROWS: SBUL		
Slaughtered heifer output	ROWS: SHEI		
Slaughtered male calf output	ROWS: SCAM		
Slaughtered female calf output	ROWS: SCAF		
Slaughtered piglet output	ROWS: SPIG	C_PIGF	Swine, slaughtered
Slaughtered lamb output	ROWS: SLAM	C_LSGE	Sheep, goats, horses, asses, mules and hinnies, slaughtered
Slaughtered chicken output	ROWS: SCHI	C_PLTR	Poultry, slaughtered
Beef meat	ROWS: BEEF	C_BFVL	Meat of bovine animals, fresh, chilled, or frozen
Pork meat	ROWS: PORK	C_PORK	Meat of swine, fresh, chilled, or frozen
Sheep and goat meat	ROWS: SGMT	C_SGMT	Meat of sheep, goats, and equines, fresh, chilled, or frozen
Poultry meat	ROWS: POUM	C_POUM	Meat and edible offal of poultry, fresh, chilled, or frozen
Fodder rich protein	ROWS: FPRO	C_ANFD	Prepared animal feeds
Fodder rich energy	ROWS: FENE		
Fodder other	ROWS: FOTH		
Rice milled	ROWS: RICE	C_RICE	Rice, milled or husked
Molasse	ROWS: MOLA	C_SUGA	Processed sugar
Starch	ROWS: STAR	C_OFOD	Other food products
Processed sugar	ROWS: SUGA	C_SUGA	Processed sugar
Rape seed oil	ROWS: RAPO	C_VOIL	Vegetable oils and fats, crude and refined; oil-cake and other solid residues, of vegetable fats or oils
Sunflower seed oil	ROWS: SUNO		
Soya oil	ROWS: SOYO		
Olive oil	ROWS: OLIO		
Other oil	ROWS: OTHO		
Rape seed cake	ROWS: RAPC		
Sunflowe seed cake	ROWS: SUNC		
Soya cake	ROWS: SOYC		
Olive cake	ROWS: OLIC		
Other cake	ROWS: OTHC		
Raw milk at dairy	ROWS: MILK	C_DAIR	Dairy products
Butter	ROWS: BUTT		
Skimmed milk powder	ROWS: SMIP		

CAPRI Classifications		Modified agricultural classification (MAC)	
Description	**Code**	**Code**	**Description**
Cheese	ROWS: CHES	C_DAIR	Dairy products
Fresh milk products	ROWS: FRMI		
Cream	ROWS: CREM		
Concentrated milk	ROWS: COCM		
Whole milk powder	ROWS: WMIP		
Nitrogen in fertiliser	ROWS: NITF	C_CHEM	Chemicals, chemical products and man-made fibres
Phospate in fertiliser [P2O5]	ROWS: PHOF		
Potassium in fertiliser [K2O]	ROWS: POTF		
Renting of milk quota	ROWS: RQUO	C_AGSV	Agricultural services
Agricultural Services Output	ROWS: SERO		
Non Agricultural Secondary Activities	ROWS: NASA		
Calcium in fertiliser	ROWS: CAOF	C_CHEM	Chemicals, chemical products and man-made fibres
Seed input	ROWS: SEED		
Plant protection inputs	ROWS: PLAP		
Pharmaceutical inputs	ROWS: IPHA		
Maintenance materials	ROWS: REPM	C_MACH	Machinery and equipment n.e.c.
Maintenance buildings	ROWS: REPB	C_CONS	Construction
Electricity	ROWS: ELEC	C_EGSW	Electrical energy, gas, steam and hot water
Heating gas and oil	ROWS: EGAS		
Fuels	ROWS: EFUL	C_RPET	Coke, refined petroleum products and nuclear fuels
Lubricants	ROWS: ELUB		
Agricultural Services input	ROWS: SERI	C_AGSV	Agricultural services
Human consumption market	COLS: HCOM	I_HHLD	Households; non-profit institutions serving households
Stock changes on market	COLS: STCM	I_STCH	Changes in inventories
Imports total	COLS: IMPT	I_ROWD	Rest of the world
Exports total	COLS: EXPT		
Other taxes on production	ROWS: TAXO	T_TXAC	Other taxes on production
Other subsisides on production	ROWS: SUBO	T_SBAC	Other subsidies on production
CAP premium effective	ROWS: PRME	T_SBAC	Other subsidies on production
Losses on market	COLS: LOSM	T_TTRM	Trade and transport margins
Total output	ROWS: TOOU	X_TOOU	Total output
Total intermdiate input	ROWS: TOIN	X_TOIN	Total intermdiate input
Gross value added at producer prices	ROWS: GVAP	X_GVAP	Gross value added at producer prices
Gross value added at basic prices	ROWS: GVAB	X_GVAB	Gross value added at basic prices
Fixed capital consumption equipment	ROWS: DEPM	X_DEPM	Fixed capital consumption equipment
Fixed capital consumption buildings	ROWS: DEPB	X_DEPB	Fixed capital consumption buildings
Depreciation others	ROWS: DEPO	X_DEPO	Depreciation others
Net value added at basic prices	ROWS: NVAB	X_NVAB	Net value added at basic prices

Table 17. Modified Agricultural Classification and Correspondence to GTAP Sector Classification

Modified agricultural classification (MAC)		GTAP Classifications	
Description	**Code**	**Code**	**Description**
Other wheat	A/C_OWHE	WHT	Wheat
Durum wheat	A/C_DWHE	WHT	Wheat
Barley	A/C_BARL	GRO	Cereal grains nec
Grain maize	A/C_MAIZ	GRO	Cereal grains nec
Other cereals	A/C_OCER	GRO	Cereal grains nec
Paddy rice	A/C_PARI	PDR	Paddy rice
Rape seed	A/C_RAPE	OSD	Oil seeds
Sunflower seed	A/C_SUNF	OSD	Oil seeds
Soya seed	A/C_SOYA	OSD	Oil seeds
Other oil plants	A/C_OOIL	OSD	Oil seeds
Other starch and protein plants	A/C_STPR	OCR	Crops nec
Potatoes	A/C_POTA	OCR	Crops nec
Sugar beet	A/C_SUGB	C_B	Sugar cane, sugar beet
Fibre plants	A/C_FIBR	PFB	Plant-based fibers
Grapes	A/C_GRPS	V_F	Vegetables, fruit, nuts
Fresh vegetables, fruit, and nuts	A/C_FVEG	V_F	Vegetables, fruit, nuts
Live plants	A/C_LPLT	OCR	Crops nec
Other crop products	A/C_OTCR	OCR	Crops nec
Fodder crops	A/C_FODD	OCR	Crops nec
Set aside	A_SETA	OCR	Crops nec
Raw milk from bovine cattle	A/C_COMI	RMK	Raw milk
Bovine cattle, slaughtered	A/C_LCAT	CTL	Bovine cattle, sheep and goats, horses
Swine, slaughtered	A/C_PIGF	OAP	Animal products nec
Raw milk from sheep and goats	A/C_SGMI	RMK	Raw milk
Sheep, goats, horses, asses, mules and hinnies, slaughtered	A/C_LSGE	CTL	Bovine cattle, sheep and goats, horses
Eggs	A/C_EGGS	OAP	Animal products nec
Poultry, slaughtered	A/C_PLTR	OAP	Animal products nec
Wool and animal hair; Silk-worm cocoons suitable for reeling	A/C_ANHR	WOL	Wool, silk-worm cocoons
Other animals, live, and their products	A/C_OANM	OAP	Animal products nec
Agricultural services	A/C_AGSV	OBS	Business services nec
Products of forestry, logging and related services	A/C_FORE	FRS	Forestry
Fish and other fishing products; services incidental of fishing	A/C_FISH	FSH	Fishing
Coal and lignite; peat	A/C_COAL	COA	Coal
Crude petroleum and natural gas; services incidental to oil and gas extraction excluding surveying	A/C_COIL	OIL_GAS	Oil and Gas
Uranium and thorium ores	A/C_URAN	OMN	Minerals nec
Metal ores	A/C_MEOR	OMN	Minerals nec
Other mining and quarrying products	A/C_OMIN	OMN	Minerals nec
Rice, milled or husked	A/C_RICE	PCR	Processed rice
Other food products	A/C_OFOD	OFD	Food products nec
Processed sugar	A/C_SUGA	SGR	Sugar
Vegetable oils and fats, crude and refined; oil-cake and other solid residues, of vegetable fats or oils	A/C_VOIL	VOL	Vegetable oils and fats
Dairy products	A/C_DAIR	MIL	Dairy products
Meat of bovine animals, fresh, chilled, or frozen	A/C_BFVL	CMT	Bovine meat products
Meat of swine, fresh, chilled, or frozen	A/C_PORK	OMT	Meat products nec
Meat of sheep, goats, and equines, fresh, chilled, or frozen	A/C_SGMT	OMT	Meat products nec
Meat and edible offal of poultry, fresh, chilled, or frozen	A/C_POUM	OMT	Meat products nec
Beverages	A/C_BEVR	B_T	Beverages and tobacco products
Prepared animal feeds	A/C_ANFD	OFD	Food products nec
Tobacco products	A/C_TOBA	B_T	Beverages and tobacco products
Textiles	A/C_TEXT	TEX	Textiles
Wearing apparel; furs	A/C_GARM	WAP	Wearing apparel

Modified agricultural classification (MAC)		GTAP Classifications	
Description	**Code**	**Code**	**Description**
Leather and leather products	A/C_LETH	LEA	Leather products
Wood and products of wood and cork (except furniture); articles of straw and plaiting materials	A/C_WOOD	LUM	Wood products
Pulp, paper and paper products	A/C_PULP	PPP	Paper products, publishing
Printed matter and recorded media	A/C_MDIA	PPP	Paper products, publishing
Coke, refined petroleum products and nuclear fuels	A/C_RPET	P_C	Petroleum, coal products
Chemicals, chemical products and man-made fibres	A/C_CHEM	CRP	Chemical, rubber, plastic products
Rubber and plastic products	A/C_PLST	CRP	Chemical, rubber, plastic products
Other non-metallic mineral products	A/C_NMMP	NMM	Mineral products nec
Basic metals	A/C_FRMT	I_S_NFM	Ferrous metals and Metals nec
Fabricated metal products, except machinery and equipment	A/C_FAME	FMP	Metal products
Machinery and equipment n.e.c.	A/C_MACH	OME	Machinery and equipment nec
Office machinery and computers	A/C_OFMA	ELE	Electronic equipment
Electrical machinery and apparatus n.e.c.	A/C_ELMA	OME	Machinery and equipment nec
Radio, television and communication equipment and apparatus	A/C_COEQ	ELE	Electronic equipment
Medical, precision and optical instruments, watches and clocks	A/C_MEIN	OME	Machinery and equipment nec
Motor vehicles, trailers and semi-trailers	A/C_MOTO	MVH	Motor vehicles and parts
Other transport equipment	A/C_TREQ	OTN	Transport equipment nec
Furniture; other manufactured goods n.e.c.	A/C_FURN	OMF	Manufactures nec
Secondary raw materials	A/C_RECY	OMF	Manufactures nec
Electrical energy, gas, steam and hot water	A/C_EGSW	ELY_GDT	Electricity and Gas manufature, distribution
Collected and purified water, distribution services of water	A/C_WATR	WTR	Water
Construction work	A/C_CONS	CNS	Construction
Trade, maintenance and repair services of motor vehicles and motorcycles; retail sale of automotive fuel	A/C_SMOT	TRD	Trade
Wholesale trade and commission trade services, except of motor vehicles and motorcycles	A/C_WTRD	TRD	Trade
Retail trade services, except of motor vehicles and motorcycles; repair services of personal and household goods	A/C_RTRD	TRD	Trade
Hotel and restaurant services	A/C_HORE	TRD	Trade
Land transport; transport via pipeline services	A/C_TRLD	OTP	Transport nec
Water transport services	A/C_TRWA	WTP	Water transport
Air transport services	A/C_TRAR	ATP	Air transport
Supporting and auxiliary transport services; travel agency services	A/C_TROT	OTP	Transport nec
Post and telecommunication services	A/C_TECO	CMN	Communication
Financial intermediation services, except insurance and pension funding services	A/C_FINA	OFI	Financial services nec
Insurance and pension funding services, except compulsory social security services	A/C_INSU	ISR	Insurance
Services auxiliary to financial intermediation	A/C_FIAX	OFI	Financial services nec
Real estate services	A/C_ESTA	OBS	Business services nec
Renting services of machinery and equipment without operator and of personal and household goods	A/C_MARE	OBS	Business services nec
Computer and related services	A/C_COMP	OBS	Business services nec
Research and development services	A/C_RESC	OBS	Business services nec
Other business services	A/C_OTBS	OBS	Business services nec
Public administration and defence services; compulsory social security services	A/C_PUAD	OSG	Public Administration, Defense, Education, Health
Education services	A/C_EDUC	OSG	Public Administration, Defense, Education, Health
Health and social work services	A/C_HESO	OSG	Public Administration, Defense, Education, Health
Sewage and refuse disposal services, sanitation and similar services	A/C_SANI	OSG	Public Administration, Defense, Education, Health
Membership organisation services n.e.c.	A/C_ORGA	OSG	Public Administration, Defense, Education, Health
Recreational, cultural and sporting services	A/C_CULT	ROS	Recreational and other services
Other services	A/C_OTSV	ROS	Recreational and other services
Private households with employed persons	A/C_PRHH	ROS	Recreational and other services

Appendix 2: Definition of Tax Rates

Table 18. Definition of Tax Rates

Tax	Definition	Comments
Value added type taxes (VAT) (D.211)	A value added type tax (VAT) is a tax on goods or services collected in stages by enterprises and which is ultimately charged in full to the final purchasers.	VAT is recorded net in the sense that: a) outputs of goods and services and imports are valued excluding invoiced VAT; b) purchases of goods and services are recorded inclusive of non-deductible VAT. VAT is recorded as being borne by purchasers, not sellers, and then only by those purchasers who are not able to deduct it. The greater part of VAT is therefore recorded in the system as being paid on final uses, mainly on household consumption. A part of VAT may, however, be paid by enterprises, mainly by those which are exempted from VAT. For the total economy, VAT is equal to the difference between total invoiced VAT and total deductible VAT (see point 4.27).
Taxes and duties on imports excluding VAT (D.212)	Taxes and duties on imports excluding VAT (D.212) comprise compulsory payments levied by general government or the Institutions of the European Union on imported goods, excluding VAT, in order to admit them to free circulation on the economic territory, and on services provided to resident units by non-resident units.	These payments include: a) import duties (D.2121): these consist of customs duties, or other import charges, payable according to customs tariff schedules on goods of a particular type when they enter for use in the economic territory of the country of utilisation; b) taxes on imports, excluding VAT and import duties (D.2122):
Taxes on products, except VAT and import taxes (D.214)	Taxes on products, except VAT and import taxes (D.214) consist of taxes on goods and services that become payable as a result of the production, export, sale, transfer, leasing or delivery of those goods or services, or as a result of their use for own consumption or own capital formation.	This heading includes, in particular: a) excise duties and consumption taxes (other than those included in taxes and duties on imports); b) stamp taxes on the sale of specific products, such as alcoholic beverages or tobacco, and on legal documents or cheques; c) taxes on financial and capital transactions, payable on the purchase or sale of non-financial and financial assets, including foreign exchange. They become payable when the ownership of land or other assets changes, except as a result of capital transfers (mainly inheritances and gifts). They are treated as taxes on the services of intermediaries; d) car registration taxes; e) taxes on entertainment; f) taxes on lotteries, gambling and betting, other than those on winnings; g) taxes on insurance premiums; h) other taxes on specific services: hotels or lodging, housing services, restaurants, transportation, communication, advertising; i) general sales or turnover taxes (excluding VAT type taxes): these include manufacturers' wholesale and retail sales taxes, purchase taxes, turnover taxes; j) profits of fiscal monopolies which are transferred to the State, except those exercising a monopoly over the imports of some good or services (included in D2.122). Fiscal monopolies are public enterprises which have been granted a legal monopoly over the production or distribution of a particular kind of good or service in order to raise revenue and not in order to further the interests of public economic or social policy. When a public enterprise is granted monopoly powers as a matter of deliberate economic or social policy because of the special nature of the good or service or the technology of production – for example, public utilities, post offices and telecommunications, railways and so on – it should not be treated as a fiscal monopoly. As a general rule, fiscal monopolies are typically engaged in the production of goods or services which may be heavily taxed in other countries; they tend to be confined to the production of certain consumer goods (alcoholic beverages, tobacco, matches, etc.) or fuels; k) export duties and monetary compensatory amounts collected on exports.

Tax	Definition	Comments
Subsidies on products (D.31)	Subsidies on products (D.31) are subsidies payable per unit of a good or service produced or imported. The subsidy may be a specific amount of money per unit of quantity of a good or service, or it may be calculated ad valorem as a specified percentage of the price per unit. A subsidy may also be calculated as the difference between a specified target price and the market price actually paid by a buyer. A subsidy on a product usually becomes payable when the good is produced, sold or imported. By convention, subsidies on products can only pertain to market output (P.11) or to output for own final use (P.12).	Import subsidies (D.311) consist of subsidies on goods and services that become payable when the goods cross the frontier for use in the economic territory or when the services are delivered to resident institutional units. They may include losses incurred as a matter of deliberate government policy by government trading organisations whose function is to purchase products from non-residents and then sell them at lower prices to residents. Other subsidies on products (D.319) include: a) subsidies on products used domestically: these consist of subsidies payable to resident producers in respect of their production which is used or consumed within the economic territory; b) losses of government trading organisations whose function is to buy the products of resident producers and then sell them at lower prices to residents or non-residents, when they are incurred as a matter of deliberate government economic or social policy; c) subsidies to public corporations and quasi-corporations to compensate for persistent losses which they incur on their productive activities as a result of charging prices which are lower than their average costs of production as a matter of deliberate government or European economic and social policy; d) direct subsidies on exports payable directly to resident producers when the goods leave the economic territory or the services are provided to non-residents – except repayments at the customs frontier of taxes on products previously paid and waiving of the taxes that would be due if the goods were to be sold or used inside the economic territory.

European Commission

EUR 24010 EN – Joint Research Centre – Institute for Prospective Technological Studies

Title: Construction of Social Accounting Matrices for EU27 with a Disaggregated Agricultural Sector (AgroSAM)

Authors: Marc Müller, Ignacio Pérez Domínguez, Stephan Hubertus Gay

Luxembourg: Office for Official Publications of the European Communities
2009

EUR – Scientific and Technical Research series – ISSN 1018-5593
ISBN 978-92-79-13625-2
DOI 10.2791/29021

Abstract

This technical report provides an overview on the AgroSAM project carried out at the Institute for Prospective Technological Studies (IPTS). The objective was to create a set of Social Accounting Matrices with a disaggregated agricultural sector for the 27 EU Member States by combining national Supply- and Use Tables with data from the agricultural sector model CAPRI. In general, three stages of can be distinguished: First, the compilation of consolidated macroeconomic indicators for EU27. Second, the combination of different datasets from EuroStat into a set of Social Accounting Matrices with aggregated agricultural and food-industry sectors. Third, the disaggregation of these sectors based on CAPRI data. The methods applied for the balancing of the datasets at the three stages draw heavily on the concept of Cross Entropy estimation. Particularly the structural deviations of agricultural sector and economy-wide data created a need to specify in which cases comparatively large deviations from recorded agricultural data could be tolerated, and in which cases not. For this purpose, Cross Entropy procedures proved to be extremely useful. The finally estimated AgroSAM are a first step into the direction of creating a harmonized database for agricultural policy analysis within the Modelling Platform iMAP hosted at IPTS.

The mission of the Joint Research Centre is to provide customer-driven scientific and technical support for the conception, development, implementation and monitoring of European Union policies. As a service of the European Commission, the Joint Research Centre functions as a reference centre of science and technology for the Union. Close to the policy-making process, it serves the common interest of the Member States, while being independent of special interests, whether private or national.

How to obtain EU publications

Our priced publications are available from EU Bookshop (http://bookshop.europa.eu), where you can place an order with the sales agent of your choice.

The Publications Office has a worldwide network of sales agents. You can obtain their contact details by sending a fax to (352) 29 29-42758.

The mission of the JRC is to provide customer-driven scientific and technical support for the conception, development, implementation and monitoring of EU policies. As a service of the European Commission, the JRC functions as a reference centre of science and technology for the Union. Close to the policy-making process, it serves the common interest of the Member States, while being independent of special interests, whether private or national.

JRC
EUROPEAN COMISSION

LF-NA-24010-EN-C

Publications Office

ISBN 978-92-79-13625-2

9 789279 136252