TRADE4SD
Fostering the positive linkages between trade and sustainable development

Programme: H2020-EU.3.2.1.3. - Empowerment of rural areas, support to policies and rural innovation
Topic: RUR-21-2020 - Agricultural markets and international trade in the context of sustainability objectives
Call: H2020-RUR-2020-2
Type of action: Research and Innovation Action (RIA)
Duration of the project: 01 June 2021 – 31 May 2025

Deliverable 3.2: Report on the modelling of social and distributional impact of trade and sustainability policies

Tamas Révész*[1], Martin Banse[2], Mona Abdelhady[2], Sergio Rene Araujo Enciso[4], Jan Hagemejer[3], Błażej Jendrzejewski[3], Vitaliy Krupin[3], Verena Laquai[2], Holger Matthey[4], Katarzyna Zawalińska[3]

* Deliverable leader

CORVINUS UNIVERSITY OF BUDAPEST[1]
THÜNEN INSTITUTE[2]
CASE[3]
FAO[4]

Workpackage No. 3.
Due date: 30 September 2023 (M28)
Actual date: 29 September 2023
Dissemination level: Public

This project has received funding from the European Union's Horizon 2020 Research and Innovation programme under grant agreement No 101000551 – TRADE4SD

About TRADE4SD Project

Trade is a central factor in shaping not only global, but also regional and local development. Trade policy has an especially important part to play in achieving the UN Sustainable Development Goals (SDGs). The premise of the TRADE4SD project is that trade has the power to produce positive outcomes when the policies which define the rules of the game are framed and designed in a way to promote access to markets, fair prices and standards of living for farmers, as well as alleviating rural poverty and ensuring sustainable farming practices. Addressing the relation between trade and SDGs requires an integrated approach to policy-making and inclusive governance.

The main objective of the TRADE4SD project is to contribute to build new opportunities for fostering the positive sustainability impacts of trade supported by improved design and framing of trade policy at national, EU and global level, including WTO modernisation, increased policy coherence at different domains including agricultural, energy, climate, environmental and nutritional policies.

To meet this objective, the project will develop an integrated and systematic approach that combines quantitative models from different perspectives, and several qualitative methods recognising that SDGs and trade are highly context-related. On the one hand, a robust analysis of economic, social and environmental impacts is given by using diverse but integrated modelling techniques and qualitative case studies. On the other hand, a wide consultation process is implemented involving stakeholders both in the EU and in partner countries as well as those with a wide international scope of activity, providing opportunities for improved understanding, human capital building, knowledge transfer and dissemination of results. To this extent, the consortium involves, as co-producers of knowledge, a number of research and stakeholder participants with different backgrounds who will use their networks to facilitate the civil society dialogue and build consensus on the subject of gains from trade in view of sustainability.

Project Consortium

No.	Participant Organisation Name	Country
1	Corvinus University of Budapest (CORVINUS)	HU
2	University of Kent (UNIKENT)	UK
3	Consiglio per la Ricerca in Agricoltura e l'Analisi dell'Economia Agraria (CREA)	IT
4	Johann Heinrich von Thünen-Institut, Bundesforschungsinstitut für ländliche Räume, Wald und Fischerei (THUENEN)	DE
5	The University of Sussex (UOS)	UK
6	University of Ghana (UG)	GH
7	Luonnonvarakeskus (LUKE)	FI
8	Centrum Analiz Spoleczno-Ekonomicznych-Fundacja Naukowa (CASE)	PL
9	Food and Agriculture Organization of the United Nations (FAO)	IT
10	Institut national de recherche pour l'agriculture, l'alimentation et l'environnement (INRAE)	FR
11	Confederazione Generale Dell'Agricoltura Italiana (CONFAGRICOLTURA)	IT
12	Truong Dai Hoc Kinh Te Thanh Pho Ho Chi Minh (UEH)	VN
13	Luminaconsult Sprl (LUMINA)	BE

Table of contents

List of Tables

List of Figures

List of abbreviations

CDE	Constant Difference of Elasticities
CET	Constant Elasticity of Transformation
CGE	Computable General Equilibrium
C-D	Cobb-Douglas
COICOP	Classification of consumption categories
CPA	Classification of Products by Activity (of the EU)
CPC	Central Product Classification
ESTAT	Eurostat
EU	European Union
EU-SILC	EU Statistics on Income and Living Conditions
FAO	Food and Agriculture Organization of the UN
GAMS	General Algebraic Modeling System
GH¢	Ghana cedi, currency in Ghana from 2007.
GHG	Greenhouse gas
GLSS	Ghana Living Standard Survey
GSEC2	GTAP Sector classification, version 2
GSEC3	GTAP Sector classification, version 3
GTAP	Global Trade Analysis Project
HBS	Household Budget Survey
IIASA	International Institute for Applied System Analysis
ISIC	International Standard Industrial Classification
NACE	EU Statistical classification of economic activities
NUTS	Nomenclature of Territorial Units for Statistics
SAM	Social Accounting Matrix
SDG	Sustainable Development Goal
SDI	Sustainable Development Indicator
SNA	System of National Accounts
SSP	Shared Socioeconomic Pathways (Sustainability)
SWIID	Standardized World Income Inequality Database
UN	United Nations
UNDP	UN Development Program
UNDESA	UN Dep. of Economic & Social Affairs
WB	The World Bank
WP3	Work Package 3 of the project
WTO	World Trade Organization

1 INTRODUCTION

The role of Work Package 3 (WP3) of the TRADE4SD project is the 'quantitative model-based analysis of the sustainability impacts of agricultural trade'. Within WP3 four different but interrelated tasks were specified. Task 2 (from the point of view of the whole project Task 3.2) was 'Estimating the social and distributional impact of trade and sustainability policies'. To accomplish this, we used primarily the MAGNET model. The MAGNET model is a multiregional computable general equilibrium (CGE-) model based on the GTAP-model and has various extensions related to environmental, energy and agricultural issues.

The 2030 Agenda for Sustainable Development containing the Sustainable Development Goals (SDGs) were adopted at the United Nations Sustainable Development Summit on 25 September 2015. The United Nations 2030 Agenda includes 17 Sustainable Development Goals (SDGs) intended to apply universally to all countries. In November 2016 European Commission issued its 'first Communication' on the next steps for a sustainable European future which explains how the Commission's 10 political priorities contribute to implementing the UN 2030 Agenda for Sustainable Development and how the EU will meet the SDGs in the future.

Annual reviews of the EU SDG indicator set serve to provide for continuous policy relevance and to enhance the statistical quality. To each of the current 17 SDGs 6 sustainable development indicators (SDIs) are primarily attributed. Special attention was given to the possibility to include these indicators (at least by their proxy categories) into the MAGNET model. The MAGNET model can be regarded a good candidate for analysing the so far somewhat neglected/overlooked within-country social- and distributional effects of trade and agricultural policies. As Allen et al. (2016) concluded (see Figure 2 of their paper and the related discussion) after reviewing the strengths and weaknesses of 80 contemporary modelling tools in the context of national development planning for the SDGs:

'In terms of sustainability dimensions, the most common approach was the integration of economic and environmental variables in the model. A total of 54 models (or 68%) integrated these two dimensions to varying degrees, highlighting the rapidly growing catalogue of economy-environment models. The social dimension of sustainable development is by far the least addressed, with only 17 models (or 21%) including social variables within their modelling framework, and often with very limited coverage (most commonly a few health-related or nutrition-related variables). The modelling of social variables can therefore also be considered as an important gap in modelling capabilities.'

More concretely they point out that '*key gaps or thematic issues requiring further model development include poverty, health, education, gender, inequality, sustainable consumption and production, biodiversity and governance-institutions.*'

They also observed that *'only eight models (10%) met the two screening criteria of 'policy relevant' and 'integrated'. The shortlist comprised one top-down CGE model (MAGNET), one top-down system dynamics model (Polestar), and six hybrid models (IMPACT; International Futures; Threshold 21; EC4MACS; InVEST; and LowGrow).'*

However, the authors found that due to the highly technical nature and complexity, the MAGNET model is less transparent, less easy to use, and more cost-, time- and effort-requiring for model development. Fortunately, the GTAP-model – which is the core of the MAGNET model – has been extensively documented in a recent article (Corong et al. (2017)) and since then the developers of the MAGNET-model themselves have published many other papers and monographies highlighting the various modules of the model and its applications. If *'the modelling of social variables can therefore also be considered as an important gap in modelling capabilities'* this is even more evident in particular in the case of the SDIs, most of which are hard to be found in any of the 80 models reviewed by Allen et al. (2016). Therefore, to fill this gap, in our research we concentrated on identifying those social – in particular distributional – indicators which may be linked with the SDIs in the MAGNET model. Then by splitting the household sector to various social groups in the MAGNET model and by running different simulations we estimated the benchmark and future level of various distributional indicators.

Therefore, CGE-modelling is a rather complex activity which has many steps. In the present report we describe these steps in a logical order. More detailed presentation and discussion of the individual steps and aspects of the modelling can be found in background papers, interim and internal reports which we duly refer to.

This text is the final deliverable D3.2 of work package 3 in the TRADE4SD project. It builds on the earlier D3.1 deliverable report titled 'Report on linking SDG indicators with models in the TRADE4SD toolbox' (Laquai et al. (2022)) and by concentrating on the SDG indicators which have a social dimension uses household microdata – primarily the Ghanaian household survey but also the surveys for various EU- and developing countries – to disaggregate the household sector of the MAGNET-model to render possible the assessment of the expected development of these SDG indicators by running various trade and agriculture policy simulations. Disaggregation of the household sector is accomplished in different break-downs to measure the various aspects of the 'inequality within countries' component of the SDG10 ('Reduce inequality within and among countries'). In addition, by computing endogenously the labor incomes and consumption of each household group by product/sector one can assess how the given groups' poverty status, employment, nutritional level, access to various goods and services, etc. are expected to change as a result of these policies. By running similar simulation by the other models of WP3 – the scenario of which also include more energy- and environmental policy related components – the coverage of estimated SDG indicators will increase so that an integrated energy-environment-social assessment of these policies will be possible. This will be done in cooperation of the workgroups dealing with Task 3.3 and 3.4 and the results will be included in the deliverable reports related with these tasks.

After the general introduction and overview of the goal and context of this deliverable report, the structure of the paper can be summarized as follows:

In Chapter 2 the strengths and weaknesses of the CGE-models are discussed from the point of view of estimating social/distributional effects and some key relevant features of the MAGNET model are presented. Special attention is paid to the benefits and challenges of introducing multiple households into CGE models and to the possible economic savings and consumption function of the represented household groups. The capacities of the CGE-models depend very much on the availability and quality of the data. Therefore, Chapter 3 and 4 deal with Household Budget Surveys (HBS) of the EU countries and selected developing countries. Since we have chosen Ghana to test and illustrate the developed method for estimating social/distributional effects, special emphasis is given to the Ghanaian GLSS5 and GLSS7 HBSs in Chapter 4. Accordingly, Chapter 5 presents the social/distributional effects of the baseline and trade liberalization simulation of the MAGNET-model using the GLSS5 data and splitting the household sector to various social groups. The last chapter concludes by summarizing the achieved results of the research, the remaining issues and the possible/desirable directions of the future research.

2 MODELS FOR ESTIMATING THE SOCIAL AND DISTRIBUTIONAL IMPACT OF TRADE AND SUSTAINABILITY POLICIES

The United Nations Department of Economic and Social Affairs (UNDESA) and the United Nations Development Programme (UNDP) have pioneered a series of modelling tools. Early initiatives included economy-wide modelling and microsimulation methodologies. Five quantitative modelling tools are being used by UNDESA and UNDP to help countries assess sustainable development policy options. One of them is a microsimulation model which offers insights for policies to eradicate poverty, reduce inequality, enhance food security and broaden access to energy, among others. The technique has been applied to policies related to taxes or subsidies, cash or in-kind transfers, and expanded access to modern energy, among other examples.

However, apart from the data availability challenges of such microsimulation models, they are not sufficiently 'integrated', i.e. they do not represent the national level and global repercussions and interlinkages of the various categories of sustainable development (see Allen et al., 2016).

To reveal and quantify linkages between trade and trade policies with the SDIs, in the project we use partial equilibrium (PE) models (AGEMOD and AGLINK_COSIMO) and general equilibrium (GE) models (MAGNET and GTAP_CGE box). In view of the latter, special attention will be given to the possibility to include these indicators (at least by their proxy categories) into the MAGNET model - which is regarded to be a model where many SDGs-related policies are addressed. MAGNET can be further extended so that it includes more social impacts of international trade and trade-related policies.

The standard GTAP model (Version 6.2 of September 2003) was the starting point for developing the MAGNET model. GTAP is a general equilibrium model covering all sectors of the economy (agriculture, manufacturing and services) as opposed to partial equilibrium models such as AGMEMOD, which focuses on subsets of an economy. In addition, GTAP is a global model, covering all regions and major countries in the world.

2.1 The 'regional household' of the GTAP-model

For every region in the model there is a single representative household. Following roughly the presentation of Kuiper and Shutes (2014) in the following subsection we will review the meaning of and the assumptions about the 'regional household' in the GTAP model.

The regional household supplies factors (land, skilled and unskilled labor, capital and natural resources) to the production sectors. The income of the 'regional household' consists of income earned from land, labor and capital, as well as income raised from taxes.
The regional household allocates its income over private expenditures, government expenditures and savings. In terms of the GTAP categories and notations expenditures need to match income earned in each region (denoted by r):

$$\sum_e VOM(e,r) - VDEP(r) + INDTAX(r) = PRIVEXP(r) + GOVEXP(r) + SAVE(r),$$

where eVOM(e,r) is the total (gross) income from factor e, VDEP(r) is the total depreciation in the region, INDTAX(r) is the (net) indirect tax revenue, PRIVEXP(r) is the private consumption expenditure, GOVEXP(r) is the government consumption expenditure and SAVE(r) is the savings.

The allocation of the income to these three components is governed by a Cobb-Douglas utility function. However, the Cobb-Douglas function yields constant and non-negative expenditure shares and therefore does not allow for the negative savings that occur in the disaggregated GTAP database. Verikios and Hanslow (1999) address this problem in version 4 of the GTAP database by offsetting the required increase in savings by a reduction in depreciation in regions with negative savings. However, this approach does require VDEP to be larger than the negative savings. Unfortunately, this condition is not satisfied in seven of the 24 regions in the V8 2007 database; we therefore cannot rely on this approach to accommodate negative savings (Woltjer – Kuiper (2014)).

Negative savings in a region can be interpreted as a transfer of income from other regions. In the GTAP database, savings balance with net investment at the global level. This implies that the negative savings of countries are in fact deducted from the savings of all other countries. This approach to modelling savings and investment allows us to introduce transfers. These transfers make the balancing of savings over all regions explicit and thereby remove negative savings. They do not alter the total income of the regions. Countries with positive savings will now spend part of their income on transfers, which are deducted from their savings. Countries with negative savings receive transfers that offset their negative savings, thus keeping their total income balanced as well (Woltjer – Kuiper (2014)).

Clearly, the above approach is questionable, and even impossible to maintain when one disaggregates the household sector to various socio-economic groups. The various options for introducing multiple households in the model are discussed in the following section.

2.2 The possible methods for disaggregating the ‘regional household’

There have been numerous calls for a more detailed accounting system for the GTAP regional household, especially as it relates to estimating the potential impacts of policies and global shocks on poverty, sustainable and inclusive growth. The requirements for the break-down of the household sector to groups is that these groups be representative of key demographic and income groups, such as rural and urban households, and secondly, that there have clear linkages to the sources of the household’s income.

The MyGTAP model (Walmsley and Minor, 2013) is an extension of the standard GTAP CGE model that splits the regional household into three main agents - government, private households and investment - and links income to factor payments and taxes in each. Furthermore, the MyGTAP model includes multiple private household types for a single region of choice. This yields a global model, not unlike many single country models, which can be used for multiple household analysis including distributional issues. In all regions the regional household, which collects income, saves and allocates income to the private household and government, in the standard GTAP model, is removed and replaced with explicit income flows to households and government. Households receive income from the supply of factor services, remittances, foreign income and transfers and save a share of their income. The government receives income from tax revenues and aid payments and (dis)saves a share of its income which forms the internal balance.

Essentially, two approaches have been used in the literature to include multiple household types in the database of global CGE models: the weights approach (or SAM approach) adopted in MyGTAP (Minor and Walmsley (2013)) and the household survey approach adopted in MIRAGE-HH (Bouët et al. (2013)). The MyGTAP data procedure currently enables multiple household types and other developing country features such as remittances and aid to be added for a single region of the GTAP database. It requires the user to define the weights for the splitting of household consumption and factor income to households, although default weights are provided as a starting point. These weights could be taken from household survey data or national SAM accounts. MyGTAP makes use of an entropy procedure to bring the user supplied weights in line with the GTAP data. While the user can interact with the procedure by viewing the output at each stage of the iteration, it is preferable to avoid automatic balancing procedures (Kuiper and Shutes (2014)).

1. Household survey approach MIRAGE-HH (Bouët et al., 2013). In MIRAGE-HH, the multiple household types are constructed using a ‘bottom-up’ approach from the household survey data. The advantages of this approach include the use of the most recent household survey data, no limitation on the number of household types that can be included and updating of the GTAP values where the household survey data are deemed to be more accurate. The disadvantages are that the process is very time-consuming and is limited to countries for which household survey data are available. These are indeed rich data sets but costly to construct in terms of processing time.

2. The SAM approach represents a method which uses household accounts in national SAMs that have been constructed from household survey data. First, we present the scheme of the SAM which can be constructed from the GTAP database itself (Fig. 1.)

Kuiper and Shutes (2014) demonstrate the SAM approach with the SAM for Ghana created by Breisinger et al. (2007). This Ghana national SAM includes more detail than the GTAP SAM of the same region, including 70 commodities, 142 activities (with regional production for 27 agricultural activities), 13 factors of production and 9 households. The nine representative household groups are location based: Accra, Urban Coast, Urban Forest, Urban South, Urban North, Rural Coast, Rural Forest, Rural South and Rural North. The scheme of the Ghana national SAM can be seen in the following figure, which also highlights those 11 differences in which parts of the table are different from the original GTAP SAM:

Figure 1. The scheme of the SAM for Ghana with 9 household groups (where the numbers refer to the steps of the adjustment procedure)

	Imported commodities	Domestic commodities	Activities	Factors	Private household	Taxes	Government	Capital	Margins	Rest of world	Total
Imported commodities			Imported intermediate inputs		Household demand for imports (2) (7)		Government demand for imports (2)	Investment demand for imports			Import demand
Domestic commodities			Domestic intermediate inputs		Household domestic demand (2) (7)		Government domestic demand (2)	Investment domestic demand	Exports of transport services	Exports of goods and services	Domestic demand
Activities		Supply matrix									Output demand
Factors			Payments to factors								Factor income
Private household: Accra, Urban coast, Urban forest, Urban south, Urban north, Rural forest, Rural south, Rural north, Rural Accra				Distributed factor income (3) (4)	Inter-household transfers (5)		Government transfers to households (6)				Private household income
Taxes	Import duties	Export duties	Sales, factor, production taxes	Factor income taxes (9)	Sales taxes (10)		Sales taxes	Sales taxes			Tax revenue
Government				Factor income (3)		Tax revenue (3)					Governme nt income
Capital				Depreciation (11)	Household savings (1) (8)		Government savings (1)		Trade balance	Trade balance	Savings
Margins	Transport margins on imports								Imports of transport margins		Margin income
Rest of world	Imports of goods and services										Payments to rest of world
Total	Supply of imports	Domestic supply	Domestic output	Factor payments	Private household expenditure	Tax payments	Government expenditure	Investment	Margin expenditure	Income from rest of world	

Source: GSS-IFPRI (2007)

The accounting steps which resulted the 11 differences between the original GTAP SAM and the GTAP SAM with multiple household types (see Figure 1) are the following:

1. Separate private household and government savings
2. Separate private household and government expenditure
3. Explicit income flows to the single private household and government
4. Explicit factor income flows to multiple households
5. Inter-household transfers
6. Government transfers to multiple households
7. Demand for domestic and imported commodities by household type
8. Household savings by household type
9. Household specific direct taxes
10. Sales taxes by household type
11. Capital stock holdings and depreciation by household type and government.

The above steps are explained and detailed in Kuiper and Shutes (2014). Here we can mention only that the above transfer categories may be disaggregated further.
It should be also noted that Ghana is used here as an example, but the method presented is relevant for any country for which a national SAM with household detail is available (see section 6 for a list of readily available national SAMs with household detail for other regions). However, Kuiper and Shutes (2014) give a list of those developing countries where at the time of writing SAMs with multiple households were available.

Since disaggregating the household sector has to be done also along the regional dimension we reviewed those EU-level models which are calibrated for the NUT2 regions.

The greatest challenge in building a regional general equilibrium model for all EU27 NUTS2 regions is the *database construction*. The main steps needed to construct such a database, including the Social Account Matrices (called IOTNUTS2 and described by Mueller and Ferrari, 2011), can be summarized as follows:

Addressing regional heterogeneity requires multi-sector data on a sub-national scale. Such datasets as are available are usually not sufficiently detailed, which gave rise to numerous non-survey methods to generate regional IOTs based on combinations of regional indicators and national datasets. At national level, some attempts to construct consistent regionalised tables have been pursued, mainly by National Statistical Offices (NSO) following survey-based methods (i.e. Finland, OFS) or national research institutes following non-survey-based methods and link them to multi-sectoral regionalised national models. To the best of our knowledge, a complete set of SAMs for all the EU NUTS2 does not yet exist and this work fulfils this deficiency in the literature. Altogether 280 SAMs at the NUTS 2 level for most European countries are available. The regional SAMs were enriched with details for agriculture from the CAPRI data base (Britz and Witzke (2014), Britz and van der Mensbrugghe (2018)).

One of the most well-known multiregional CGE-model is the RHOMOLO model (RHOMOLO, 2018) of the European Commission. Although its latest technical description (downloadable from the https://joint-research-centre.ec.europa.eu/tedam/rhomolo-model_en website) contains such model equations where the income, savings and consumption of the households of each NUTS2 regions are determined endogenously so that even the consumption is detailed by products, the also downloadable dataset does not contain these details. Only the aggregate (one-sector) national Input-Output tables are disaggregated to NUTS2 regions (i.e. a multiregional input-output table is shown) but no data for the transfers and savings and the product composition of the consumption is given. The technical description of the compilation of the dataset (JRC (2023), also downloadable from the https://publications.jrc.ec.europa.eu/repository/handle/JRC132883 website) claims that the regionalization of the national level data was based on the NUTS2 level regional accounts data and the structural business statistics data of the Eurostat. However, neither of these contains NUTS2 level data of household consumption and its composition. Instead, these are estimated by using proportionality assumptions (based, for example, on the share of the regions in the total national value added, or in the total output of the given sector) and by standard balancing techniques so that the total consumption is assumed to be proportional to the total household incomes of the regions (coded as (NAMA_10R_2HHINC in the Eurostat Database). Obviously, the full dataset needed for the calibration of the model's (not even too complicated) household behaviour equations is much more comprehensive and detailed than what they share with the public as 'publicly available RHOMOLO data'.

In summary, the rather vague documentation of these regional datasets makes one suspect that these models are not too rich in terms of details and specificities, but it is worth studying the issue further in the near future.

2.3 The CGEBox model

GAMS based global CGE models which come close to the GTAP Standard model have been available for a while, such as the well-known GTAPinGAMS implementation, see Rutherford and Arbor (2005) and Lanz and Rutherford (2016).
However, the first faithful replication was only recently available, which provided the starting point for the work on CGEBox, see Britz and van der Mensbrugghe (2016), combined with a Graphical User Interface (GUI).
The original code is largely based on the GAMS code of ENVISAGE (van der Mensbrugghe, 2008) and therefore comprises many features found in ENVISAGE. That rendered it inviting to not only replicate version 7 of the GTAP Standard model, but to also allow for variants and extensions based on a modular concept, see Britz and van der Mensbrugghe (2018). More specific details about CGEBox, its method and mechanisms are described in the official model's documentation (Britz, 2021).

2.4 The MAGNET model

MAGNET is a computable general equilibrium model (Woltjer et al., 2014). It is an extension of the GTAP model (Hertel, 1997) so the standard GTAP model and its database is the core of MAGNET. It covers all sectors of the economy (agriculture, manufacturing and services) and all regions and major countries in the world. The model is used especially for trade, biofuel, agricultural and other policy analyses as well as for long-term projections of the world economy. It has been developed at Wageningen Economic Research and is applied and further extended at Wageningen Economic Research, the Thünen Institute and the Joint Research Centre of the European Commission.

The MAGNET model is based on neo-classical microeconomic theory. On the consumption side, one household per region is distinguished. The household incomes are allocated to private and government consumption and savings using fixed budget shares. In GTAP, private (household) consumption behaviour is modelled via a Constant Difference of Elasticity (CDE) function, which is a relatively flexible, non-homothetic function allowing for non-constant marginal budget shares. It is calibrated by GTAP using data on income and price elasticities of demand. Since the use of the CDE-function in practice results in constant income elasticities over time – leading to unrealistically high levels of consumption of food items in fast growing economies – the extended MAGNET model allows income elasticities to adjusted over time so that income elasticities of consumption are decreasing function of purchasing power parity corrected real GDP per capita. This approach can be implemented by activating the consumption module.

In MAGNET production technology is represented by a nested CES-function which implies constant returns to scale but the CDE consumption function means that income has an impact on the spending pattern. Thus, from the demand side there is a rational to distinguish different types of households. Differences in household endowment ownership patterns will create differences in income and thus spending patterns, changing total demand compared to a model with a single representative household and thus market prices.

The GTAP database (Aguiar et al., 2019) is the core database of the MAGNET. However, multiple satellite databases complement it (Woltjer et al., 2014) due to the various extensions of MAGNET compared to GTAP. The various extensions in MAGNET are modelled in a modular way so that they can be switched on and off. This makes MAGNET flexible and ready to be applied to various research questions.

The MAGNET household module is based on MyGTAP which is a version of the GTAP model developed by Minor and Walmsley, see Walmsley and Minor (2013) and Minor – Walmsley (2012). In this module the different households (groups) receive income from factor payments according to their distribution of factor ownership. This results in different levels of income for different households. The proportion of income that is saved is household specific and the rest of the income is spent on consumption demand according to the household specific pattern of consumption.

MyGTAP also includes remittances, foreign income flows and aid payments. These are however difficult to integrate into the flexible MAGNET system because of the international nature of the flows. Remittances entering a country with multiple households must originate from another country which in the MAGNET set-up has a representative household formulation. To be consistent however, the remittances should be included both as an item of household income in the recipient country and as a household expenditure on the donor country which would entail changing the specification of household expenditure in standard GTAP regions. Given this complication and the fact that these international flows are not necessary for the modelling of multiple households in MAGNET, the decision was taken to omit this portion of the code until a proper representation of the remittances can be found.

2.5 Model assumptions about the behaviour of the social groups

Transfer incomes are heavily underreported in Household Budget Surveys (e.g. 16 % of the households of the GLSS5 survey in Ghana did not report any incomes!) and are rather difficult to explain by economic theory. We assume that they are proportional to the labor incomes or at least their ratio is exogenous, which on the individual household level would be quite unrealistic assumption but on a household group level there are reasonably arguments to support it. For example, the pension incomes are quite strongly correlated with earlier earnings, so that if in a household group both active earners and their retired counterparts (in other respect similar households) are included then the pension income (which in the EU is the largest component of the transfer incomes) of this group may be assumed to be proportional to the labor income of the group. Note, that since the ratio of transfer incomes to labor incomes may be different across household groups, this proportionality assumption does not mean that on the level of the whole population such proportionality exists.

Savings are also rather underrepresented in the HBSs or even not represented at all. Computing savings residually as the difference between incomes and expenditures is also quite questionable in particular when the household groups are defined as income or consumption deciles. It was already observed by Deaton (1997) that since both consumption and income data contain a large component of statistical errors, temporary effects and other random effects, the residually computed savings are quite distorted, even meaningless.
Nonetheless, since most models – like the aforementioned RHOMOLO model – assume fixed savings rates (i.e. savings/income ratios) savings may be assumed to be proportional to income at least on the household level, while naturally allowing for different savings rates across the defined household groups.

Note, that since the components of the savings are not known (this is obvious in the case of residually computed savings but it is also well-known for economic statisticians that the components of savings are rather differently reported in the surveys and even defined so differently that renders their cross-country comparison practically impossible) we have to deal only with the net savings. The same can be argued for the transfers. Also note, that many respected modelers like André Lemelin claim that net categories (like changes in stocks, net taxes, net exports, etc. in the input-output tables) should not be attempted to be explained as such net categories but rather their plus and minus components have to be separated first and then these components have to be estimated separately. Therefore, we did not attempt to try to work out more sophisticated assumptions about the relationship between incomes and net transfers and/or net savings.

Since both (net) transfers and (net) savings are assumed to be proportional to labor income this implies that the total consumption expenditure is proportional to the income too.

Finally note, that when analysing the simulation results, we usually compare two simulations and hence we are interested only in the *marginal* behaviour of the households. In other words, we are only interested in that when the income of certain households increase how they spend their additional income. Our proportionality assumptions therefore are restricted to the relationship of such incremental values.

2.6 Formal description of the chosen method

The formal description of the method chosen for this report can be summarized in the following steps/assumptions:

- Relative shares of the households (groups, denoted by the h index) in the labor income of economic (production) sectors (denoted by the i index) are assumed to be constant (and denoted by $s^{w}_{h,i}$).
- Total labor income (which may be regarded to be the real wage earnings) of the households (denoted by y^{w}_{h}) is the sum of their labor income received from the sectors (denoted by $w_{h,i}$) :
  $y^{w}_{h} = \sum_i s^{w}_{h,i} \cdot w_i$, where w_i is the total renumeration of labor paid by sector i.
- Total income (denoted by y_h) is proportional to the total labor income of the given household so that the proportionality factors (denoted by r_h) differ across households:
  $y_h = r_h \cdot y^{w}_{h}$
- The demand for each (composite, domestic and import together) consumption good (q^{i}_{h}, where the consumption goods are distinguished by their producing sectors) is computed so that the initial level (denoted by qo^{i}_{h}) is modified by the given good's income and own-price elasticities and by the cross-price elasticities (the income elasticities are denoted by $\gamma_{i,h}$ while the price elasticities by $\psi_{i,j,h}$, where the j index represents the sector which affects the demand for the product of sector i) applied to the computed indices of the group-specific labor incomes (denoted by ω_h and which is computed as $\omega_h = y_h / yo_h$, where yo_h is the benchmark level of y_h) and the model computed consumer prices (denoted by $p_{j,h}$, which may be interpreted as real prices):

  $q^{i}_{h} = qo^{i}_{h} \cdot {\omega_h}^{\gamma_{i,h}} \cdot \Pi_j \, ({p_{j,h}}^{\psi_{i,j,h}})$,

  where the Π_j operator represents that the following expressions have to be computed for each j sector and the results have to be multiplied together.

  Note, that the $\gamma_{i,h}$ and $\psi_{i,j,h}$ elasticities and the $p_{j,h}$ prices may not be different for the individual household groups. However, this formulation leaves room for such extremely social and distributional policy relevant model simulations which take into account such policies which influence the groups-specific/local prices (e.g. by group-specific indirect taxes and subsidies) so that – without rationing and in kind provisions/distributions – the poor still can consume a certain minimum amount of the basic goods (see the SDIs related to accessibility, deprivation, hunger, nutrition, etc.).

  As mentioned before, these income elasticities may change with welfare. Therefore, when applying these elasticities to the household groups in simulations for the 2014-2030 period we used the average income elasticities reported by the simulation results of the standard MAGNET model (i.e. without multiple households) for the 2014-2018 and the 2025-2030 periods.

  Also note, that since the (weighted) average income elasticities are approximately 1, and the price level (general price index) is also around 1 (see in Table 5.1. below) or at least unchanged (as usual for the Walrasian-type of CGE-models, which determine only the relative prices and relative incomes) the total consumption expenditure per labor income ratio remains approximately unchanged (this can also be observed in Table 5.1.).

3 The European Household Budget Surveys

Since for arbitrary and complex household groups usually no statistical tables are available, to compute the average value of each economic variables (wealth, income, consumption, etc.) for the above defined (individual) household groups one must have access to the micro data of the available household income and expenditure surveys. Such surveys are available for the EU countries. These surveys are presented in the following sections.

3.1 The Eurostat microdata

The Luxembourg Income Study (LIS)[1] is a data archive and research center dedicated to cross-national analysis and is home to two databases:

- The **Luxembourg Income Study Database** (LIS) is the largest available income database of harmonised microdata collected from about 50 countries in Europe, North America, Latin America, Africa, Asia, and Australasia spanning five decades.

- The **Luxembourg Wealth Study Database** (LWS), is the first cross-national wealth micro database in existence.

The authors of the LIS wish to enable, facilitate, promote, and conduct cross-national comparative research on socio-economic outcomes and on the institutional factors that shape those outcomes. The website contains a collection of working papers based on the above databases[2]. However, these papers deal almost exclusively with the effects of social policy and taxation, while no trade policy impact studies can be found.

Household Finance and Consumption Survey:

The other international micro database which contains detailed data on the wealth and incomes of households is the 'Euro System **Household Finance and Consumption Survey**' (HFCS) of the European Central Bank (ECB)[3]. The survey is based on 84,000 interviews conducted in 18 EUR area countries, as well as Poland and Hungary, mainly in 2013, 2014 and 2017. The first wave of the HFCS was conducted mainly in 2010 and was published in 2013. The HFCS questionnaire consists of two main parts:

- questions relating to the household as a whole, including questions on real assets and their financing, other liabilities and credit constraints, private businesses, financial assets, intergenerational transfers and gifts, consumption and saving.

- questions relating to individual household members, covering demographics (for all household members), employment, future pension entitlements and income (for household members aged 16 and over).

[1] LIS Cross-National Data Center in Luxembourg (lisdatacenter.org)
[2] http://www.lisdatacenter.org/lis-wp-webapp/app/search-workingpapers
[3] https://www.ecb.europa.eu/stats/ecb_surveys/hfcs/html/index.en.html

Note, that data on consumption is available only in a rather aggregated form. Concretely, the survey has the following consumption related variables:

- HB2300 monthly amount paid as rent
- HI0210 amount spent on utilities
- DOCOGOOD Amount spent on consumer goods and services
- DOFOODC Food expenditure in a month at home/outside home
- DOFOODCH Food expenditure in a month (at home)
- DOGIFTINHER Amount of received gifts and inheritances
- DOTRIPSH Annual expenses on trips and holidays

Obviously, such incomplete set of the consumption expenditures is insufficient to conduct a CGE-model analysis of the household groups. Note, however, that the survey contains a quite valuable variable, i.e. the current or past *industry affiliation* of each adult person in the surveyed/interviewed household. Concretely, these variables are the following:

- PE0400 main employment – NACE code
- PE0450 previous employment - NACE for unemployed
- PE0470 main employment - NACE for retired or other inactive

The HFCS website also contains a list of research papers[4] based on their data, however, apparently none of them are related to trade policies.

EU Statistics on Income and Living Conditions:

The third household micro database is the **EU Statistics on Income and Living Conditions** (EU-SILC)[5], which aims to collect timely and comparable cross-sectional and longitudinal data on income, poverty, social exclusion and living conditions.

The EU-SILC project was launched in 2003 based on a 'gentlemen's agreement' between 6 Member States (Belgium, Denmark, Greece, Ireland, Luxembourg and Austria) and Norway. The legal basis entered into force in 2004 and since then the country coverage was extended and now covers all EU countries plus Iceland, Norway, Turkey, Switzerland, North Macedonia, Montenegro, Serbia, Albania, and Kosovo.
s
The EU-SILC provides two types of data:

- Cross-sectional data over a given time or a certain time period with variables on income, poverty, social exclusion and other living conditions.

- Longitudinal data on individual-level changes over time, observed periodically over a 4-year period.

[4] https://www.ecb.europa.eu/pub/economic-research/research-networks/html/researcher_hfcn.en.html
[5] European Union Statistics on Income and Living Conditions - Access to microdata - Eurostat (europa.eu)

Note, that although the EU-SILC database also records the NACE (Rev.2) 2 digit-level code (variable code: PL111) of the industry affiliation of the household members, this seems to be 'anonymized' in the datasets made available for scientific research (see the P-file worksheet of the SILC Variables - from 2014.xls file)[6].

Information on social exclusion and housing conditions is collected mainly at household level, while information on labor, education and health is obtained from individuals aged 16 and over. Income variables at detailed component level are also mainly collected from individuals.

The last but apparently most useful household micro dataset is the Eurostat's collection of **Household Budget Surveys** (HBSs). Since we found it to be the most important dataset for our research project and therefore, we requested this from the Eurostat, in the next section it will be discussed separately.

There are various databases which contain the **Gini-coefficients** of income inequality. The best (most comprehensive and updated and freely downloadable) such dataset is the **Standardized World Income Inequality Database** (SWIID). The goal of the SWIID is to meet the needs of those engaged in broadly cross-national research by maximizing the comparability of income inequality data while maintaining the widest possible coverage across countries and over time. It incorporates data from the OECD Income Distribution Database, the Socio-Economic Database for Latin America and the Caribbean generated by CEDLAS and the World Bank, Eurostat, the World Bank's PovcalNet, the UN Economic Commission for Latin America and the Caribbean, national statistical offices around the world, and academic studies while minimizing reliance on problematic assumptions by using as much information as possible from proximate years within the same country. The data collected by the Luxembourg Income Study is employed as the standard. The SWIID currently incorporates comparable Gini indices of disposable and market income inequality for 198 countries for as many years as possible from 1960 to the present; it also includes information on absolute and relative redistribution. A full description of the SWIID, the procedure used to generate it, and an assessment of the SWIID's performance in comparison to the available alternatives is presented in Solt (2020). To use the SWIID in statistical analyses, datasets formatted for use in Stata and R are available for download.

Note, that these Gini-indicators were used for adjusting/correcting obviously underreported incomes of the socioeconomic groups defined in the models of our modelling toolbox. This will be reported in the subsequent sections and chapters of this report.

3.2 The Eurostat collection of Household Budget Surveys

The HBSs are conducted in all EU Member States and are focusing mainly on household expenditure on goods and services. Since the survey is conducted based on a gentlemen's agreement, each Member State decides the objectives, methodology and frequency of the survey. Although there have been continuous efforts to make the data comparable across countries and over time, differences remain. The surveys vary between countries in terms of

[6] {GroupName}} – Könyvtár (europa.eu)

frequency, timing, content or structure. Consumption is reported in a 2-3 digits level COICOP-classification.

Eurostat has been collating and publishing these data every five years since 1988. The last two collection rounds were 2015 and 2020.

In the Household Budget Surveys section of the Eurostat Database one can find various data tables about the HBS returns/results. However, in these tables the data are presented only for some one-dimensional socioeconomic groups like the income quintiles, the age-groups or activity status, but not for any multidimensional groups.

Concretely the following table shows the available break-downs of the household sector:

- Mean consumption expenditure by socio-economic category of the reference person (hbs_exp_t131)
- Mean consumption expenditure by number of active persons (hbs_exp_t132)
- Mean consumption expenditure by income quintile (hbs_exp_t133)
- Mean consumption expenditure by type of household (hbs_exp_t134)
- Mean consumption expenditure by age of the reference person (hbs_exp_t135)
- Mean consumption expenditure by degree of urbanisation (hbs_exp_t136)
- Mean consumption expenditure by main source of the household's income (hbs_exp_t137)

Note, that in the above list the 'socio-economic category' just means the activity status where the employees ('workers') are split to groups of manual and non-manual workers. More disappointingly, for this split only the 'Purchasing power standard (PPS) per adult equivalent' and the 'Purchasing power standard (PPS) per household' indicators are reported. Therefore, to conduct a reasonable analysis one must get access to the micro datasets of the HBS.

The microdata is available for reference years: 2010 and 2015 (https://ec.europa.eu/eurostat/web/microdata/household-budget-survey). To protect the anonymity of respondents (persons, households), the access to microdata is restricted and for researchers only a partially anonymized version, i.e. the so-called scientific-use datafiles are made available.

For the survey reference year 2015 the HBS scientific use files contain 2 files (one containing the household related records and another one containing the personal records) separately for each of the following 23 EU member states in Excel format:

BE, BG, CZ, DK, DE, EE, FR, IE, IT, EL, ES, HR, HU, CY, LV, LT, LU, PL, PT, RO, SK, SE, FI.

For, example, for Hungary the following two files could be downloaded:

- **HU_MFR_hh.xlsx**: containing records (rows) of 7185 households

- **HU_MFR_hh.xlsx**: containing records (rows) of 16785 persons belonging to the above households.

However, Eurostat grants access to these micro-data only to *recognised research entities* and only through a rather bureaucratic process. In April 2022 our research group managed to submit a request for the 2015 EU HBS data. Eventually we got the data at the beginning of July 2022. Along with the data files the most recent user manual was also provided (HBS_UserManual_2015_ver1_3.pdf file).

More details of the acquisition process can be found in our earlier paper (Révész, 2022c). In the next section we summarize how we actually used the Eurostat HBS datasets.

3.3 Processing the Eurostat Household Budget Survey data

The General Algebraic Modelling System (GAMS, see on https://www.gams.com/) software and programming language was chosen to write a program code for reading in the HBS and auxiliary data from the Excel-files (through its GDXXRW facility) and processing them. The GAMS program developed to process the Eurostat HBS data is called HBSEU.GMS.

However, the GAMS have some strict requirements about the format of the data of the Excel-file to be read in. Therefore, our first task was to modify the acquired scientific-use datafiles accordingly. Details of these modifications are also described in our aforementioned paper (Révész, 2022c).

The data processing revealed several characteristics which limited the scope or depth of possible policy analyses. For example, for some unknown reasons the *household expenditures in abroad* (in the 12 main COICOP-categories break-down and even in aggregate form) contained only zeros. One may wonder how this could be dropped from the total consumption expenditures of the resident households without distorting their consumption level.

Since the NUTS-2 regional codes were suppressed (apparently due to the many bad data) only the 'major socio-economic regions', i.e., the NUTS-1 codes are available in the sample. It is a really disappointing fact, since it prevents us to analyse the socioeconomic impact of trade policies and trade shocks on the most affected (e.g., less developed agricultural) regions and to elaborate related policy recommendations, although 'Regions eligible for support from cohesion policy have been defined at NUTS 2 level'.

Although in the file the ME04 variable shows the *economic sector of employment* of the household members, this was available only in the ISIC one-letter codes break-down, i.e., only for 20 branches (see in the Eurostat provided HBS User Manual[7]). Since agriculture (coded A) and Manufacturing industry (coded C) are only 2 of these 20 branches this posed a difficult task and required various assumptions and auxiliary information to transform (disaggregate, etc.) the related employment and labor cost data to the GTAP sector classification (which break-down is needed by our CGE-models).

Strangely, the Eurostat provided HBS datasets do not contain income categories similar to those of the national accounts. In particular labor incomes are not shown separately. The given income categories are the following:

- EUR_HH012 Income in kind from employment (wages and salaries in kind)
- EUR_HH023 Income in kind from non-salaried activities

[7] fb5d8371-08fe-4ecf-bca6-b40984fde0b6 (europa.eu)

- EUR_HH032 Imputed rent
- EUR_HH095 Monetary net income (total monetary income minus income taxes)
- EUR_HH099 Net income (HH099 = HH095 + HH012 + HH023 + HH032)

Regarding the ‘monetary net income’ the HBS User Manual remarks the following:

‘When source of income does not concern any individual, but the household as a whole, it is allocated to the household's record. This is why property income, income-in-kind (except those from paid employment) and housing benefits are not included in the individual's record. Consequently, the sum of the individual incomes does not necessarily equal to the household's income.’

The personal record of the HBS contains only the following income category:

- EUR_MF099 Total income from all sources (net amount) corresponding to each single member of the family

For checking the reliability of the HBS data (weighted up to the whole population) we acquired the GTAP10.1 database (dated from January 13, 2021) and the data for the household consumption expenditures in the so-called COICOP-classification (available in the Eurostat database in 2-digits code break-down). Since the Eurostat HBS consumption categories are also in COICOP break-down (although in certain cases even 5-6 digits code detailed) the *representativity of the HBS data* could be checked in the common 2-digits code break-down by comparing them to the corresponding figures of the consumption statistics.

Since so far, we have processed the 2015 Eurostat HBS data only for Hungary, Bulgarian and Portugal, the following Table 3.1. presents the representativity of these samples. Bulgaria was selected as supposedly the closest to Middle-Eastern and Asian developing countries (in welfare level, importance of agriculture, social structure, hidden economy), while Portugal was selected to serve as a proxy for Latin American countries and former Portugal colonies (in Africa). These proxies will be used to estimate/impute missing data of the household surveys of the selected developing countries (for example, see in Chapter 4 how Portuguese consumption patterns were used to estimate the Brazilian ones). Finally, Hungary was selected mainly because the main contributor to this report and the scientific leaders of the project are connoisseurs of the Hungarian household budgets and in general the Hungarian statistical system and this greatly facilitated the checking of the Eurostat HBS data.

Note, that although in average the Portuguese HBS data seem to be somewhat more representative than the Hungarian HBS data, in the case of the expenditures on food – which is more relevant for our research – the opposite is true. In addition, extreme (5-fold) overrepresentation of the ‘maintenance and repair of the dwelling’ in the Portuguese HBS might be due to the improper inclusion of certain investment expenditures while the corresponding national account data contains only the consumption goods.

Further analysis of the table would reveal similar methodological issues, which one may bear in mind when using HBS data for Ghana or other developing countries too. However, since the focus of our past research and this report is not the analysis of the European household groups, this we leave for the interested reader.

Table 3.1. Representativity of the Hungarian, Bulgarian and Portuguese HBS consumption data for 2015 (in M €)

Consumption categories		**Hungary**			**Bulgaria**			**Portugal**		
Category name	**COICOP code**	weighted totals of **HBS** consumption data	**National Accounts** data for domestic consumption	HBS/ Nat.Acc. **ratio**	weighted totals of **HBS** consumption data	**National Accounts** data for domestic consumption	HBS/ Nat.Acc. **ratio**	weighted totals of **HBS** consumption data	**National Accounts** data for domestic consumption	HBS/ Nat.Acc. **ratio**
Food	CP011	6804.2	8751.3	***0.78***	4360.8	5351.7	***0.81***	11171.7	19821.6	***0.56***
Non-alcoholic beverages	CP012	671.5	1448.8	***0.46***	247.0	470.1	***0.53***	789.5	1169.6	***0.67***
Alcoholic beverages	CP021	405.3	1694.1	***0.24***	198.8	482.9	***0.41***	569.8	1728.7	***0.33***
Tobacco	CP022	598.9	2049.5	***0.29***	529.0	990.3	***0.53***	735.8	1923.8	***0.38***
Narcotics	CP023	3.4	428.0	***0.01***	0.0	80.1	***0.00***	0.0	112.6	***0.00***
Clothing	CP031	867.1	1487.3	***0.58***	386.8	702.4	***0.55***	2103.7	5914.0	***0.36***
Footwear	CP032	397.4	562.7	***0.71***	212.7	224.7	***0.95***	795.2	2007.0	***0.40***
Actual rentals for housing	CP041	712.0	511.0	***1.39***	106.7	327.7	***0.33***	2132.6	3809.8	***0.56***
Imputed rentals for housing	CP042	7384.9	6920.5	***1.07***	3363.6	3044.3	***1.10***	16758.5	13358.8	***1.25***
Maintenance and repair of the dwelling	CP043	313.3	104.5	***3.00***	215.9	293.7	***0.74***	887.7	174.9	***5.08***
Water supply and miscellaneous services relating to the dwelling	CP044	1606.9	989.9	***1.62***	351.0	490.6	***0.72***	1659.6	1213.1	***1.37***
Electricity, gas and other fuels	CP045	3918.9	2887.2	***1.36***	1713.3	1797.7	***0.95***	5247.7	4384.5	***1.20***
Furniture and furnishings, carpets and other floor coverings	CP051	134.8	554.0	***0.24***	143.0	379.2	***0.38***	466.5	1593.0	***0.29***
Household textiles	CP052	57.9	282.5	***0.20***	34.9	50.4	***0.69***	178.6	583.3	***0.31***
Household appliances	CP053	204.6	554.3	***0.37***	137.2	258.9	***0.53***	460.8	1026.4	***0.45***
Glassware, tableware and household utensils	CP054	81.1	328.8	***0.25***	36.2	255.1	***0.14***	67.1	443.7	***0.15***
Tools and equipment for house and garden	CP055	84.8	175.5	***0.48***	32.0	265.8	***0.12***	100.3	267.2	***0.38***
Goods and services for routine household maintenance	CP056	654.1	619.4	***1.06***	242.4	228.9	***1.06***	2049.2	2148.7	***0.95***
Medical products, appliances and equipment	CP061	1187.3	1254.8	***0.95***	689.6	1318.9	***0.52***	2701.5	1915.0	***1.41***

Out-patient services	CP062	294.3	978.1	***0.30***	119.9	209.2	***0.57***	1778.1	3898.0	***0.46***
Hospital services	CP063	60.1	214.2	***0.28***	63.1	299.5	***0.21***	143.0	315.5	***0.45***
Purchase of vehicles	CP071	534.4	1538.7	***0.35***	74.6	526.1	***0.14***	3100.8	5101.2	***0.61***
Operation of personal transport equipment	CP072	2405.6	4670.6	***0.52***	893.3	2503.5	***0.36***	7658.8	7810.7	***0.98***
Transport services	CP073	467.8	1020.8	***0.46***	215.2	1490.5	***0.14***	993.6	2204.7	***0.45***
Postal services	CP081	6.0	32.8	***0.18***	2.3	152.6	***0.02***	11.2	127.4	***0.09***
Telephone and telefax equipment	CP082	61.7	72.1	***0.86***	21.6	181.2	***0.12***	146.1	100.0	***1.46***
Telephone and telefax services	CP083	2121.5	2043.8	***1.04***	698.8	1090.3	***0.64***	2550.7	2706.6	***0.94***
Audio-visual, photographic and information processing equipment	CP091	185.1	494.2	***0.37***	73.3	487.9	***0.15***	359.4	905.7	***0.40***
Other major durables for recreation and culture	CP092	2.5	4.0	***0.62***	9.4	82.0	***0.11***	18.6	92.3	***0.20***
Other recreational items and equipment, gardens and pets	CP093	425.6	966.0	***0.44***	101.2	266.8	***0.38***	808.2	1543.8	***0.52***
Recreational and cultural services	CP094	619.8	1746.8	***0.35***	163.8	872.0	***0.19***	948.3	2977.1	***0.32***
Newspapers, books and stationery	CP095	394.3	525.0	***0.75***	258.1	203.2	***1.27***	858.1	1259.9	***0.68***
Package holidays	CP096	353.8	284.1	***1.25***	103.1	354.7	***0.29***	330.5	403.8	***0.82***
Education	CP10	279.7	965.1	***0.29***	80.8	298.1	***0.27***	1882.5	1880.7	***1.00***
Catering services	CP111	978.6	3659.3	***0.27***	661.6	1520.4	***0.44***	6745.1	10550.6	***0.64***
Accommodation services	CP112	130.2	869.2	***0.15***	75.3	600.6	***0.13***	586.9	3752.8	***0.16***
Personal care	CP121	950.2	1029.1	***0.92***	417.7	349.5	***1.20***	2382.8	2629.2	***0.91***
Prostitution; other services n.e.c.	CP122_127	153.4		***0.21***	68.8		***0.13***	380.4		***0.16***
Prostitution	*CP122*		548.7	***0.00***		77.3	***0.00***		644.0	***0.00***
Personal effects n.e.c.	CP123	52.2	248.8	***0.21***	50.8	187.7	***0.27***	304.5	1373.0	***0.22***
Social protection	CP124	101.1	341.7	***0.30***	55.7	56.8	***0.98***	856.1	1567.7	***0.55***
Insurance	CP125	838.2	614.3	***1.36***	86.5	213.3	***0.41***	1680.3	2576.5	***0.65***
Financial services n.e.c.	CP126	119.1	1970.1	***0.06***	1.2	441.9	***0.00***	31.8	2621.4	***0.01***
Other services n.e.c.	*CP127*		169.7	***0.00***		470.7	***0.00***		1691.2	***0.00***
Total	**TOTAL**	**37623.7**	**56611.3**	**0.66**	**17296.8**	**29949.2**	**0.58**	**83431.3**	**122329.5**	**0.68**

Source: Authors calculations based on Eurostat data (EO2:EV49 range of the HU sheet of the HBSEUout_HU.xlsx, HBSEUout_BG.xlsx and HBSEUout_PT.xlsx files)

Note, however, low apparent representativity of certain expenditures may be due not only to underreporting, but also due to (sampling) weights of the HBS observations which are inappropriate for the given consumption category, to the consumption of foreign tourists (which by construction is not part of the HBSs but part of the domestic consumption data of the consumption statistics and the private consumption expenditure category of the GTAP10.1 database) and to imprecise mapping of COICOP categories with GTAP sectors.

However, our goal was the disaggregation of the GTAP10.1 data to various household groups and hence it was more important to create a *mapping between the COICOP and the sectors of the GTAP10.1 database* to see which HBS consumption category corresponds to which sectors of the GTAP10.1 database. The sector classification of the GTAP10.1 database is called GSEC3 and contains 65 sectors (see Aguiar et al., 2019 and Appendix 1).

After having studied earlier attempts to work out correspondence between household budget survey categories available in COICOP break-down and GTAP sectors (also via the CPC, see Sahin and van der Mensbrugghe (2007), Cazcarro et al. (2020) and Luu et al. (2020)) and the content of the GTAP sectors carefully, we elaborated such a mapping by using the report of the APRAISE project (EPU-NTUA, 2013), the Appendix of which contains the correspondence of the GTAP sectors with the ISIC/NACE 2 industry codes.

Details of the mapping process can be found in Révész (2022c). The created detailed matchings/mapping is stored in the MapHBSGTAP sheet of the AuxilData.xlsx file (i.e. the other Excel input file of the HBSEU.GMS GAMS program where the auxiliary data are put together), while their explanations (written mostly in cell notes) are given in the C3:D68 range of the HU sheet of the HBSEUout.xlsx file (i.e., the Excel output file of the program).

The GAMS program then weights the HBS expenditure data by the sampling weight of the given households and transforms them into the (GSEC3) GTAP sector break-down using the above-described mapping.

Subsequently, by computing the EUR/USD cross-exchange rate, these weighted totals are *converted to million $*. These divided by the corresponding private consumption expenditure figures of the GTAP10.1 (value of the sum of the VDPA and VIPA matrices) show the 'quasi-representativity' of the (weighted) HBS sample. Since our HBS data refer to year 2015 while the GTAP10.1 data are for 2014, this ratio can only be regarded to be an approximate measure of the representativity of the given HBS expenditure data. Naturally, in some cases (for some GTAP sectors) this ratio is higher than 1, which means 'overrepresentation', while a ratio lower than 1 means that the related consumption expenditure is underrepresented (due to underreporting, inadequate weights or mapping, etc.). Also note, that the mentioned private consumption expenditure data of the GTAP10.1 database are at 'agents' prices', which is similar to the purchasers (consumer) prices, but where the trade margin is separated out from the expenditure on the products and accounted as (direct) purchase of the trade sector's services. As a consequence, the quasi-representativity figures of the Eurostat HBS data for the 'trd' sector are quite low.

Table 3.2. illustrates this with the Hungarian HBS data. In general, studying the possible causes of the table's extremely low or high HBS/GTAP ratios may reveal further methodological and/or estimation problems, which ought to be addressed before using the data in model simulations and drawing far-reaching conclusions. For example, the ratio of 80 for the non-ferrous metals (sector 'nfm') and the ratio of 0 for the iron and steel production (sector 'i-s')

might be due to the misallocation of certain metal products (to the 'nfm' sector instead of the 'i-s' sector or vice versa) in the HBS data processing or in the GTAP database. Similar problem can be detected in the case of the forestry products (sector 'frs') and the wood- and wood products (sector 'lum').

Table 3.2. Comparison of the transformed 2015 Hungarian HBS data and the GTAP10 data

Sector	HBS15 M €	GTAP10.1 M $	*HBS15 /GTAP10.1*	Sector	HBS15 M €	GTAP10.1 M $	*HBS15 /GTAP10.1*
pdr	0,00	0,18	*0,00*	**rpp**	151,25	76,90	*1,97*
wht	0,00	188,05	*0,00*	**nmm**	190,07	29,02	*6,55*
gro	0,00	49,55	*0,00*	**i_s**	0,00	1,92	*0,00*
v_f	896,42	440,70	*2,03*	**nfm**	22,47	0,28	*80,24*
osd	0,00	16,34	*0,00*	**fmp**	128,20	38,48	*3,33*
c_b	0,00	1,35	*0,00*	**ele**	164,75	108,25	*1,52*
pfb	0,00	0,59	*0,00*	**eeq**	215,22	61,78	*3,48*
ocr	146,65	103,17	*1,42*	**ome**	187,25	168,70	*1,11*
ctl	0,00	2,97	*0,00*	**mvh**	522,82	207,98	*2,51*
oap	320,64	160,34	*2,00*	**otn**	11,73	20,21	*0,58*
rmk	309,63	72,14	*4,29*	**omf**	477,54	423,48	*1,13*
wol	0,00	0,33	*0,00*	**ely**	1724,38	1685,36	*1,02*
frs	709,77	116,79	*6,08*	**gdt**	1441,38	378,12	*3,81*
fsh	40,95	20,11	*2,04*	**wtr**	426,56	298,61	*1,43*
coa	43,29	13,29	*3,26*	**cns**	136,89	99,79	*1,37*
oil	0,00	0,00		**trd**	233,79	8183,98	*0,03*
gas	0,00	90,51	*0,00*	**afs**	1108,81	5305,39	*0,21*
oxt	0,00	5,98	*0,00*	**otp**	460,56	1341,15	*0,34*
cmt	37,10	120,24	*0,31*	**wtp**	0,29	16,52	*0,02*
omt	2072,18	2049,68	*1,01*	**atp**	0,91	116,79	*0,01*
vol	250,18	151,66	*1,65*	**whs**	6,04	255,38	*0,02*
mil	760,08	783,75	*0,97*	**cmn**	2127,51	1641,40	*1,30*
pcr	40,88	8,23	*4,97*	**ofi**	119,09	2169,21	*0,05*
sgr	97,84	52,86	*1,85*	**ins**	838,20	1210,37	*0,69*
ofd	2139,55	798,69	*2,68*	**rsa**	1286,58	1825,52	*0,70*
b_t	1679,07	3669,90	*0,46*	**obs**	1279,76	487,71	*2,62*
tex	59,13	65,46	*0,90*	**ros**	1049,07	3769,80	*0,28*
wap	865,47	145,93	*5,93*	**osg**	30,96	541,42	*0,06*
lea	397,44	69,92	*5,68*	**edu**	279,73	1369,95	*0,20*
lum	0,00	17,84	*0,00*	**hht**	455,57	1858,88	*0,25*
ppp	254,47	100,67	*2,53*	**dwe**	7384,87	8210,19	*0,90*
p_c	1932,08	2315,11	*0,83*				
chm	1061,59	250,03	*4,25*	**Total**	**37623,7**	**53868,5**	***0,70***
bph	1047,03	83,55	*12,53*				

Source: Authors own calculations based on the Eurostat HBS data and the GTAP10.1. data (B290:J326 range of the HU sheet of the HBSEUout_HU.xlsx file); the explanation of the sector codes can be found in Appendix 1.

Therefore, the next step of the GAMS program was to estimate the trade margin component of the consumption expenditures. Similarly, household expenditures for unmatched or underrepresented GTAP sectors also had to be imputed, so that the related originally reported

HBS data were replaced by 'imputed' data estimated proportionately to an appropriately selected proxy expenditure category.
The next task of the GAMS program was to create household groups according to the following grouping (stratification) criteria:

- by the *major region* of their residence
- by the *age group* (band) of the main earner
- by the household per capita *expenditure terciles combined with the major region* affiliation (i.e., the elements of the set of groups are the Cartesian product of the regional and expenditure sets)
- by the *economic sector* of the main earner
- by the per capita *consumption deciles*
- by the per capita *income deciles*

In doing these groupings we faced various serious data-, technical- and/or methodological problems (see also in Révész, 2022c). Eventually the GAMS program computes the *group- and sector specific (weighted) incomes, expenditures, employment and populations*. Here we also faced various problems. Notably, to estimate the full-time equivalent employment we had to apply some reasonable assumptions, since the survey did not contain the 'hours worked' category or any details of the 'part time employment' category.

Unfortunately, the income categories of the Eurostat HBS not only can neither be matched with the national accounts (SNA) categories (as described in section 3.2.) nor with the GTAP data. Since in the GTAP data and model the private households are not separated out from the 'regional household' (which practically receives the whole primary income and represents the whole final demand) and the secondary income distribution is missing almost completely from the GTAP database (partly because not breaking-down the 'regional household' to its private households, state households, non-profit organisations and firms components, which renders the transfers between these component agents irrelevant in the accounting framework of the GTAP database) there are only a few categories in the GTAP data which might correspond to the household incomes. In fact, the GTAP database contains only the "Net factor costs by sectors and factors of production" (denoted by VFM) as such a category. From the point of view of the owners of the production factors, the 'net factor costs' (VFM) may be called as the gross factor incomes since it still includes (at least in principle) the employees' social security contributions (or 'workers' social security contribution' abbreviated subsequently as WSSC) and the personal income tax (P.I.T.). In any case, the VFM matrix contains the labor costs by types of labor, the capital costs and the costs of 'land' which presumably includes the land rent too.

Imputed rent income of the HBS (called EUR_HH032 as an income variable and EUR_HE042 as an expenditure category) corresponds to the gross capital income of the dwelling sector (coded by 'dwe') element of the VFM matrix. Therefore, the imputed rent income HBS data can be adjusted proportionately to this element of the VFM matrix.

Because of these different break-downs, definitions and representations of the transfer incomes in the Eurostat HBS data (and the HBSs of the developing countries) and the GTAP data we did not attempt to disaggregate our CGE-models household transfer variables to household groups.

More detailed discussion of the matter can be found in Révész (2023). However, as far as the imputed rent and the labor incomes are concerned, we have elaborated a method which distributes them to the individual households or household groups and transforms them into the 65 GSEC3 sectors break-down.

The so estimated GTAP-sector x household groups dimension (initial) labor income matrix had to be adjusted row-wise to the corresponding GTAP10.1 data (i.e. the total labor incomes by sectors computed from the VFM matrix). To retain the originally observed/reported share of the given household groups in the total personal income too (which meant that the matrix elements had to be adjusted column-wise too), the RAS biproportional (two-directional) matrix adjustment algorithm was used. The two-directional adjustment of the (matrix of the) so far estimated labor *incomes* was done by the GAMS program. For this *two-directional matrix adjustment* problem the GAMS program uses the *RAS algorithm*, concretely its simplified version (see the ‘Algorithm 2’ section at the ‘Iterative proportional fitting’ - Wikipedia website). To demonstrate our capability to form combined (two-criteria defined) household groups Table 3.3. shows the results for the region-consumption level terciles combined groups.

Table 3.3. RAS-estimated labor incomes of the Hungarian households by GTAP sectors and combined (NUTS-1 region, per capita consumption level) household groups

year 2014, values are in million US$

	1st major region			**2nd major region**			**3rd major region**			
	Low cons.	**Aver. cons.**	**High cons.**	**Low cons.**	**Aver. cons.**	**High cons.**	**Low cons.**	**Aver. cons.**	**High cons.**	**All Hou-seholds.**
codes	*Reg1 Poor*	*Reg1 Midd*	*Reg1 Rich*	*Reg2 Poor*	*Reg2 Midd*	*Reg2 Rich*	*Reg3 Poor*	*Reg3 Midd*	*Reg3 Rich*	**Total**
pdr	0.0	0.0	0.0	0.1	0.1	0.1	0.1	0.2	0.1	**0.5**
wht	7.7	5.9	3.9	26.1	30.8	29.8	49.6	79.1	29.3	**262.2**
gro	7.4	5.6	3.8	25.0	29.6	28.6	47.6	75.9	28.1	**251.6**
v_f	12.4	9.5	6.3	42.0	49.7	48.0	79.9	127.4	47.2	**422.4**
osd	3.8	2.9	1.9	12.8	15.2	14.7	24.4	38.9	14.4	**129.1**
c_b	0.2	0.2	0.1	0.8	1.0	1.0	1.6	2.5	0.9	**8.4**
pfb	0.3	0.3	0.2	1.2	1.4	1.3	2.2	3.5	1.3	**11.8**
ocr	1.0	0.8	0.5	3.4	4.0	3.9	6.5	10.3	3.8	**34.2**
ctl	0.9	0.7	0.5	3.2	3.8	3.6	6.0	9.6	3.6	**32.0**
oap	17.5	13.4	8.9	59.3	70.0	67.7	112.5	179.6	66.5	**595.3**
rmk	3.4	2.6	1.7	11.4	13.5	13.0	21.7	34.6	12.8	**114.7**
wol	0.4	0.3	0.2	1.2	1.5	1.4	2.4	3.8	1.4	**12.5**
frs	2.1	1.6	1.1	7.2	8.5	8.2	13.7	21.9	8.1	**72.5**
fsh	0.2	0.1	0.1	0.6	0.7	0.7	1.2	1.9	0.7	**6.2**
coa	0.0	0.3	1.3	0.9	2.5	2.6	0.7	1.9	1.4	**11.6**
oil	0.0	0.5	2.2	1.6	4.3	4.5	1.3	3.2	2.3	**19.8**
gas	0.0	0.4	1.9	1.4	3.7	3.9	1.1	2.8	2.0	**17.2**
oxt	0.0	1.0	4.5	3.4	8.8	9.2	2.6	6.6	4.8	**41.0**
cmt	1.0	1.3	1.8	1.6	2.4	2.1	2.4	2.1	1.2	**15.9**
omt	20.6	25.9	37.5	32.7	49.8	44.0	49.1	42.3	25.2	**327.1**
vol	2.2	2.7	3.9	3.4	5.2	4.6	5.2	4.4	2.6	**34.3**

mil	6.1	7.7	11.2	9.7	14.8	13.1	14.6	12.6	7.5	**97.4**
pcr	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	**0.9**
sgr	0.7	0.9	1.3	1.2	1.8	1.6	1.8	1.5	0.9	**11.7**
ofd	24.7	31.2	45.1	39.3	59.9	52.9	59.1	50.8	30.3	**393.3**
b_t	10.6	13.3	19.3	16.8	25.6	22.6	25.3	21.7	13.0	**168.1**
tex	8.3	10.4	15.1	13.2	20.0	17.7	19.8	17.0	10.2	**131.6**
wap	13.5	17.0	24.6	21.4	32.6	28.8	32.2	27.7	16.5	**214.4**
lea	5.6	7.0	10.1	8.8	13.5	11.9	13.3	11.4	6.8	**88.4**
lum	6.5	8.2	11.9	10.3	15.7	13.9	15.5	13.4	8.0	**103.4**
ppp	24.2	30.5	44.2	38.5	58.7	51.8	57.9	49.8	29.7	**385.5**
p_c	3.0	3.8	5.5	4.8	7.3	6.4	7.2	6.2	3.7	**47.6**
chm	18.2	22.9	33.1	28.9	44.0	38.9	43.4	37.3	22.3	**289.0**
bph	15.2	19.1	27.6	24.1	36.7	32.4	36.2	31.1	18.6	**241.0**
rpp	45.5	57.3	83.0	72.3	110.1	97.3	108.7	93.5	55.8	**723.5**
nmm	20.4	25.7	37.3	32.5	49.5	43.7	48.8	42.0	25.1	**325.0**
i_s	14.2	17.9	25.8	22.5	34.3	30.3	33.9	29.1	17.4	**225.4**
nfm	11.1	14.0	20.2	17.6	26.8	23.7	26.5	22.8	13.6	**176.3**
fmp	39.2	49.3	71.4	62.3	94.8	83.7	93.6	80.5	48.0	**622.8**
ele	47.8	60.2	87.2	76.0	115.7	102.2	114.2	98.2	58.6	**760.1**
eeq	55.9	70.5	102.1	89.0	135.4	119.7	133.7	115.0	68.7	**889.9**
ome	51.7	65.1	94.2	82.1	125.1	110.5	123.5	106.2	63.4	**821.7**
mvh	48.0	60.5	87.6	76.4	116.3	102.7	114.8	98.7	58.9	**764.1**
otn	7.8	9.8	14.2	12.3	18.8	16.6	18.6	16.0	9.5	**123.5**
omf	14.1	17.8	25.8	22.5	34.2	30.2	33.8	29.0	17.3	**224.8**
ely	0.0	41.5	27.0	20.2	58.7	84.6	20.3	31.9	34.9	**319.0**
gdt	2.5	5.5	4.8	6.7	4.7	4.5	9.0	5.2	3.9	**46.8**
wtr	19.9	44.4	39.2	54.2	38.2	36.6	73.0	41.9	31.4	**378.8**
cns	150.6	180.6	185.5	134.6	182.0	181.0	225.5	179.1	124.0	**1542.8**
trd	322.3	330.9	807.0	273.3	349.6	359.5	365.4	390.6	370.2	**3568.9**
afs	69.5	101.7	111.1	85.6	92.5	123.6	100.9	65.3	48.7	**799.0**
otp	101.3	186.4	164.9	83.3	133.4	170.1	127.6	121.3	80.0	**1168.3**
wtp	1.4	2.6	2.3	1.2	1.9	2.4	1.8	1.7	1.1	**16.4**
atp	2.3	4.3	3.8	1.9	3.1	3.9	2.9	2.8	1.8	**27.0**
whs	26.1	48.1	42.5	21.5	34.4	43.9	32.9	31.3	20.6	**301.2**
cmn	25.7	137.1	680.3	76.3	103.0	128.2	57.0	41.3	85.0	**1333.8**
ofi	33.4	143.8	417.3	34.6	134.6	274.8	53.1	125.2	134.5	**1351.4**
ins	5.7	24.6	71.3	5.9	23.0	47.0	9.1	21.4	23.0	**231.0**
rsa	26.8	81.2	118.9	0.0	61.0	21.7	25.3	16.9	75.7	**427.6**
obs	190.7	604.0	1661.3	140.1	307.0	323.3	240.6	271.4	229.8	**3968.1**
ros	74.6	86.5	300.1	86.5	93.3	121.8	133.7	118.9	76.9	**1092.4**
osg	139.5	295.6	481.0	245.4	305.5	332.4	506.8	383.7	329.1	**3019.1**
edu	117.7	267.2	477.2	195.5	269.4	339.8	259.8	475.6	428.1	**2830.3**
hht	158.4	312.0	658.6	262.2	465.2	564.4	454.5	541.6	491.0	**3907.8**
dwe	1.2	1.8	4.0	1.2	1.8	2.2	2.0	2.1	1.7	**17.9**
Total	**2043.1**	**3596.1**	**7238.6**	**2752.3**	**4160.2**	**4511.5**	**4309.5**	**4532.8**	**3453.2**	**36597.3**

Source: Authors calculation (B497:L566 range of the HU sheet of the HBSEUout_HU.xlsx file); the explanation of the sector codes can be found in Appendix 1.

Then, the total *private consumption expenditure* is computed from the GTAP data as the sum of the VDPA and VIPA categories. Then, by distributing this among the household groups proportionately to their HBS reported total consumption, the 'GTAP-data consistent' total consumption expenditure by household groups are estimated. To make the HBS data usable in the GTAP-based models, like the MAGNET model, the matrix of the so far estimated HBS-based household consumption expenditure data had to be adjusted both row-wise (i.e., each elements to the corresponding element of the sum of the VDPA and VIPA categories) and column-wise (i.e., to the corresponding element of the above estimated 'GTAP-data consistent' total consumption expenditure by household groups category).

Note, that these two-directional adjustments of the HBS data to the GTAP10.1 (and GTAP10.1 consistent group specific total consumption) data implicitly convert the euro values to dollars and 'backdate' (as opposed to 'update') them from 2015 to 2014.

Table 3.4. illustrates the results of this RAS-adjustment not by displaying the absolute figures but rather the estimated consumption structure or patterns of the Hungarian income deciles.

Table 3.4. RAS-estimates for the 2014 consumption patterns of the Hungarian households by GTAP sectors and per capita income decile groups

unit: per cent (%)

	1st decile	**2nd decile**	**3rd decile**	**4th decile**	**5th decile**	**6th decile**	**7th decile**	**8th decile**	**9th decile**	**10th decile**	**All H.holds.**
codes	**dec1**	**dec2**	**dec3**	**dec4**	**dec5**	**dec6**	**dec7**	**dec8**	**dec9**	**dec10**	**Average**
pdr	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	*0.00*
wht	0.33	0.27	0.35	0.25	0.36	0.28	0.27	0.33	0.23	0.23	*0.28*
gro	0.11	0.09	0.12	0.08	0.12	0.09	0.09	0.11	0.07	0.08	*0.09*
v_f	1.17	1.13	1.15	1.18	1.22	1.04	1.19	1.15	1.12	0.90	*1.10*
osd	0.03	0.04	0.03	0.03	0.03	0.03	0.03	0.03	0.03	0.03	*0.03*
c_b	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	*0.00*
pfb	0.03	0.00	0.02	0.02	0.02	0.01	0.00	0.00	0.02	0.00	*0.01*
ocr	0.10	0.16	0.18	0.17	0.23	0.22	0.28	0.25	0.34	0.22	*0.22*
ctl	0.00	0.00	0.00	0.01	0.00	0.00	0.01	0.01	0.00	0.00	*0.00*
oap	0.31	0.32	0.30	0.28	0.28	0.24	0.33	0.27	0.25	0.19	*0.27*
rmk	0.15	0.13	0.13	0.12	0.12	0.10	0.10	0.10	0.09	0.06	*0.10*
wol	0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	*0.00*
frs	0.39	0.28	0.25	0.22	0.21	0.19	0.18	0.15	0.11	0.06	*0.18*
fsh	0.05	0.04	0.04	0.03	0.04	0.05	0.04	0.05	0.05	0.04	*0.04*
coa	0.06	0.05	0.07	0.05	0.06	0.03	0.04	0.05	0.03	0.01	*0.04*
oil	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	*0.00*
gas	1.28	1.68	1.81	1.89	1.90	1.88	1.98	1.94	2.04	1.49	*1.79*
oxt	0.11	0.08	0.08	0.08	0.09	0.07	0.07	0.06	0.06	0.04	*0.07*
cmt	0.20	0.25	0.13	0.43	0.21	0.23	0.41	0.38	0.17	0.31	*0.28*
omt	6.24	5.15	4.84	4.67	4.46	4.10	4.10	3.73	3.63	2.54	*4.09*
vol	1.03	0.84	0.81	0.77	0.69	0.68	0.64	0.64	0.53	0.34	*0.65*
mil	1.97	1.84	1.83	1.86	1.78	1.76	1.71	1.78	1.64	1.38	*1.71*

pcr	0.11	0.09	0.09	0.07	0.06	0.06	0.06	0.06	0.05	0.04	*0.06*
sgr	0.58	0.49	0.45	0.45	0.46	0.34	0.37	0.26	0.27	0.13	*0.35*
ofd	3.33	2.72	2.68	2.55	2.49	2.31	2.30	2.23	2.18	1.61	*2.33*
b_t	9.07	8.50	7.79	6.93	5.89	6.47	6.46	6.26	5.83	5.21	*6.57*
tex	0.24	0.45	0.37	0.29	0.56	0.80	0.47	0.69	0.71	0.40	*0.51*
wap	1.57	1.77	1.84	1.82	1.95	1.65	1.73	1.56	1.71	2.08	*1.79*
lea	1.11	1.18	1.01	1.06	1.12	1.00	0.98	0.89	0.98	1.03	*1.03*
lum	0.11	0.08	0.07	0.06	0.06	0.05	0.05	0.04	0.03	0.02	*0.05*
ppp	0.28	0.34	0.35	0.39	0.32	0.33	0.29	0.31	0.37	0.29	*0.32*
p_c	2.42	3.10	3.64	3.46	3.88	3.50	3.66	3.14	3.71	3.62	*3.46*
chm	2.60	2.23	2.23	2.09	1.95	2.07	1.87	1.82	1.62	1.33	*1.89*
bph	1.08	1.22	1.41	1.50	1.37	1.54	1.55	1.62	1.50	0.97	*1.36*
rpp	0.42	0.50	0.52	0.54	0.49	0.55	0.50	0.59	0.48	0.44	*0.50*
nmm	0.08	0.14	0.15	0.16	0.12	0.16	0.28	0.13	0.18	0.13	*0.15*
i_s	0.00	0.00	0.00	0.00	0.00	0.01	0.01	0.00	0.00	0.00	*0.00*
nfm	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	*0.00*
fmp	0.18	0.24	0.22	0.20	0.29	0.35	0.33	0.36	0.37	0.32	*0.30*
ele	0.43	0.63	0.64	0.60	0.94	0.80	0.57	0.86	0.87	0.82	*0.74*
eeq	0.20	0.28	0.34	0.39	0.37	0.44	0.36	0.44	0.49	0.46	*0.39*
ome	0.50	0.98	0.80	0.81	1.05	0.84	1.17	1.34	0.80	0.83	*0.92*
mvh	1.30	1.01	3.66	3.76	3.35	4.36	1.74	4.59	2.10	6.43	*3.55*
otn	0.38	0.38	0.42	0.25	0.29	0.39	0.30	0.13	0.11	0.15	*0.26*
omf	0.80	1.01	1.27	0.93	1.18	1.10	1.08	1.12	1.06	1.26	*1.10*
ely	4.68	4.15	4.13	4.00	3.98	3.56	4.03	3.58	3.73	2.82	*3.74*
gdt	0.52	0.57	0.58	0.60	0.59	0.57	0.60	0.58	0.59	0.43	*0.55*
wtr	0.62	0.53	0.49	0.47	0.45	0.45	0.42	0.41	0.39	0.28	*0.43*
cns	0.06	0.12	0.07	0.22	0.10	0.15	0.19	0.15	0.17	0.15	*0.14*
trd	13.35	12.65	13.00	12.48	12.45	12.00	11.86	11.53	11.18	9.78	*11.75*
afs	8.00	7.18	6.28	6.23	6.00	7.90	8.34	6.24	8.03	9.84	*7.63*
otp	2.12	2.58	1.94	2.09	2.25	2.23	2.01	1.78	1.55	1.47	*1.93*
wtp	0.00	0.00	0.00	0.15	0.00	0.00	0.00	0.00	0.00	0.06	*0.02*
atp	0.07	0.09	0.05	0.12	0.14	0.46	0.66	1.97	0.93	1.76	*0.78*
whs	0.33	0.83	0.06	0.13	0.29	0.03	0.24	0.00	0.29	0.99	*0.37*
cmn	2.67	3.10	3.15	2.99	2.88	2.64	2.75	2.67	2.71	2.36	*2.75*
ofi	2.71	3.13	2.90	2.67	3.10	3.55	3.19	2.71	2.91	3.72	*3.12*
ins	1.08	1.06	1.49	1.52	1.90	1.78	1.79	2.14	2.25	1.82	*1.75*
rsa	4.16	3.44	2.58	2.75	2.31	2.16	2.24	2.16	2.59	2.66	*2.64*
obs	0.76	0.72	0.83	0.77	1.05	0.84	0.94	0.91	1.10	1.30	*0.97*
ros	3.73	4.23	4.05	4.66	5.57	5.72	5.17	5.30	6.91	7.50	*5.59*
osg	0.30	0.80	0.87	0.28	0.80	0.45	0.38	0.95	0.87	1.36	*0.78*
edu	2.91	2.63	2.33	3.12	2.57	1.84	1.60	2.09	1.32	0.95	*1.97*
hht	1.00	1.51	1.65	2.69	1.97	2.54	3.57	3.42	3.18	3.41	*2.67*
dwe	10.54	11.00	11.47	11.61	11.57	10.99	12.39	11.94	13.44	11.82	*11.78*
Total	**100,0**	**100,0**	**100,0**	**100,0**	**100,0**	**100,0**	**100,0**	**100,0**	**100,0**	**100,0**	***100,0***

Source: Authors own calculation (EA148:EL217 range of the HU sheet of the HBSEUout_HU.xlsx file); the explanation of the sector codes can be found in Appendix 1.

Table 3.4. illustrates how the consumption patterns depend on the per capita income. For example, the percentage share of food, utilities, land transport and education in the total consumption expenditures is higher in the upper income deciles, while the share of petrol (‘p_c’), car purchases (‘mvh’), and most non-material services (‘obs’, ‘ros’, ‘osg’ and ‘hht’) is lower than in the expenditures of the other deciles. Decreasing expenditure share of education with higher income looks odd but may be explained by the fact that many high-income households spend much on education abroad (which – as mentioned before – is not included in the Eurostat HBS data) or in the black economy as tutoring fee. This highlights the importance of studying the methodology behind the data and even supplementing/modifying them when needed and possible to avoid wrong conclusions from analysing/modelling them.

More detailed and more precise documentation of the above steps of the consumption matrix estimating process can be found in the referred paper (Révész, 2022c) and in the code of the HBSEU.GMS program in which for each step and almost each statement an explanatory comment is written.

We also processed the 2015 Eurostat Household Budget Survey (HBS) data for other countries. The first important finding is that even Eurostat cannot ensure the "consistency" of the HBSs of each country, i.e., their methodological uniformity and similar sample quality. To illustrate this, we use the HBS data of two EU countries with a similar size and population to Hungary (i.e., Bulgaria and Portugal) to dispel the possible opinion that the differences in the samples are explained by the size or development of the country. In the following paragraphs mainly those main features of the Bulgarian and Portuguese HBS samples are presented which are important from the point of view of multi-sector modelling and differ considerably from the Hungarian one.

The *Bulgarian HBS sample* is extremely small: it contains only 2,766 households, just over a third of the Hungarian one. As I mentioned, the variable of the person's income is missing (zero) from the Bulgarian personal data file. This makes it extremely difficult to estimate the sectoral origin of household incomes. The use of the data for regional analysis is further complicated by the fact that *there are only 2 ‘major regions*’ (NUTS-1 regions of the Eurostat) in Bulgaria.

Although *Portugal contains 3 large regions*, they have been defined in the most unusable way possible for regional analysts and modelers. Because *the whole of continental Portugal forms one large region*, while two small islands, Madeira and the Azores, form the remaining 2 large regions. To the credit of the Portuguese sample, it contains 11,398 households, one and a half times more than the Hungarian one.

The reliability and international comparability of the HBS data could be better understood based on the experience gained by processing the data of several other countries too. In any case, the modeler must be aware of these challenges and be ready to correct some data and data processing techniques and assumptions.

3.4 Main results of the Eurostat Household Budget Surveys-based analyses

Since changes in each household has an effect on the Gini index, it is understandably the most widely accepted measure of income or wealth inequality. Naturally, the value of the index also depends on what type of income or wealth is taken into account. The Wikipedia homepage (List of countries by income equality - Wikipedia) shows the different Gini indices for each country of the world as reported by the OECD, World Bank, United Nations, Eurostat, CIA, etc. organisations. In addition, we have acquired various Excel-files containing the time series for the World Bank (WB) and Eurostat (ES) Gini-indices. Then we compared them with the Gini-index computed from the corresponding country's HBS survey data.

For the Hungarian HBS the Gini-index turned out to be 0.199. This is significantly lower than the 2015 Hungarian Gini-index reported by the WB (0.304) or the ES (0.282). Similarly, for Bulgaria the original Gini-coefficient of the HBS sample was 0.214 while the ES reported Gini-index was 0.37. For Portugal the original Gini-coefficient of the HBS sample was 0.273 much closer to the official (ES reported) 0.34 Gini-coefficient. (For 2015 the WB reported the Portuguese to have been 0.355).

There is no available explanation for these differences. However, we may assume these are caused by various factors and features of the HBSs like the following:

- Incomes are more seriously underreported/underrepresented than in other, more income focused surveys,
- Since filling in the HBS questionnaire is more time consuming and requires more stability in lifestyle and skills, the less educated poor and busier rich are less likely to be included in the sample (Szabóné, 1996),
- By registering details of the consumption expenditures, reported incomes which seem to be inconsistent with the consumption habits are more likely to be checked and corrected

In any case, a model had to be developed by which the individual incomes are adjusted upwards disproportionately, so that the resulting Gini-index be closer to that of those of the WB and ES. The main methodological considerations of this adjustment were the following:

- A proxy variable (indicator) had to be found for the missing incomes
- The reported income itself is not the best candidate for such proxy variable: for some households zero or even negative income (entrepreneurial loss?) is reported, and omitting a part of the income – by definition – decreases the reported income (i.e., they cannot be assumed to change parallelly)
- Consumption may be a good proxy, but since it is not perfectly correlated to income, imputing even a proportionate part of the consumption to the incomes will not increase the Gini-index, but rather decreases it,
- Using a progressive (more than linear) function of the consumption as the proxy variable may require a strong progressivity (due to the mentioned less inequal consumption levels), which, however, may enlarge the possible errors in the reported consumption level

Based on the above considerations we experimented with various own-developed income-adjustment methods (see the earlier version in Révész, 2023). Eventually we found a satisfactory method which distributes the aggregate consumption discrepancy among the households. This revised method has the following two main steps:

1. econometric linear regression estimation of the **savings rate** of income deciles (as a function of the per capita consumption, which is a proxy for the less reliable income variable):

$s_i (c_i) = s^0_i + m_i \cdot c_i$, where s^0_i is the intercept and m_i is the *marginal propensity to save*

2. The aggregate ΔC consumption difference (National accounts – weighted HBS) is distributed among households proportionately to their **savings gap** (difference of econometrically estimated $s_i (c_i)$ and reported $y_i - c_i$ savings)

$a_i = y_i + \Delta C \cdot [s_i (c_i) - (y_i - c_i)] / q$, where $q = \sum_i [s_i (c_i) - (y_i - c_i)]$ is the total gap.

In step 1 to avoid the possibly distorting („noise") effect of weird/outlier individual data, the so-called kernel regression was applied for the data of the per capita consumption deciles (as data for 10 observations).

The results of the income adjustments were the following:

For **Hungary** the adjustment increased the income by 44 % on average (ranging from 26 to 76 % from the bottom to the top income decile). The Gini-index increased to 0.251 (so that the EUSILC figure target is 0.282).

Similarly, for **Bulgaria** the adjustment increased the income by 66 % on average (ranging from 28 to 125 % from the bottom to the top income decile) and the Gini-index increased to 0.299 (so that the target is 0.37).

For **Portugal** the adjustment increased the income by 41 % on average (ranging from 7 to 90 % from the bottom to the top income decile). The Gini-index increased to 0.366 (while the target range based on the WB and EUSILC official figures are 0.34 – 0.355).

We finish the presentation of our Eurostat HBS-related work with a note, that HBS data-based disaggregation of the household sector of certain EU-countries may be important also for those trade and agricultural policies which affect both some developing countries and their European trade partners.

4 Household Budget Surveys for Ghana

The method of estimating the consumption and labor income (‘wage’-) matrices was extended to selected developing countries. Ghana was already chosen as a target country in the project with partners in Ghana involved. In the following steps, the work for Ghana is presented in more detail. Additionally, we will include Brazil and India as the – in terms of population and/or agricultural output – largest Latin American and Asian developing countries, where the structural changes in the household sector (in particular, change in the income distribution across social groups, change in the consumer preferences and patterns) may have significant feedback to the world economy. However, the available survey data for these countries were rather different in scope, data quality and methodological clarity.

In Brazil only income survey (Pesquisa Nacional por Amostra de Domicílios - PNAD) was conducted for 2014. The data and documentation of this survey can be found on the following website:
https://www.ibge.gov.br/en/statistics/social/labor/20620-summary-of-indicators-pnad2.html

The Gini index of the sample turned out to be 0.539. This is not only satisfactorily close to the officially reported Gini-index for this survey (=0.526) but also close to the World Bank reported Gini-index of 0.521. Thus, no adjustment for the income (distribution) is needed.

The main steps of the processing of the 2014 Brazilian Income Survey (mostly in a GAMS program) included the estimate of the matrix of consumption expenditures (using as proxy the estimated Portuguese consumption matrix) and labor incomes (in which matrices the rows represent the GTAP-sectors and the columns the *per capita income deciles*) and their adjustment by the RAS-method to the corresponding GTAP10.1 data, while retaining the estimated/computed shares of the given household groups in the total consumption/labor income.

The closest Indian Household Budget Survey for our (i.e. the GTAP10’s) 2014 reference year is the Human Development Survey – II for 2011-12 (coded as IHDS-II). Its data (in 14 separate datasets) and methodological documentation can be downloaded from the https://www.icpsr.umich.edu/web/pages/DSDR/ihds-II-data-guide.html and the referred India Human Development Survey-II (IHDS-II), 2011-12 (umich.edu) website.

Interestingly, the Gini-index of the Indian survey sample (for the per capita incomes) turned out to be 0.5 while the World Bank reported figure is only 0.357. Difference may be due to the many negative household incomes (due to negative agricultural income, see our reference to Round (2003) in section 4.1.). Apart from this, there is no need for corrections of the incomes. Comparing with the Indian national accounts data, the representativity of the survey’s total consumption is 69.2 %.

In theory, the estimation of consumption and labor income matrices consistent with GTAP10.1 data and with the survey’s income/consumption shares required the same main steps what we mentioned in the case of Brazil. Nevertheless, the different content of the surveys required many different technical procedures.

The processing of the Brazilian and Indian data was made by the HBSEU_BR.GMS and HBSEU_IN.GMS GAMS-programs and their main results are extracted to the HBSBRout.xlsx and HBSINout.xlsx Excel-files respectively, while the rest are stored in files in GDX format. Many notes were written (even in the Excel-cells and the GAMS-codes) about the methodological and technical difficulties of the estimation process. The details of the whole process are described in the longer and more technical background paper.

4.1 Main features and processing of the Ghana Household Budget Survey for 2005/6

We got access to the Ghana Living Standard Survey fifth round (GLSS 5) which contains household and individual level data for part of 2005 and 2006. 8687 households and 37128 persons were interviewed. Only a general report was available (Ghana Statistical Service, 2008), the missing codebook was found in the World Bank microdata collection (see the Ghana - Living Standards Survey V 2005-2006 - World Bank SHIP Harmonized Dataset link which also duly publishes the main statistical indicators for each variable but without any critical remarks about the quality/reliability of the data).

Investigation of the received Excel files and their comparison with the World Bank documentation and the content of the above report revealed that our Excel-files are incomplete, the break-down of the food consumption (only 2 categories) and the industry of employment (10 branches) are less detailed than what is reported in the general report (10 categories and 17 branches). In the WB documentation a 4-digits industry code is also included, but ¾ of the respondents left it blank. We could not find any variables in the personal data sheet (concretely on the 2005_L sheet of the Ghana_GLSS5.xlsx file which we compiled from the original and apparently constructed variables of the survey data) which is related to the sector-affiliation of the given person. On household level (on the 2005_H or 2005_I sheet) we found such a variable (HHINDUSTRY) but it does not include all sectors displayed in Table 4.7. of the report (stored in the glss5_report.pdf file) (concretely "Fishing", "Hotels & Restaurants", "Real estates", "Education", "Health", "Private households" and the "Extraterritorial organizations" sectors of the table are missing from our data files). It is unclear from which variable Table 4.7. was computed from.

7997 households named the sector of the head of household. More than half of them, 4508 works primarily in the agriculture and fishing sector, 730 in the manufacturing sector and 1175 in the trade sector (called 'commerce' in the survey). All households have regional code and urbanization code. 8662 households reported educational attainment code.

Even assuming that the value of the own-production is not included in the incomes, total incomes and consumption seem to be inconsistent. Even the survey report states (page *vii – viii* and 105) that the average per capita household income is just 397 GH¢ (which is 433 $ converted by the 1 $ = 0.92 GH¢ exchange rate of June 2006) while the per capita household expenditure was 644 GH¢. Table 9.4. on page 96 shows that of this amount 284 GH¢ is the 'actual' food expenditure, while 50 GH¢ is the 'imputed' food expenditure. It is impossible to reconcile these figures with the other consumption related figures mentioned in the paper. For example, page 85 informs us that the value of the average per capita food consumption of own produce of food is 347 GH¢, while Table 9.11. claims that the per capita food consumption was

543 GH¢ (342 GH¢ ‘cash expenditure’ + 201 GH¢ ‘home produced’). In any case, it seems that the 397 GH¢ per capita total income does not include the value of own-production.

Data for cash- and in-kind wages also seem to be contradictory with each other and with total income. Note, that alone the average reported amount of the related World Bank documentation called ‘Bonus, social allowances derived from wage job’ (INC_SOCA variable, which might be some in-kind wages with some family support component) is higher in the sample (12.6 million cedis) then the total income (INC_TOT_G variable) the average of which is 11.8 million cedis. Since this is partly due to some really astronomical reported figures for INC_SOCA (the largest such recorded income is 5.475 billion ¢, which is about 600 thousand $!) in computing per capita income deciles we took into account only as much of it (by adding it to the reported total income) which did not exceed the reported total income.

More importantly, consumption expenditures are frequently so much misreported or -recorded that not only the expenditure on certain goods shows astronomical figures but in these households the share of these goods in the total consumption expenditure and the consumption/income ratio is unbelievably high. Concretely, there are 413 households where the consumption/income ratio is over 100 (!), for 1250 households this ratio is over 10, for 2127 households the ratio is over 5 and for 4236 households (almost half of the sample!) the ratio is over 2, which is still hardly believable. It is also hard to believe that reported/recorded consumption expenditure is lower than income only for 2499 households, less than 30 % of the sample. However, this might be due to the omission of the value of the consumption of the own-produced good from the agricultural and total income. This has been already discussed by Round (2003), who (see endnote 16 on page 181) claims that ‘household respondents have difficulty in separating out the intermediate costs of enterprises from consumer expenditures, so in the GLSS, as in LSMS surveys of other African countries, the majority of the estimates turned out to be negative.’

Note, that 40 households did not report any consumption (expenditure + of own produce), of which 25 did not report any incomes either. On the other hand, 10 households reported total consumption over 10^9, which is 100 thousand GH¢ (~110 thousand $). Since this alone cannot be taken for granted to be an error, we sorted the households by the share of the largest consumption expenditure item in their total consumption expenditure. We found that for 228 households this share is over 90 %.

Eventually of these 228 households we investigated the data of those 90 households (about 1 % of the sample) where this share was the highest. We found that the very high values were reported/recorded for the following consumption categories:

food bought, food purchased, non-alcoholic drinks, communication services, transportation, cloths, household textiles, telephone, newspapers, hotels and catering, household repairs and education.

Apart from the case of food we could not identify any patterns in these outliers. Then we had to inspect carefully the whole record of the related households, their size, occupation, sources of income, consumption patterns, etc. to find out what could be the cause of the error and what could be the true value of the misreported/-recorded figure. Depending on the estimated magnitude of the true value, we divided the original figure either by 1,000 or (more rarely) by 10,000. Dividing by 10,000 implicitly assumed that the cause of the error could be the misinterpretation of the reported figures in cedis (¢) as if they were already in GH¢ (introduced just during the processing of the survey data so that 1 GH¢ = 10000 ¢). Therefore, in the data processing these figures were multiplied by 10,000 meaning to convert them to cedis (which they should not have done since the figures were already in cedis). Dividing by 1,000 implicitly assumes that the given figure was misinterpreted as what was reported in millions instead of thousands, thousands instead of units or in billions instead of millions. This could result in erroneously converting them to the smaller unit of measurement by multiplying them by 1,000.

Eventually we corrected the largest expenditure item of 56 households. Details of these corrections (the corrected figures are highlighted by red or blue coloured fonts) can be found in Appendix 4 and in the GAMSINPsort sheet of the HBSwage_cons_corrected_simul.xlsx file. After these corrections the computed income- and consumption percentiles of the sample have become plausible. The inter-group Gini-index for the consumption deciles decreased from 0.677 to 0.412 (The World Bank reported that the Gini index for Ghana is 0.43) and the share of the top decile in the total consumption was approximately halved. However, the Gini-index for Ghana is quite high, hence analysing impacts on income groups must be quite important.

Processing of the data required more or less similar steps as we outlined in section 3.3. Computing the labor income transformation matrix from the above-mentioned quite aggregated 10 branches to the 65 GSEC3 sectors required a disaggregation. When one sector code covered more than one GTAP sectors the survey's employment and labor income figures were distributed among the related GTAP sectors according to the proportions of the aggregate labor income data of the GTAP database (computed from the VFM category). For the mentioned missing branches, it was assumed that they are included in the closest not-missing branch. For example, the missing health care services was assumed to be reported under the 'Education' branch and the 'Hotels & Restaurants' was reported under the not-missing 'Trade' branch.

Then, using the labor income transformation matrix the labor income matrix was computed and adjusted by the RAS-method to the corresponding GTAP data (computed from the VFM category) while retaining the income deciles shares in the total labor income.

Following this the consumption transformation matrix was elaborated to transform the consumption expenditures from the 47 consumption categories to the 65 GSEC3 sectors. This was a rather complicated process, so we describe its individual steps only in a concise way:

First, a French consumption transformation matrix (named as CPA_COICOP_FR in the HBSwage_cons_corrected_simul.xlsx Excel file) was borrowed from the set of European consumption transformation matrices estimated by Cai and van Dyck (2020) which transforms the consumption expenditures from COICOP consumption categories to the CPA product by activity break-down (practically the ISIC sector classification).

Second, the 2015 Austrian consumption transformation matrix was borrowed from the NACEtoGTAP matrix of the IOtoGTAP.gdx file created by Révész (2019). This transforms the consumption expenditures from the CPA classification to the 57 branches of the GSEC2 sector classification of the GTAP9 database. For this the CPA data of agriculture and food industry had to be split to their component GSEC2 sectors proportionately to their shares in the Ghanaian VDPA+VIPA consumption expenditure data in the GTAP database.

Finally, by matrix (post-)multiplying the NACEtoGTAP matrix by the CPA_COICOP_FR matrix we got a GSEC2 x COICOP dimension consumption transformation matrix. Since the 47 consumption categories of the GLSS5 do not correspond exactly to the COICOP classification and the content of the 65 GSEC3 sectors are partly different from the 57 GSEC2 sectors, further reclassifications were needed to create the desired GSEC3 x GLSS5 consumption categories dimension consumption transformation matrix (called Wants_GSEC3_GH in the HBSwage_cons_corrected_simul.xlsx Excel file and the HBSEU_GH.GMS program).

Then, as we did before for other countries survey data, we imputed the consumption of the trade margins (as consumption from the 'trd' sector) proportionately to the material products purchased and the 'imputed rent' (as consumption from the 'dwe' sector) to the table. Finally, after accomplishing the transformation of the survey data by this consumption transformation matrix (more precisely by the consumption transformation *coefficient* matrix computed from the absolute figures) the RAS-method was applied again to adjust the survey figures (i.e., the 'initial' consumption matrix) to the corresponding GTAP data while retaining the income deciles shares in the total consumption expenditures.

The RAS-adjustments performed quite well both for the labor incomes and the consumption expenditures. However, due to the many assumptions (disaggregations, imputations, proxies, etc.) and auxiliary matrices used in the process, further checks, refinements and improvements may be useful.

4.2 The Ghana Household Budget Survey for 2016/7 and the possibilities of its future use

However, the GLSS5 survey is quite old, therefore we investigated the possibility to use the more recent GLSS7 survey for 2016-17, which is closer to the benchmark year of the GTAP10 database and the MAGNET model. Comprehensive comparison of the GLSS5 and GLSS7 surveys would be quite interesting and useful, but here we can only highlight some of the strengths and weaknesses of these surveys and cannot go into details. Depending on the decision of the project leadership and the stakeholders, in the near future we may try to use the GLSS7 survey in our models or at least use certain parts of it to build into (merge with) the GLSS5 data.

The size of the GLSS7 sample is significantly larger than that of GLSS5. It contains 14009 households and 58844 persons. The data were made available to us by our Ghanaian partner Prof. Peter Quartney and his colleague in the Ghana Statistical Services, Ralph nii Amrah. A great improvement is that for the GLSS7 we also received the questionnaires and the manuals describing the meaning of the codes (which by the way turned out to be straightforward, since follows strictly the section code and serial number of the related question of the questionnaire) and the methodology. The official 'main report' was also provided for us (Ghana Statistical Services, 2019). Its Table 10.6. shows the consumption expenditures by the 12 main COICOP-categories, Table 10.9. the same by income quintiles, Table 10.10. the expenditure on 11 food categories and 34 non-food categories, Table 10.1.8 and Table 10.25. the main components of the labor and non-labour incomes and Table 10.26. the other (not consumption) expenditures like taxes, contributions, gifts, dowry, funeral, and other expenses.

We are considering using various software packages and methods for making the many and occasionally rather big files of the GLSS7 dataset usable efficiently. Apart from the various technical issues, the main challenge for us is that we have not found such components (files) of the dataset which contain the processed version of the data of annualized labor incomes and constructed variables for the food consumption by detailed food categories.

In addition, despite of Table 5.14. of the main report (which shows the employment by 21 branches so that for all, but the extraterritorial organization branch the figures are positive), in the methodological description we could not find any even such a character variable for the industry affiliation of the persons/households as it was the HHINDUSTRY variable in the GLSS5. Clearly, without such information it would not be possible to estimate the labor income matrix, at least not in a better way than we could do it with the GLSS5 data.

In the also available Ghana SAM 2005 - regional.xlsx file (SAM constructed by the IFPRI, see GSS – IFPRI, 2007) and the similar SAM compiled for 2015 (documented in the 2015 Ghana SAM Report.pdf file) the household accounts are disaggregated to regions and municipality types (urban - rural). In the future these data might be also used for computing social effects of certain simulation scenarios and be compared with our MAGNET-model based disaggregated results using the same groupings.

Instead, at least for the time being, we concentrated on feeding in the estimated data to our CGE-models and running various simulations with them. These are described in the following chapter.

5 Model simulations using the GLSS5 Household Budget Surveys for Ghana

Acquiring and processing the household budget survey data and other household related (macroeconomic, etc.) statistical data rendered possible the estimation of social and distributional effects of trade and agricultural policies by using the CGE-models of our modelling toolbox. In this section we intend to demonstrate this capability of our workgroup by running various simulations and evaluating their results. For this purpose, the MAGNET-model was chosen for its easier accessibility and more flexibility regarding the representation of the household sector. Social and distributional effects were computed for Ghana by splitting the household sector according to the 6 different break-downs (stratifications) which were described in the previous sections.

The MAGNET-model was run with the SSP1 scenario with GDP- and population growth forecasts. The SSP (Shared Socio-economic Pathways) scenarios describe future global socio-economic conditions, including associated emissions of greenhouse-gases (GHGs). The SSPs have been proposed as a new set of scenarios to be used as a basis of future climate research (Van Vuuren et al., 2014, O'Neill et al., 2017). The SSP1 is the scenario of green growth (sustainable development). The SSP database is published by IIASA (International Institute for Applied System Analysis) (see the description of the SSP scenarios at the https://tntcat.iiasa.ac.at/SspDb/dsd?Action=htmlpage&page=80 link and the details of their assumptions at the 1-s2.0-S095937801630067X-mmc1.docx (live.com) link). Two simulations were made: the first with the baseline scenario and the second with a liberalization scenario for the Ghanaian agri-food-trade. Of the simulation results the household sector related variables and parameters were extracted and used in computing the social and distributional effects.

5.1 Preparations of the data for the disaggregated model simulations

Since the given simulations of the MAGNET-model were run in 68 sectors break-down, our first task was to elaborate a *transformation matrix from the 65 GSEC3 sectors to these 68 sectors* (most of which has a one-to-one correspondence with the GSEC3 sectors, but some sectors are more aggregated and others represent subsectors separated out of GSEC3 sectors for more detailed energy-, agricultural- and environmental-policy analysis). The "AggrScheme" sheet of the HBSwage_cons_corrected_simul.xlsx file contains the provisional transformation coefficient matrix from the 65 GTAP sectors to these 68 sectors.

Secondly, since the MAGNET-model output displayed these 68 sectors in a different order than what their original (defined) order is a reordering was needed for those household group related categories (i.e., the matrices of consumption expenditures and labor incomes) and model parameters which were displayed in different (i.e., the original) sector order. For the reordering of the 68 x 68 size matrix of own- and cross-price elasticities even a permutation matrix was developed.

Since the MAGNET model's benchmark values for 2014 turned out to be different from the GTAP10.1 data to which the estimated survey labor income and consumption data had been adjusted (it turned out that while the MAGNET used the GTAP10 data dated 19/03/2020, the

HBS data processing used the GTAP10 data dated 11/01/2021) the processed HBS data had to be adjusted (further) to the MAGNET model's corresponding benchmark data for 2014.

5.2 The disaggregated model simulations

Then we ran both simulations in a disaggregated way, i.e. by using the model provided sectoral labor income and price indices and the related elasticity parameters.
To compare the results of the simulations with the benchmark and with each other we chose 2030 as the reference (terminal) year. 2030 was selected partly because it is the target year of sustainability policies and strategies and also because it is sufficiently far from the present (and from the benchmark) to see the longer run effects of the different parameter changes and tendencies but not so far from the present that the forecasts could not be compared with any informal but definite expectations of the future state of world.

Note, that the income- and price elasticities change over time in the MAGNET-model essentially depending on the level of the GDP at ppp (at purchaser's power parity, which represents the 'welfare') as it is suggested by the CDE-demand function used in the GTAP-models. As explained in section 2.6. for the disaggregated simulations for the 2014-2030 period, we used the average income elasticities reported by the simulation results of the MAGNET model for the 2014-2018 and the 2025-2030 periods.

Table 5.1. shows the elasticities used in the model and the main results for the model for the aggregate household sector and for the relevant economic sectors.

As one can see by comparing the first and last numerical columns of Table 5.1. that the MAGNET-model uses significantly different elasticities than what we estimated from the consumption matrix estimated for the per capita income deciles. In some cases, the estimates seem to be more realistic (although due to the above discussed lack of details in the consumption data of the GLSS5 the estimates could be made uniformly only for groups of GTAP/MAGNET-model sectors). In any case, it would deserve a more thorough consideration of the matter especially comparing these results with the income elasticities also estimated from the GLSS5 by Breisinger et al. (2008) (see in Appendix 3).

To understand the household group specific results, one has to know the figures of the 2030/2014 labor income indices of the model simulations. These are also displayed in separate columns of the table.

Even with the same income- and price elasticities the disaggregated calculations resulted in occasionally significantly different consumption levels and -patterns due to the different percentage changes in the labor income and different initial consumption patterns of the individual social groups. Change in the labor incomes of the social groups, in turn, depend on whether they are relatively more or less represented in those economic sectors in which the labor income increased the most or the least. By integrating this module formally into the MAGNET model, one can develop further this labor income determination so that the social groups labor supply in the individual sectors depend on the relative sectoral wage rates. This is usually done by introducing a Constant Elasticity of Transformation (CET-) function and by maximizing the wage earnings with given constraint on the total labor supply by groups. Obviously, such mechanism will have impact on the equilibrium wage rates by sectors.

Table 5.1. Selected households related features of the MAGNET-model's 'Baseline' and 'Trade liberalization' scenarios for Ghana

serial number and code	average income elasticity	own-price elasticity	Benchmark values (for 2014)		Baseline scenario					Trade liberalization scenario					Memo:
of the relevant MAGNET-model sectors	computed from the parameters of the MAGNET-model		Private consumption volume in M $	Labor income in M $	consumer price index 2030/ 2014	Private consumption volume in 2030, M $	Volume index of private consumption '30/'14	Labor income in 2030, M $	Value index of labor income '30/'14	consumer price index 2030/ 2014	Private consumption volume in 2030, M $	Volume index of private consumption '30/'14	Labor income in 2030, M $	Value index of labor income '30/'14	*income elasticities estimated from the GLSS5*
1 pdr	**-0.001**	**-0.168**	370.56	157.93	0.605	504.01	1.360	135.02	0.855	0.598	502.87	1.357	135.63	0.859	***0.253***
2 wht	**0.0065**	**-0.162**	129.64	1.30	1.004	163.02	1.257	1.69	1.293	0.860	166.49	1.284	2.20	1.689	***0.253***
3 grain	**0.0065**	**-0.187**	1089.75	689.78	0.610	1506.00	1.382	560.39	0.812	0.599	1504.06	1.380	553.57	0.803	***0.253***
4 oils	**0.0065**	**-0.167**	324.05	205.28	0.645	438.92	1.354	211.00	1.028	0.634	438.23	1.352	214.51	1.045	***0.253***
5 sug	**0.0065**	**-0.159**	0.01	0.24	0.659	0.01	1.400	0.30	1.233	0.647	0.01	1.400	0.30	1.250	***0.000***
6 hort	**0.153**	**-0.302**	5048.93	3276.17	0.654	8282.48	1.640	3160.02	0.965	0.640	8278.69	1.640	3165.81	0.966	***0.253***
7 crops	**0.0745**	**-0.159**	0.78	1209.15	0.884	1.05	1.343	1531.53	1.267	0.787	1.07	1.379	1578.81	1.306	***0.253***
8 oagr	**0.0745**	**-0.208**	35.31	58.60	0.655	51.39	1.455	68.42	1.168	0.647	51.26	1.452	69.24	1.182	***0.120***
9 cattle	**0.3745**	**-0.208**	3.32	5.02	0.636	6.13	1.850	3.86	0.770	0.611	6.16	1.858	3.52	0.702	***0.300***
10 othctl	**0.3745**	**-0.208**	7.76	11.81	0.634	14.37	1.851	8.93	0.757	0.609	14.43	1.859	8.05	0.682	***0.300***
11 pltry	**0.3745**	**-0.214**	215.06	61.04	0.629	400.17	1.861	59.72	0.978	0.621	399.35	1.857	58.52	0.959	***0.253***
12 wol	**0.3745**	**-0.207**	0.04	0.00	0.834	0.07	1.838	0.01	5.000	0.808	0.07	1.865	0.01	8.000	***0.000***
13 pigpls	**0.3745**	**-0.21**	55.64	15.53	0.640	103.27	1.856	16.36	1.053	0.630	103.01	1.851	16.02	1.031	***0.253***
14 milk	**0.153**	**-0.208**	3.15	2.83	0.585	5.02	1.595	1.76	0.620	0.547	5.06	1.609	1.53	0.538	***0.253***
15 bfmt	**0.3745**	**-0.214**	215.75	164.49	1.302	344.16	1.595	311.84	1.896	1.224	352.71	1.635	283.49	1.723	***1.136***
16 othcmt	**0.3745**	**-0.215**	222.48	148.53	1.205	360.59	1.621	224.68	1.513	1.057	379.32	1.705	173.44	1.168	***1.136***
17 pulmt	**0.3745**	**-0.213**	178.76	47.34	1.022	304.80	1.705	38.42	0.812	0.850	320.98	1.796	19.57	0.413	***1.136***
18 othmt	**0.3745**	**-0.213**	182.44	153.10	1.206	301.48	1.652	258.10	1.686	1.091	319.58	1.752	216.09	1.411	***1.136***
19 dairy	**0.153**	**-0.219**	311.59	74.15	1.000	440.82	1.415	76.81	1.036	0.866	458.01	1.470	57.77	0.779	***1.136***
20 sugar	**0.2335**	**-0.199**	176.30	0.01	1.047	265.84	1.508	0.06	9.143	0.952	269.64	1.529	0.09	12.857	***1.136***
22 vol	**0.63**	**-0.203**	305.54	119.11	1.110	619.37	2.027	139.16	1.168	0.913	656.65	2.149	96.01	0.806	***1.136***

23 pcr	**0.0065**	**-0.172**	473.58	9.86	0.956	600.41	1.268	13.22	1.341	0.811	616.08	1.301	9.84	0.998	***1.136***
24 ofd	**0.512**	**-0.289**	2931.79	666.89	0.970	5571.11	1.900	1341.36	2.011	0.905	5666.88	1.933	1286.57	1.929	***1.343***
25 feed	**0.346**	**-0.194**	0.00	2.22	0.862	0.00	2.000	7.42	3.349	0.862	0.00	2.000	7.58	3.419	***1.343***
26 wfish	**0.7925**	**-0.23**	731.73	305.90	1.023	1714.99	2.344	922.77	3.017	1.013	1712.04	2.340	909.43	2.973	***0.859***
27 aqcltr	**0.7925**	**-0.209**	41.71	5.91	1.222	94.16	2.257	11.96	2.025	1.258	93.28	2.236	11.71	1.982	***0.859***
28 fishp	**0.346**	**-0.199**	167.94	19.53	1.010	278.03	1.656	55.44	2.839	0.898	283.29	1.687	53.82	2.756	***0.859***
29 frs	**1.23**	**-0.232**	0.39	231.35	1.024	1.16	2.997	452.49	1.956	1.023	1.16	2.992	455.80	1.970	***0.195***
30 plan	**1.23**	**0**	0.00	0.06	0.000	0.00		0.13	2.276	0.000	0.00		0.14	2.328	***0.000***
32 petro	**0.846**	**-0.292**	838.80	1.31	1.870	1638.59	1.953	2.10	1.608	1.870	1632.50	1.946	2.07	1.589	***1.673***
38 gas	**0.975**	**-0.267**	1.67	0.00	1.311	3.91	2.340	0.00		1.311	3.90	2.332	0.00		***2.208***
39 coa	**0.975**	**-0.267**	0.50	0.00	2.202	1.03	2.048	0.00		2.202	1.03	2.042	0.00		***0.654***
40 ely	**0.975**	**-0.291**	662.05	171.61	0.982	1673.33	2.527	521.45	3.039	0.978	1669.80	2.522	520.20	3.031	***1.375***
42 ely_g	**0.975**	**0**		11.41				10.79					10.57		***0.000***
44 ely_h	**0.975**	**0**		82.14				154.53	1.881				154.63	1.882	***0.000***
51 othcrp	**1.0645**	**-0.311**	2162.23	77.57	1.047	5703.95	2.638	188.75	2.433	1.046	5692.56	2.633	191.02	2.463	***0.677***
52 fert	**1.0645**	**0**		6.93				14.33	2.066				14.02	2.021	***0.000***
53 f_chem	**1**	**0**		5.28				3.39	0.641				3.38	0.639	***0.000***
54 othind	**1.133**	**-0.378**	4039.56	4269.88	0.895	11638.04	2.881	12683.05	2.970	0.893	11623.27	2.877	12731.91	2.982	***0.976***
55 pel	**1.134**	**-0.232**	0.00	0.01	1.110	0.01	2.667	0.02	3.000	1.099	0.01	2.667	0.03	3.125	***0.000***
56 gas_dist	**1.913**	**-0.267**	0.18	0.02	1.071	0.90	5.096	0.05	2.368	1.071	0.89	5.051	0.05	2.368	***0.000***
57 trans	**1.913**	**-0.281**	545.27	2796.22	1.386	2640.93	4.843	6270.10	2.242	1.379	2622.67	4.810	6245.89	2.234	***2.373***
58 ser	**1.913**	**-0.462**	3029.05	7365.82	1.216	15668.30	5.173	26349.44	3.577	1.207	15607.47	5.153	26079.84	3.541	***1.191***
Total/ average	**1.045**	**-0.310**	**24503.29**	**22432.6**	**1.02**	**61341.77**	**2.50**	**55814.43**	**2.49**	**1.00**	**61454.47**	**2.51**	**55345.95**	**2.47**	**0.995**

Source: Authors calculations (see in the C35:R96 range of the ‘Content’ sheet of the HBSwage_cons_corrected_simul.xlsx file)

In any case, the main group-specific results of the model simulations with disaggregated household sector are summarized in Table 5.2.

Table 5.2. Estimated change in the labor income of differently defined social groups in the baseline and trade liberalization scenario

Name of the social groups	Code of the social groups	Labor income index 2030/2014		Name of the social groups	Code of the social groups	*Labor income index 2030/2014*	
		Baseline scenario	Trade liberali-zation scenario			Baseline simulation	Trade liberali-zation scenario
Average of all households:		**2.488**	**2.467**				
Deciles of the per capita consumption				**Deciles of the per capita income**			
1st decile	dec1	2.340	2.321	1st decile	dec1	*2.488*	*2.467*
2nd decile	dec2	2.029	2.017	2nd decile	dec2	1.882	2.269
3rd decile	dec3	2.241	2.227	3rd decile	dec3	1.683	2.095
4th decile	dec4	2.302	2.284	4th decile	dec4	1.743	2.198
5th decile	dec5	2.393	2.374	5th decile	dec5	1.988	2.322
6th decile	dec6	2.425	2.406	6th decile	dec6	2.092	2.310
7th decile	dec7	2.444	2.423	7th decile	dec7	2.326	2.353
8th decile	dec8	2.475	2.455	8th decile	dec8	2.485	2.456
9th decile	dec9	2.712	2.684	9th decile	dec9	2.668	2.504
10th decile	dec10	2.716	2.691	10th decile	dec10	2.549	2.628
Regions				**Industry affiliation groups**			
UpperWest	reg0	2.306	2.409	Agriculture, forestry and fishing	1	1.810	1.803
Ashanti	reg1	2.585	2.743	Mining and quarrying	2	3.004	2.980
BrongAhafo	reg2	2.225	2.435	Manufacturing	3	2.639	2.599
Central	reg3	2.510	2.610	Electricity, gas, steam and air conditioning supply	4	2.962	2.945
Eastern	reg4	2.344	2.499	Water supply; sewerage, waste management and remediation activities	5	1.000	*1.000*
Greater Accra	reg5	2.920	2.598	Construction	6	3.577	3.541
Northern	reg6	1.970	2.228	Wholesale and retail trade; repair of motor vehicles and motorcycles	7	3.364	3.330
UpperEast	reg7	1.825	2.254	Transportation and storage	8	2.295	2.285

Volta	reg8	2.342	2.442	Accommodation and food service activities	9	1.000	*1.000*
Western	reg9	2.175	2.422	Information and communication	10	1.000	*1.000*
Age-groups				Financial and insurance activities	11	3.105	3.076
Young	15_29	2.705	2.679	Real estate activities	12	1.000	*1.000*
Middle	30_44	2.639	2.614	Professional, scientific and technical activities	13	1.000	*1.000*
Senior	45_59	2.376	2.358	Administrative and support service activities	14	1.000	*1.000*
Old	60_Inf	2.216	2.200	Public administration and defence; compulsory social security	15	3.577	3.541
Combined age-gender groups				Education	16	1.000	*1.000*
Female Young	F15_29	2.762	2.674	Human health and social work activities	17	1.000	*1.000*
Female Middle	F30_44	2.786	2.950	Arts, entertainment and recreation	18	1.000	*1.000*
Female Senior	F45_59	2.545	3.149	Other service activities	19	3.118	3.089
Female Old	F60_Inf	2.162	2.610	Households as employers and producers for own-consumption	20	1.000	*1.000*
Male Young	M15_29	2.703	2.661	Extraterritorial organisations	21	1.000	*1.000*
Male Middle	M30_44	2.613	2.584	Not specified	22	1.000	*1.000*
Male Senior	M45_59	2.352	2.372				
Male Old	M60_Inf	2.227	2.381				

Source: Authors calculations (see in the C342:J383 range of the 'Content' sheet of the HBSwage_cons_corrected_simul.xlsx file)

Note, that in the last column the precise values of 1.000 represent branches for which no employment was reported in the GLSS5. Red figure for the bottom decile indicates that since this decile reported zero income in the survey (as discussed earlier) we assumed that their income increased in the same degree as that of the average household. A similar assumption was made for their consumption.

From Table 5.2. one can develop ideas about which social groups benefit the most from the changes over time and from the trade liberalization and which ones benefit less or even are affected adversely.
The main household sector related (aggregate) macroeconomic and distributional indicators of the two simulations are presented in Table 5.3.

Table 5.3. Selected socioeconomic indicators for Ghana computed by the MAGNET-model with multiple households

	Benchmark values for 2014	Baseline simulation results for 2030	Trade liberalization simulation results for 2030
Volume index of total household consumption	100	250.3	250.8
Volume index of total labor income	100	244.7	247.7
Consumer price index	100	101.7	99.6
Value index of total labor income	100	248.8	246.7
Consumption/Labor income ratio, per cent	109.2	111.7	110.6
Share of agricultural labor income in the total, per cent	25.4	10.3	10.5
Share of food industry labor income in the total, per cent	6.2	4.3	3.9
Share of agricultural households in total labor income, per cent	42.8	31.1	31.3
Share of agricultural households in total consumption, per cent	46.3	33.6	33.7
Share of tradable sectors affiliated households in total labor income, per cent	67.53	57.28	57.34
Share of tradable sectors affiliated households in total consumption, per cent	61.80	50.06	50.16
Share of young led households in total labor income, per cent	10.16	11.05	11.03
Share of young led households in total consumption, per cent	13.49	14.55	14.55
Share of households of Greater Accra in total labor income, per cent	28.11	32.99	32.95
Share of households of Greater Accra in total consumption, per cent	20.16	26.89	26.81
Ratio of female led households' labor income to that of male led households, per cent	15.91	16.81	16.76
Ratio of female led households' consumption to that of male led households, per cent	33.39	33.91	33.83
Inter-group Gini-index for the labor income of the per capita labor income deciles	0.6856	0.7063	0.7060
Inter-group Gini-index for the consumption of the per capita labor income deciles	0.1209	0.1924	0.2504
Inter-group Gini-index for the labor income of the per capita consumption deciles	0.2734	0.3136	0.3131
Inter-group Gini-index for the consumption of the per capita consumption deciles	0.2688	0.3227	0.3223

Source: Authors calculations (see in the C342:J383 range of the ‘Content’ sheet of the HBSwage_cons_corrected_simul.xlsx file)

We can make the following summary remarks to the figures of Table 5.3. As far as the aggregate model's macroeconomic results are concerned:

- Trade liberalization is beneficial although until 2030 only moderately. It affects the labor income more than the consumption and apparently decreases the price level (but in such Walrasian-type CGE-model one has to see what the 'numeraire' of the model is). As a related feature the trade liberalization figure for the labor incomes is lower in nominal value but higher in volume.
- The consumption/labor income ratio may depend on the transfer incomes and the possibly changing savings rates as well but looks quite stable in the simulations.

Regarding the disaggregated model results:

- The share of agricultural and food industry labor incomes decrease seriously but not in absolute volume. In the trade liberalization simulation, the share of agricultural income decreases somewhat less, but that of the food industry decreases further.
- Although the share of agricultural and food industry affiliated households in the total labor income seriously decreases but less than the agricultural and food industry labor income itself and the trade liberalization is slightly beneficial for them.
- Similar results – but to somewhat lesser extent – can be observed for the labor income shares of the tradable sectors affiliated households.
- The share of households with a young head in total labor income increases over time but practically does not differ in the two simulations.
- The share of households living in the Greater Accra region in total labor income and consumption is high and increases further over time considerably but to a somewhat less extent in the trade liberalization simulation.
- The ratio of female led households' labor income and consumption to that of male led households increases over time considerably but to a somewhat lesser extent in the trade liberalization simulation.
- The inter-group Gini-index for the labor income of the per capita labor income deciles is extremely high, which is mainly because about 16 % of the households did not report any incomes. As well-known, consumption is distributed among income deciles more evenly and improves significantly.
- The inter-group Gini-index for the labor income and consumption of the per capita consumption deciles increases over time considerably but to a somewhat less extent in the trade liberalization simulation.

To store the main results of the data processing and simulations we created the HBSwage_cons_corrected_simul.xlsx Excel-file in which the 'Ghana' sheet contains the simulation results for Ghana. The 'Content' sheet gives English and Hungarian (for the consortium leadership and management) description of the content of the individual sheets and it also contains the summary tables related to the simulations for Ghana. From the underlying formulas one can trace back the data on which basis these indicators were computed.

To understand (and to be able to explain) the disaggregated results of the MAGNET model scenarios, on the *'Content' sheet of the HBSwage_cons_corrected_simul.xlsx reporting Excel-file* we elaborated several summary tables (see in rows 33-383).

The table in the A129:Q198 range of the 'Content' sheet shows the *consumption patterns* (structure) of the selected (and differently defined) social groups.

The table in the A202:Q271 range of the 'Content' sheet shows the *share of sectors in the total labor income* of the selected (and differently defined) social groups.

Note, that in the current disaggregated simulations we have not overruled the aggregate model's result for the consumption and labor income because the goal was to estimate the social (distributional) effects and not the possible macroeconomic feedbacks (repercussions) of the consequences of the different group-specific indices of the consumption and labor income levels.

Nevertheless, in the HBSwage_cons_corrected_simul.xlsx Excel-table we created a table (see in the A274:S339 range of the 'Content' sheet) which illustrates how the different break-downs (groupings) of the household sector would affect the estimates for the *level and structure of the aggregate private consumption*. However, since these would only be the estimated demands, the equilibrium prices and consumption would be somewhat different, although somewhere between the model's results for the aggregate consumption and the estimated demands.

We can generally characterize the result by saying that, in general, the changes over the 2014-2030 period are significant in both simulations, while the differences between the trade liberalization and baseline simulations are mostly quite small, but their direction is reasonable. Obviously, a more sector specific (e.g., agricultural support, environmental protection measures affecting the various sectors quite differently) and targeted redistributive policy scenario would result in more pronounced differences. Such scenarios may be used after coordinating our simulations with the environmental and energy-policy oriented scenarios of the CGEbox model to be developed in the next months by the other research teams of our WP3 workgroup.

Further analysis of the result will be carried through after receiving constructive feedback from the stakeholders, research partners and in general, from the scientific community.

6 Conclusions and directions of further research and dissemination

Although facing many data availability, software and methodological challenges we managed to develop a strong capability to model social and distributional effects of trade and agricultural policies for various countries, in particular, for Ghana.

By acquiring the data of the Ghana Living Standard Survey for 2005/6 (GLSS5) we identified those variables of the survey, which may be used for forming various household groups, including the agricultural and non-agricultural, rich and poor, rural and urban households. By accomplishing such groupings, we estimated their labor incomes and consumption expenditures consistently with the 2014 benchmark data of the MAGNET global CGE-model. However, the GLSS5 is known to have been unable to record properly the incomes of the households, in particular the income from agricultural activity. When investigating the problem, we revealed many outlier expenditure figures too in the survey. By elaborating various cross-checks, we identified about 1 % of the sample in which households apparently one expenditure item must have been wrongly recorded. By duly correcting them we achieved a much more reliable sample. We developed a facility/module, which by applying the MAGNET-model's income and price elasticities for the individual household groups of the given stratification estimates their labor income and consumption expenditures by sector of origin consistently with the simulation results of the MAGNET model with aggregate household sector.

We ran two simulations with 2030 as the final year. The *baseline simulation* – while it is also based on the SSP1 scenario – just revealed the long-run consequences of present tendencies and the continuation of the normal developments in world trade processes. This highlighted not only the serious decrease of the GDP-share of the agriculture sector – and to some extent that of the food industry sectors – but also showed how this affects the income and consumption of the agriculture- and food industry-affiliated households, or more generally, who are involved in the most affected value chains. We elaborated various inequality measures which represent precisely or approximately the SDG 10 related SDIs (10_41 Income quintile share ratio, 10_50 Income share of the bottom 40 %, etc.). The model-estimated group-specific consumption expenditure levels and patterns also render possible the assessment of the direction and the magnitude of the changes in other SDIs (notably the 01_10 People at risk of poverty or exclusion, 02_20 Agricultural factor income per annual work unit, 05_20 Gender pay gap in unadjusted form, 07_20 Final energy consumption in households per capita, 08_10 Real GDP per capita and 08_30 Employment rate SDIs). For Ghana in particular, the computed agricultural production level of partly forested regions may be connected with SDI 15_10 (Share of forest area) assuming that in those regions without serious capital injection, the agricultural production can be increased only by continuing deforestation.

The other, the *trade liberalization simulation* (implemented via removing tariffs on the exports and imports of Ghana) was also based on the SSP1 scenario. Comparison of the results of the trade liberalization and baseline simulations showed that the trade liberalization increases the total real income by 1.2 per cent and – in addition – the share of agriculture-affiliated households in the total labour income by 0.2 percentage points. While most of the agricultural sectors seem to be quite resilient to the trade liberalization, it has a clear tendency to replace the unprocessed food by the products of the food industry (including the imported food products) in the private consumption. Model computed changes in the household groups' consumption patterns do not suggest that the trade liberalization may cause serious nutritional problems for any social groups. However, more disaggregated results of the partial equilibrium models of our WP3's modelling toolbox (using FAO data and forecasts) may reveal detailed nutritional effects and identify those foodstuffs which may become less available/affordable for the poor.

Apart from agriculture-affiliated households, households affiliated to tradable sectors benefit more than the average household from the trade liberalization, while households in the Greater Accra region and women led households benefit less. The young led households and the poor benefit from trade liberalization in the average degree.

We have made significant progress to be able to get to know and use the GLSS7 data for Ghana and improve the simulation results by them. More and more elaborated scenarios may be developed which are more relevant to the current world-market issues and reflect the new tendencies.

Clearly, more targeted, redistributive, sustainability, food security and nutrition-centred policy simulations would result in more characteristic changes in the SDIs and bigger differences in the group-specific effects. Such simulations, with as much integrated simulation scenarios as possible will be done with the MAGNET-model in cooperation with the research teams of the related Task 3.3. and Task 3.4. of the project. These simulations will be harmonized with the simulations made by their models.

In addition, by processing the Eurostat household budget survey data for 3 not too advanced EU-countries and the household surveys for Brazil and India, we demonstrated that such exercise could be done easier in the future for arbitrarily selected bigger EU-countries or other developing countries. By processing the fundamentally different household income/consumption surveys of Brazil and India so that the processed data can be fed into the MAGNET or other GTAP-data based CGE-model, we demonstrated that social and distributional effects can be estimated by our models even in developing countries in three different continents. Modelling Brazil and India is extremely promising since the feedbacks of the social and distributional effects in these countries on the whole world economy may be felt seriously even in the richer and faraway countries (e.g. in Ghana) by affecting the supply of certain commodities (see for example, India's recently introduced ban on the rice and sugar export with the aim to keep domestic prices from increasing and, hence, make rice and sugar affordable for the poor Indian households) and its world-market prices. Modelling Vietnam is also promising, since through our project partner in Vietnam we have received the quite long time series of the quite comprehensive reports on the Vietnam HBS returns (which publishes the results for each category in various break-downs, including the income quintiles) and we also got the underlying HBS data as well.

Naturally, fewer spectacular developments can be planned based on the achieved state of the research. For example, it would be worth reviewing the *income elasticities* of the MAGNET-model based on the HBS sample-based cross-sectional estimates of the point-elasticities by commodities. For Ghana such cross-sectional estimates have already been carried out by earlier researchers, but they published their results only in a rather aggregated way and did not have the possibility to compare them with other estimates (see Breisinger et al., 2008 and Table A.1. in Appendix 2 of their paper).

Many other auxiliary statistical data could be acquired and used in checking and improving our estimates. In particular, some social group specific *transfer incomes* may be built in the analysis and the model. Naturally, this would require the more careful investigation of the usually not too well measured and methodologically less clear transfer income categories of the household income surveys.

It would be also useful to revise and *develop further the various mappings and transformation matrices* between different classifications. Notably, the mapping of the HBS (COICOP) categories and GTAP sectors may be improved possibly by elaborating one-to-many type matchings and by estimating their (country specific?) value-shares (by using the approach of Cazcarro et al. (2020) and similar studies). For this the above mentioned newly acquired set of the consumption transformation matrices for the EU-countries may be used, as an auxiliary dataset (see the Cai – van Dyck (2020) paper). Each of these matrices show how much of the given country's household consumption expenditures on a certain 2-digits level COICOP category comes from the individual CPA product groups. Clearly, this break-down of the consumption expenditures are not detailed enough and we still have to match the CPA categories with the GTAP sectors (based mainly on EPU-NTUA (2013), as discussed in my earlier reports), but still, it may provide useful proportions to do the transformation.

Since the Input-Output Tables and the GTAP dataset shows only the "domestic consumption" as opposed to the "national consumption" (consumption of resident households), one has to estimate and then *separate out the consumption of inbound and outbound tourists* (it also requires a transformation of the tourist expenditures from the expenditure categories of the so called "tourist satellite accounts" to the needed sector break-down)

Knowing the structure of the GTAP database one may consider how to break-down the estimated labor income figures to those *types of labor* which are distinguished in the GTAP10.1 database (see in Aguiar et al., 2019). The main question is whether we can find an appropriate proxy category or categories in the HBS which might help to split the labor income of the various household groups to these types of labor.

For those countries where the HBS sample Gini indices are significantly lower than what is published by the World Bank (in its World Development Indicators) or by the Eurostat (based on the EU SILC survey) one may apply our method of augmenting the households' reported incomes differentially, so that the Gini indices of the samples get closer to those. Clearly, applying this new method for other countries may provide such feedbacks which may help to improve and generalize this method considerably. *Developing further the income adjustment model* can be carried through by possibly applying the so called 'bottom-coding' and 'top-coding' methods (including the exclusion of the outlier observations or replacing them by imputed values) suggested by Neugschwender (2020).

We may consider improving the method of adjusting the HBS data to their corresponding GTAP totals (e.g., by *integrating the adjustment of the consumption expenditures with the estimation* *of the incomes and savings*). Note, that the many (multidimensional or multidirectional) add-up consistency criteria and other theoretically or empirically imposed constraints will make the problem impossible to solve by such simple methods like the RAS. Therefore, a complex mathematical programming problem will have to be elaborated.

Depending on the suggestions of the stakeholders *further break-downs of the household sector* may be elaborated and analysed (in particular, certain combined groups might be interesting for policy makers, like young agricultural people).

References

Aguiar, A., Chepeliev, M., Corong, E., McDougall, R., van der Mensbrugghe, D. (2019): The GTAP Data Base: Version 10. Journal of Global Economic Analysis, 4(1), 1-27. Retrieved from https://www.jgea.org/ojs/index.php/jgea/article/view/77

Aguiar, A., Corong, E., McDougall, R. (2022): GTAP 11 Data Base Documentation - Chapter 2 : Guide to the GTAP Data Base https://www.gtap.agecon.purdue.edu/databases/v11/v11_doco.aspx

Allen, C., Metternicht, G., Wiedmann, T. (2016): National pathways to the Sustainable Development Goals (SDGs): A comparative review of scenario modelling tools. Environmental Science & Policy, 66, pp.199-207.

Bouët, A., Estrades, C., Laborde, D. (2013): Households heterogeneity in a global CGE model: an illustration with the MIRAGE-HH (MIRAGE-HouseHolds) model. LAREFI Working Paper 2013-01

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

Breisinger, C., Diao, X., Thurlow, J., Al-Hassan, R. M. (2008) : Agriculture for Development in Ghana - New Opportunities and Challenges, IFPRI Discussion Paper 00784, August 2008, International Food Policy Research Institute

Britz, W., van der Mensbrugghe, D. (2016): Reducing unwanted consequences of aggregation in large-scale economic models - a systematic empirical evaluation with the GTAP model. Economic Modelling 2016: 462-473

Britz, W., van der Mensbrugghe, D. (2018): CGEBox: A Flexible, Modular and Extendable Framework for CGE Analysis in GAMS – Journal of Global Economic Analysis, Volume 3 (2018), No. 2, pp. 106-177.

Britz, W., Witzke P. (2014): CAPRI model Documentation 2014. http://www.capri-model.org/docs/capri_documentation.pdf

Britz, W. (2021). CGEBox - a flexible and modular toolkit for CGE modelling with a GUI. https://www.ilr.uni-bonn.de/em/rsrch/cgebox/cgebox_GUI.pdf, Accessed May 27, 2022.

Cai, M., van Dyck, T. (2020): Bridging between economy-wide activity and household-level consumption data: Matrices for European countries, Data in Brief 30 (2020) 105395, Elsevier, https://doi.org/10.1016/j.dib.2020.105395

Cazcarro, I., Amores, A. F., Arto, I. , Kratena, Kurt (2020): Linking multisectoral economic models and consumption surveys for the European Union, Economic Systems Research, DOI: 10.1080/09535314.2020.1856044

Corong, Erwin L., Hertel, T., McDougall, R., Tsigas, M., van der Mensbrugghe, D. (2017): The Standard GTAP Model, Version 7 - Journal of Global Economic Analysis, Volume 2 (2017), No. 1, pp. 1-119.

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

EPU-NTUA (2013): Assessment of Policy Impacts on Sustainability in Europe - Baseline and exploratory scenarios, parameters and validation, Deliverable no.: D 4.1, Grant Agreement no.: 283121, Project Acronym: APRAISE, Theme: ENV.2011.4.2.1-1: Efficiency Assessment of Environmental Policy Tools Related to Sustainability.

Ghana Statistical Service (2008): Ghana Living Standards Survey – Report of the fifth round (GLSS 5), September 2008

Ghana Statistical Service (2019): Ghana Living Standards Survey (GLSS) 7 – Main Report, June 2019

GSS – IFPRI (2007): A 2005 Social Accounting Matrix (SAM) For Ghana, report of the Ghana Statistical Services (GSS) and International Food Policy Research Institute (IFPRI) under the Ghana Strategy Support Program (GSSP), Edited by Breisinger, C., J. Thurlow (IFPRI), Magnus Duncan (GSS), October, 2007

JRC (2023): A NUTS-2 European Union interregional system of Social Accounting Matrices for the year 2017: The RHOMOLO V4 dataset - JRC Working Papers on Territorial Modelling and Analysis No. 01/2023

Kuiper, M., Shutes, L. (2014): „FoodSecure Technical Paper: Expanding the household coverage of global simulation models: an application to Ghana" LEI Wageningen UR

Lanz, B., Rutherford, T. (2016): GTAPinGAMS: Multiregional and Small Open Economy Models. Journal of Global Economic Analysis, 1(2): 1-77. DOI: 10.21642/JGEA.020101AF.

Laquai, V., Krupin, V., Kavalari, A., Révész, T., Abdelhady, M., Banse, M., Hagemejer, J., Jendrzejewski, B., Zawalińska, K., Araujo Enciso, S., Matthey, H. (2022): Deliverable D3.1: Report on linking SDG indicators with models in the TRADE4SD toolbox

Luu, N. et al. (2020): Mapping trade to household budget survey: A conversion framework for assessing the distributional impact of trade policies, OECD Trade Policy Papers, No. 244, OECD Publishing, Paris. http://dx.doi.org/10.1787/5fc6181b-en

M'Barek, R., Britz, W., Burrell, A., Delincé, J. (2012): An Integrated Modelling Platform for Agri-economic commodity and Policy Analysis (iMAP) – A look back and a way forward – European Commission, Joint Research Centre, Scientific and Policy Report, Report EUR 25267 EN, ISBN 978-92-79-23554-2 (pdf) ISBN 978-92-79-23553-5 (print) ISSN 1831-9424 (online) ISSN 1018-5593 (print) doi:10.2791/78721

McDonald, S., Sonmez, Y. (2004): Augmenting the GTAP Database with Data on Inter-Regional Transactions, mimeo, Purdue University, Global Trade Analysis Program.

McDonald, S., Thierfelder, K. (2004): Deriving a Global Social Accounting Matrix from GTAP versions 5 and 6 Data. GTAP Technical Paper No. 22.

Minor, P., Walmsley, T. (2013): „MyGTAP Data Program: A Program for Customizing and Extending the GTAP Database - Multiple Households - Split Factors - Remittances - Foreign Aid - Transfers" (MyGTAP documentation) https://impactecon.com/store/

Mueller M, Ferrari E., 2011. An Inventory of Datasets for the Compilation of Regional Social Accounting Matrices for the EU. In Conference Proceedings: 14th Annual Conference on Global Economic Analysis. West Lafayette (USA): Purdue University; p. 1-32. JRC65667 (see at https://publications.jrc.ec.europa.eu/repository/handle/JRC65667 and https://www.gtap.agecon.purdue.edu/resources/res_display.asp?RecordID=3619)

Neugschwender, J. (2020): Top and Bottom Coding at LIS, Technical Working Paper Series of the Luxembourg Income Study (LIS), No. 9, September 2020

O'Neill, B.C., Kriegler, E., Ebi, K.L., Kemp-Benedict, E., Riahi, K., Rothman, D.S., Van Ruijven, B.J., Van Vuuren, D.P., Birkmann, J., Kok, K., Levy, M., Solecki, W. (2017): The roads ahead: narratives for shared socioeconomic pathways describing world futures in the 21st century, Global Environmental Change, Vol. 42, pp. 169-180

Révész, T. (2019): Further developments in the decomposition of the Eurostat net tax matrices and their transformation to GTAP format, paper presented at the "22nd Annual Conference of Global Economic Analysis" held in Warsaw, June 19-21, 2019. Corvinus University Budapest, https://www.gtap.agecon.purdue.edu/resources/res_display.asp?RecordID=5853

Révész, T. (2021): Report on linking SDG indicators with models in the TRADE4SD toolbox – Deliverable D3.1. of the TRADE4SD project

Révész, T. (2022): Report on the modelling of social and distributional impact of trade and sustainability policies – Background document of Deliverable D3.2. of the TRADE4SD project
Révész, T. (2022a): Households related micro data for the modelling of social and distributional impact of trade and sustainability policies – Background document of Deliverable D3.2. of the TRADE4SD project
Révész, T. (2022c): Processing households budget survey data for the modelling of social and distributional impact of trade and sustainability policies, – Background document of Deliverable D3.2. of the TRADE4SD project
Révész, T. (2023): Models for adjusting the Hungarian, Bulgarian and Portuguese Household Budget Surveys reported consumption and income data to the GTAP database and official Gini indexes; Background document of Deliverable D3.2. of the TRADE4SD project
RHOMOLO 2018. Lecca, P., Barbero, J., Christensen, M., Conte, A., Di Comite, F., Diaz-Lanchas, J., Diukanova, O., Mandras, G., Persyn, D. and Sakkas, S., RHOMOLO V3: A Spatial Modelling Framework, EUR 29229 EN, Publications Office of the European Union, Luxembourg, 2018, ISBN 978-92-79-85886-4, doi:10.2760/671622, JRC111861.
Round, J. (2003) : Constructing SAMs for Development Policy Analysis: Lessons Learned and Challenges Ahead, Economic Systems Research, 15:2, 161-183, DOI: 10.1080/0953531032000091153
Rutherford, T., Arbor, M. (2005): GTAP6InGAMS: The Dataset and Static Model. Prepared for the Workshop: "Applied General Equilibrium Modeling for Trade Policy Analysis in Russia and the CIS", The World Bank Resident Mission, Moscow, December 1-9, 2005, available at http://www.mpsge.org/gtap6/gtap6gams.pdf
Sahin, Sebnem, van der Mensbrugghe, Dominique (2007) : The effects of consumer demand parameters on trade policy analysis: An application tot he World bank LINKAGE model**,** Paper prepared for the 10th Annual Conference on Global Economic Analysis hosted by Purdue University, West Lafayette, Indiana, USA, June 7-9, 2007 (see also among the GTAP resources #2358 at GTAP Resources: Resource Display: The effects of consumer demand parameters on trade... (purdue.edu))
Szabó, Sándorné (1996): Estimate for the 'true' level and variance of the incomes (In Hungarian: Becslés a "valódi" jövedelemszintre és szóródásra), Statistical Review of Hungary (Statisztikai Szemle), Vol. 74., No. 2., pp. 126-134.
van Vuuren, D.P., Kriegler, E., O'Neill, B.C., Ebi, K.L., Riahi, K., Carter, T.R., Edmonds, J., Hallegatte, S., Kram, T., Mathur, R., Winkler, H. (2014): A new scenario framework for Climate Change Research: scenario matrix architecture, Climate Change, Vol. 122, pp. 373-386
Verikios, G., K. Hanslow (1999): Modelling the effects of implementing the Uruguay Round: A comparison using the GTAP model under alternative treatments of international capital mobility, Productivity Commission, Canberra, Australia. Paper presented to the Second Annual Conference on Global Economic Analysis, Denmark, 20-22 June 1999.
Walmsley, T., P. Minor (2013). MyGTAP Model: A Model for Employing Data from the MyGTAP Data Application—Multiple Households, Split Factors, Remittances, Foreign Aid and Transfers. Working papers 283496, Purdue University, Center for Global Trade Analysis, Global Trade Analysis Project.
Woltjer, G., Kuiper, M. (2014): The MAGNET Model - Module description, LEI Wageningen UR Wageningen, August 2014, MANUAL LEI 4-057
World Bank (1997): World Development Report 1997 (see the World Development Indicator Tables)

Appendix 1: Serial number, code and description of the GSEC3 sectors

Number	Code	Description
1	**pdr**	Rice: seed, paddy (not husked)
2	**wht**	Wheat: seed, other
3	**gro**	Other Grains: maize (corn), sorghum, barley, rye, oats, millets, other cereals
4	**v_f**	Veg & Fruit: vegetables, fruit and nuts, edible roots and tubers, pulses
5	**osd**	Oil Seeds: oil seeds and oleaginous fruit
6	**c_b**	Cane & Beet: sugar crops
7	**pfb**	Fibres crops
8	**ocr**	Other Crops: stimulant; spice and aromatic crops; forage products; plants and parts of plants used primarily in perfumery, pharmacy, or for insecticidal, fungicidal or similar purposes; beet seeds (excluding sugar beet seeds) and seeds of forage plants; natural rubber in primary forms or in plates, sheets or strip, living plants; cut flowers and flower buds; flower seeds, unmanufactured tobacco; other raw vegetable materials nec
9	**ctl**	Cattle: bovine animals, live, other ruminants, horses and other equines, bovine semen
10	**oap**	Other Animal Products: swine; poultry; other live animals; eggs of hens or other birds in shell, fresh; reproductive materials of animals; natural honey; snails, fresh, chilled, frozen, dried, salted or in brine, except sea snails; edible products of animal origin n.e.c.; hides, skins and furskins, raw; insect waxes and spermaceti, whether or not refined or coloured
11	**rmk**	Raw milk
12	**wol**	Wool: wool, silk, and other raw animal materials used in textile
13	**frs**	Forestry: forestry, logging and related service activities
14	**fsh**	Fishing: hunting, trapping and game propagation including related service activities, fishing, fish farms; service activities incidental to fishing
15	**coa**	Coal: mining and agglomeration of hard coal, lignite and peat
16	**oil**	Oil: extraction of crude petroleum, service activities incidental to oil and gas extraction excluding surveying (part)
17	**gas**	Gas: extraction of natural gas, service activities incidental to oil and gas extraction excluding surveying (part)
18	**oxt**	Other Mining Extraction (formerly omn): mining of metal ores; other mining and quarrying
19	**cmt**	Cattle Meat: fresh or chilled; meat of buffalo, fresh or chilled; meat of sheep, fresh or chilled; meat of goat, fresh or chilled; meat of camels and camelids, fresh or chilled; meat of horses and other equines, fresh or chilled; other meat of mammals, fresh or chilled; meat of mammals, frozen; edible offal of mammals, fresh, chilled or frozen
20	**omt**	Other Meat: meat of pigs, fresh or chilled; meat of rabbits and hares, fresh or chilled; meat of poultry, fresh or chilled; meat of poultry, frozen; edible offal of poultry, fresh, chilled or frozen; other meat and edible offal, fresh, chilled or frozen; preserves and preparations of meat, meat offal or blood; flours, meals and pellets of meat or meat offal, inedible; greaves
21	**vol**	Vegetable Oils: margarine and similar preparations; cotton linters; oil-cake and other residues resulting from the extraction of vegetable fats or oils; flours and meals of oil seeds or oleaginous fruits, except those of mustard; vegetable waxes, except triglycerides; degras; residues resulting from the treatment of fatty substances or animal or vegetable waxes; animal fats
22	**mil**	Milk: dairy products
23	**pcr**	Processed Rice: semi- or wholly milled, or husked
24	**sgr**	Sugar and molasses
25	**ofd**	Other Food: prepared and preserved fish, crustaceans, molluscs and other aquatic invertebrates; prepared and preserved vegetables, pulses and potatoes; prepared and preserved fruits and nuts; wheat and meslin flour; other cereal flours; groats, meal and pellets of wheat and other cereals; other cereal grain products (including corn flakes); other vegetable flours and meals; mixes and doughs for the preparation of bakers' wares; starches and starch products; sugars and sugar syrups n.e.c.; preparations used in animal feeding; lucerne (alfalfa) meal and pellets; bakery products; cocoa, chocolate and sugar confectionery; macaroni, noodles, couscous and similar farinaceous products; food products n.e.c.
26	**b_t**	Beverages and Tobacco products

27	**tex**	Manufacture of textiles
28	**wap**	Manufacture of wearing apparel
29	**lea**	Manufacture of leather and related products
30	**lum**	Lumber: manufacture of wood and of products of wood and cork, except furniture; manufacture of articles of straw and plaiting materials
31	**ppp**	Paper & Paper Products: includes printing and reproduction of recorded media
32	**p_c**	Petroleum & Coke: manufacture of coke and refined petroleum products
33	**chm**	Manufacture of chemicals and chemical products
34	**bph**	Manufacture of pharmaceuticals, medicinal chemical and botanical products
35	**rpp**	Manufacture of rubber and plastics products
36	**nmm**	Manufacture of other non-metallic mineral products
37	**i_s**	Iron & Steel: basic production and casting
38	**nfm**	Non-Ferrous Metals: production and casting of copper, aluminium, zinc, lead, gold, and silver
39	**fmp**	Manufacture of fabricated metal products, except machinery and equipment
40	**ele**	Manufacture of computer, electronic and optical products
41	**eeq**	Manufacture of electrical equipment
42	**ome**	Manufacture of machinery and equipment n.e.c.
43	**mvh**	Manufacture of motor vehicles, trailers and semi-trailers
44	**otn**	Manufacture of other transport equipment
45	**omf**	Other Manufacturing: includes furniture
46	**ely**	Electricity; steam and air conditioning supply
47	**gdt**	Gas manufacture, distribution
48	**wtr**	Water supply; sewerage, waste management and remediation activities
49	**cns**	Construction: building houses factories offices and roads
50	**trd**	Wholesale and retail trade; repair of motor vehicles and motorcycles
51	**afs**	Accommodation, Food and service activities
52	**otp**	Land transport and transport via pipelines
53	**wtp**	Water transport
54	**atp**	Air transport
55	**whs**	Warehousing and support activities
56	**cmn**	Information and communication
57	**ofi**	Other Financial Intermediation: includes auxiliary activities but not insurance and pension funding
58	**ins**	Insurance (formerly isr): includes pension funding, except compulsory social security
59	**rsa**	Real estate activities
60	**obs**	Other Business Services nec
61	**ros**	Recreation & Other Services: recreational, cultural and sporting activities, other service activities; private households with employed persons (servants)
62	**osg**	Other Services (Government): public administration and defense; compulsory social security, activities of membership organizations n.e.c., extra-territorial organizations and bodies
63	**edu**	Education
64	**hht**	Human health and social work
65	**dwe**	Dwellings: ownership of dwellings (imputed rents of houses occupied by owners)

Source: Aguiar et al. (2019), p. 22. and Aguiar et al. (2022)

Appendix 2: Approximate correspondence of the household sector relevant GSEC3 sectors and MAGNET-model sectors

Serial Number of the corresponding GSEC3 sectors	Description of the MAGNET-model's sectors	serial number and code
	lower case initials represent the MAGNET-model's special sectors !	**of the relevant MAGNET-model sectors**
1 pdr	Paddy Rice	**1 pdr**
2 wht	Wheat	**2 wht**
3 gro	Cereal grains nec	**3 grain**
5 osd	Oil seeds	**4 oils**
6 c_b	Sugar cane, sugar beet	**5 sug**
4 v_f	Vegetables, fruit, nuts	**6 hort**
7 pfb, part of 8 ocr	Crops nec	**7 crops**
part of 8 ocr	Other agriculture	**8 oagr**
part of 9 ctl	cattle sector	**9 cattle**
part of 9 ctl	sheep,goats,horses	**10 othctl**
part of 10 oap	poultry sector	**11 pltry**
12 wol	Wool, silk-worm cocoons	**12 wol**
part of 10 oap	Pig and other animal product	**13 pigpls**
11 rmk	Raw milk	**14 milk**
part of 19 cmt	beef meat	**15 bfmt**
part of 19 cmt	Meat: other cattle,sheep,goats,horse	**16 othcmt**
part of 20 omt	poultry meat	**17 pulmt**
part of 20 omt	Other meat product nec	**18 othmt**
22 mil	Dairy products	**19 dairy**
24 sgr	Sugar and molasses	**20 sugar**
part of 21 vol	Vegetable oils and fats	**22 vol**
23 pcr	Processed rice	**23 pcr**
part of 25 ofd	Processed food	**24 ofd**
part of 25 ofd	Animal feed	**25 feed**
part of 14 fsh	Wild fish	**26 wfish**
part of 14 fsh	Aquaculture	**27 aqcltr**
part of 25 ofd	Fish processing	**28 fishp**
13 frs	Forestry	**29 frs**
part of 8 ocr	Plantation	**30 plan**
32 p_c	Petroleum, coal products	**32 petro**
17 gas	Gas	**38 gas**
15 coa	Coal	**39 coa**
part of 46 ely	Electricity	**40 ely**
part of 46 ely	electricity from gas	**42 ely_g**
part of 46 ely	electricty from hydro	**44 ely_h**

35 rpp, part of 33 chm and 34 bph	Chemical,rubber, other plastic prods	**51 othcrp**
part of 33 chm	fertilizer	**52 fert**
	mixed fossil biochemical sector	**53 f_chem**
sectors 18, 26-31, 36-45	Other industry	**54 othind**
	pellet sector	**55 pel**
47 gdt	Gas manufacture, distribution	**56 gas_dist**
52 otp, 53 wtp, 54 atp, 55 whs	Transport sector	**57 trans**
the rest	Services	**58 ser**

Source: Authors assessment based on the MAGNET-model documentation and data (see in the A35:C94 range of the 'Content' sheet of the HBSwage_cons_corrected_simul.xlsx file)

Appendix 3: Ghanaian Household budget shares and income elasticities

Table A.1. Ghanaian Household budget shares and income elasticity

	Current budget share		Marginal budget share		Income elasticity	
	Urban	Rural	Urban	Rural	Urban	Rural
Foods	43.5	52	34.6	49	0.8	0.9
Maize	0.8	1.8	0.4	1.2	0.4	0.7
Rice and wheat	3.7	4.3	2.6	4.4	0.7	1
Roots	3	2.6	2.2	3.3	0.7	1.3
Other food	7.2	8.6	5.2	7.3	0.7	0.8
Plantain	1.2	1.1	0.9	1.3	0.8	1.3
Chicken	1.6	1.1	2	1.5	1.2	1.3
Other livestock	10.8	15.6	8.5	14.4	0.8	0.9
Fish	1.9	2.1	1.8	2.3	1	1.1
Other foods	13.3	14.7	10.9	13.2	0.8	0.9
Non-foods	46.1	37	56.6	40	1.2	1.1
Clothing	10.4	11	8.9	11	0.9	1
Other manufacturing	7	9.6	6.9	9.7	1	1
Fuels	3.8	5.1	8	3.5	2.1	0.7
Durable equipment	9.4	4.8	20.9	7.6	2.2	1.6
Water and electricity	0.5	0.1	0.7	0.2	1.4	2.1
Services	25.4	17.4	20	19	0.8	1.1

Source: Breisinger et al. (2008), p.12. (Authors' estimates using the 2005/06 Ghana Living Standards Survey (GLSS5)).

Appendix 4: Corrections of the recorded expenditures of the GLSS5

House-hold identi-fier	Code of expenditure category	Name of expenditure category (short form)	Correc-ted amount in ₵	Correc-tion was made by dividing by	*Originally recorded total expenditure in ₵*	*Bonus, social allowances derived from wage job*	*Total income (as recorded) in ₵*	*income decile affili-ation*	*Original consumption /income ratio*	*Share of the lar-gest item (original), %*
5041/5	*FD_P*	own produced food	93574	1000	103,485,504	1,500,000	3,412,572	10	21	90.42
5048/11	*FDNONALC_B*	purchased soft drinks	731241	1000	782,955,520	-	3,500,000	10	224	93.39
5064/4	*FD_P*	own produced food	112815	1000	124,701,576	6,813,334	9,780,000	10	8	90.47
5101/1	*FD_P*	own produced food	121281	1000	125,180,448	1,000,000	-	10	125	96.88
5147/8	*HSCLOTH*	clothes	300057	10000	3,017,315,840	24,504,000	24,024,000	10	62	99.45
5150/8	*RCNEWS*	newspapers, books, stationery	208460	10000	2,100,800,768	-	2,000,000	10	1050	99.23
5166/12	*HSTEXTIL*	textiles	160020	10000	1,618,426,240	12,000,000	-	10	135	98.87
5171/15	*RENT_ACT*	actual rent paid	8000006	1000	8,012,210,196	2,400,000	-	9	3338	99.85
5183/2	*COMSERVE*	communication service	800001	10000	8,032,047,104	22,657,142	150,000	10	352	99.60
5197/7	*FD_B*	purchased food	3650424	10000	36,505,759,744	1,800,000	15,600,000	10	2098	100.00
5206/14	*FD_B*	purchased food	983220	1000	987,698,944	3,650,000	9,900,000	10	73	99.55
5206/5	*FD_B*	purchased food	1644328	10000	16,501,017,600	9,650,000	6,000,000	10	1054	99.65
5209/7	TOTEDUC	education	2433342	10000	24,367,235,072	5,000,000	-	10	4873	99.86
5225/1	*FD_P*	own produced food	420704	1000	437,248,128	3,200,000	7,284,714	10	42	96.22
5232/17	*FD_P*	own produced food	486000	1000	499,094,432	3,600,000	1,920,000	10	90	97.38
5232/9	*FD_P*	own produced food	205308	1000	213,307,216	3,200,000	4,040,000	10	29	96.25
5253/7	*HSCLOTH*	clothes	920404	10000	9,229,233,152	-	-	10	not applicable	99.73
5271/10	*FD_P*	own produced food	358104	1000	385,989,440	2,900,000	2,700,000	10	69	92.78
5271/4	*FD_P*	own produced food	97225	1000	118,821,976	500,000	3,402,857	10	30	81.82
5271/5	*FD_P*	own produced food	164502	1000	194,152,592	1,000,000	37,637,144	10	5	84.73
5271/7	*FD_B*	purchased food	1832105	10000	18,405,652,480	1,076,749,984	97,522,856	10	16	99.54
5272/2	*FD_B*	purchased food	1827097	10000	18,297,470,976	36,000,000	2,508,572	10	475	99.86
5273/10	*FD_B*	purchased food	1223427	1000	1,245,199,488	13,035,714	87,360,000	10	12	98.25

5275/10	*COMSERVE*	communication service	3650374	10000	36,520,648,704	-	5,800,000	10	6297	99.95
5295/7	*HSCLOTH*	clothes	500705	1000	516,374,400	2,607,143	2,475,714	10	102	96.97
5321/4	*FD_P*	own produced food	297199	1000	311,572,096	3,000,000	2,001,572	10	62	95.39
5407/9	*FD_P*	own produced food	158958	1000	167,390,880	6,257,143	2,388,572	10	19	94.96
5408/10	*FD_B*	purchased food	1825451	10000	18,260,269,056	-	31,302,858	10	583	99.97
5408/13	*FD_B*	purchased food	658107	10000	6,602,148,864	-	2,820,000	10	2341	99.68
5409/12	*HSREPAIR*	maintenance & repair	180037	10000	1,818,589,440	25,550,000	8,317,143	10	54	99.00
5409/8	*HSCLOTH*	clothes	1000839	1000	1,025,865,472	6,778,572	3,182,714	10	103	97.56
5412/6	*FD_P*	own produced food	164388	1000	175,903,424	12,800,000	5,060,000	10	10	93.45
5430/12	*FD_B*	purchased food	377349	1000	393,199,904	2,000,000	-	10	197	95.97
5432/12	*FD_B*	purchased food	551709	1000	572,300,096	13,328,572	26,330,856	10	14	96.40
5432/13	*FD_B*	purchased food	1826365	10000	18,280,968,192	44,000,000	20,470,858	10	284	99.91
5432/14	*HOTCAT*	hotels and catering	1097126	1000	1,115,014,144	-	3,888,572	10	287	98.40
5434/12	TRSERVE	public transport	7300196	10000	73,029,795,840	60,500,000	129,000,000	10	385	99.96
5438/10	*FD_B*	purchased food	377290	1000	395,462,240	-	200,000	10	1977	95.40
5439/9	*FD_B*	purchased food	621910	1000	635,747,136	18,000,000	24,000,000	10	15	97.82
5441/13	*HOTCAT*	hotels and catering	109561	10000	1,107,049,472	1,825,000	-	10	607	98.97
5442/7	*FD_B*	purchased food	734825	1000	739,232,384	2,607,143	9,840,000	10	59	99.40
5454/9	*FD_B*	purchased food	373337	1000	377,223,232	10,950,000	-	10	34	98.97
5469/4	*FD_P*	own produced food	71917	1000	73,781,648	9,500,000	42,925,716	10	1	97.47
5469/5	*FD_P*	own produced food	51985	1000	56,874,724	4,000,000	1,920,000	9	10	91.40
5469/9	*FD_P*	own produced food	59277	1000	65,274,124	13,500,000	1,800,000	8	4	90.81
5470/13	*FD_P*	own produced food	92705	1000	100,657,800	8,864,286	1,220,000	10	10	92.10
5470/6	*FD_P*	own produced food	206809	1000	210,679,648	73,000,000	946,857	10	3	98.16
5474/15	*FD_P*	own produced food	36600	1000	39,574,628	4,000,000	536,000	9	9	92.48
5483/14	*FD_P*	own produced food	73560	1000	79,594,544	2,340,000	2,910,000	10	15	92.42
5485/17	*FD_P*	own produced food	40722	1000	45,112,624	540,000	10,657,143	10	4	90.27
5495/1	*FD_P*	own produced food	47811	1000	58,363,124	-	18,821,144	9	3	81.92
5495/13	*FD_P*	own produced food	67805	1000	73,332,000	-	2,520,000	9	29	92.46
5503/4	*FD_P*	own produced food	110397	1000	116,039,128	2,680,000	4,030,000	10	17	95.14
5523/15	*FD_P*	own produced food	42397	1000	49,120,528	-	1,140,000	7	43	86.31
5545/15	*FD_P*	own produced food	38916	1000	42,996,052	3,000,000	740,000	9	11	90.51

5549/5	*FD_P*	own produced food	18251	10000	183,860,064	720,000	206,000	10	199	99.27