PIRAMID: a new methodology to build baselines for CGE models

BY KRZYSZTOF WOJTOWICZ[a], LUIS REY LOS SANTOS[b] , UMED TEMURSHO[c], BERT SAVEYN[d], TOON VANDYCK[e], MARIE TAMBA[f] AND MATTHIAS WEITZEL[g]

[a] JRC Seville, Calle Inca Garcilaso 3, Seville, Spain 41092
(e-mail: krzysztof.wojtowicz@ec.europa.eu).
[b] JRC Seville, Calle Inca Garcilaso 3, Seville, Spain 41092
(e-mail: luis.rey-los-santos@ec.europa.eu).
[c] JRC Seville, Calle Inca Garcilaso 3, Seville, Spain 41092
(e-mail: Umed.TEMURSHOEV@ext.ec.europa.eu).
[d] JRC Seville, Calle Inca Garcilaso 3, Seville, Spain 41092
(e-mail: bert.saveyn@ec.europa.eu).
[e] JRC Seville, Calle Inca Garcilaso 3, Seville, Spain 41092
(e-mail: toon.vandyck@ec.europa.eu).
[f] JRC Seville, Calle Inca Garcilaso 3, Seville, Spain 41092
(e-mail: marie.tamba@ec.europa.eu).
[g] JRC Seville, Calle Inca Garcilaso 3, Seville, Spain 41092
(e-mail: matthias.weitzel@ec.europa.eu).

Ex ante policy assessments often need to consider economic impacts over the coming decades of for a future point in time, a task which requires a counterfactual view of how the economy would evolve in absence of the policy measure. We present a novel method (PIRAMID) to develop a baseline for CGE models by projecting a consistent combination of data into the future. In a first step, Input-Output tables are projected forward to 2050, along with satellite data and projections from demographic, macroeconomic, and energy system models. In a second step, we calibrate a CGE model to the resulting global, economy-wide dataset. The method captures structural change, is flexible to integrate various data sources and can be shared across different modelling teams, enabling data transparency and fostering an open stakeholder engagement process.

JEL codes: C68, D58, F01, Q43, Q58

Keywords: Baseline; CGE modelling; Data integration; Calibration.

1. Introduction

Computable General Equilibrium (CGE) models have become one of the most used tools for economic analysis. Originally, CGE models were developed for short-term policy assessment such as fiscal and trade policies. In this context, base year economic structure was suitable for the comparison of policy scenarios. More recently, the need to address long-term-issues, such as energy and climate policies, has motivated the development of dynamic CGE models with increasingly long time horizons. One drawback of such a use is the persistence of the initial economic structure. When analysing long-term policies, the economic structure of the base year may not be valid and, therefore, baselines are required to assess the implication of different policies in the future. A baseline describes the evolution of the economy without additional policies and serves as a benchmark for assessing economic, energy and climate policies.

The increasing interest in dynamic models has led to an increasing interest in baselines[h]. It is well known that the numeric results of a model are very dependent on the economic structure and, thus, baselines are essential for any assessment of counterfactual policy scenarios. For example, the cost of a climate mitigation policy in 2050 will depend on the preferences, productivities, efficiencies and technological options assumed in the reference path. Although the importance of the baselines is well known, still there is a wide range of different approaches on how to build a baseline and its creation is often not well

[h] In January 2018 the OECD hosted the workshop on "Shaping long-term baselines with CGE Models". More than 60 participants, representing more than 20 different models, contrast and compare the different strategies for baseline building across modelling teams. Following the meeting, in 2019 the JGEA will publish a special issue which seeks to identify state-of-the-art practices and help CGE modellers build their own baselines.

documented. This lack of consensus hinders the comparative assessment between models. When comparing results of different models, it is hard to decompose differences in policy assessment to differences in models and baseline data.

Roughly speaking, baselines in most CGE models are built in two steps. First, base year data is used to calibrate the parameters of the model. Then, key parameters such as the endowment of production factors and various productivity parameters of the model are projected over time to meet exogenously determined targets, such as GDP, aggregate energy use, GHGs emissions, etc. In our opinion, this approach presents two drawbacks. First, the number of exogenously determined target is generally very limited. The more targets are set, the more difficult calibration gets. This might make targets less achievable and might lead to strange/inconsistent parameter values. Second, baselines are created using a specific model and, therefore, are dependent on the structure of that model. This makes it difficult for other modelling teams to replicate the same baseline.

This paper presents PIRAMID (Platform to Integrate, Reconcile and Align Model-based Input-output Data) which is an alternative methodology to build a baseline. We propose to reverse the order of the two steps in the conventional baseline building. In PIRAMID, firstly, base year data structure is projected over time to match the exogenously determined targets. Thus, our methodology requires the projection of base year IOTs integrating consistently data from different external sources. As a result, we generate a data base composed of 78 IOTs, which represent the economic structure of 13 countries/regions in 6 future years. The data base is open and available for download. In a second step, the projected data is used to calibrate the parameters of the model for each period.

The literature on updating, projecting, balancing and/or estimation methodologies of IOTs is really huge by now and is still growing. As a brief summary, the methods that were shown to be performing well in practice include the so-called RAS method (see e.g. Leontief, 1941; Stone, 1961; Bacharach, 1970), the GRAS method, an extension of RAS, which allows for positive and negative elements (Temurshoev et al., 2013), the minimum sum of cross entropies (MSCE) approach (Golan et al., 1994; Golan and Vogel, 2000) and the matrix updating methods proposed by Harthoorn and van Dalen (1987) and Kuroda (1988). In this paper we use a version of the GRAS method adapted to multi-regional IO setting (MR-GRAS).

The projection of the IOTs is subject to additional exogenous assumptions, which may reflect the expected macroeconomic evolution and/or sectoral changes. The macroeconomic assumptions can reflect the evolution of the GDP, private and public consumption, investment, exports, imports, tax rates, capital and labour payments, labour force, unemployment rates, population, etc. These data can be obtained from external sources and, ideally, should be internally consistent. In addition to the macroeconomic assumptions, we can impose projections on specific sectors of the economy (e.g. energy, agriculture, transport). In this paper, we impose energy data projections from partial equilibrium energy models such as POLES and PRIMES. We fix the use and production of energy products in the projection of the IOTs. Energy data is balanced not only in value terms but also in quantities and, thus, we can replicate CO_2 emissions estimated in energy models. Similarly to Le Treut et al (2014), based on the accuracy of bottom-up models, we introduce sector specific data within an IOT framework. As shown by Le Treut et al (2014), the use of accurate energy data is highly important when analysing the general equilibrium consequences of energy and climate policies.

To our knowledge, this is the first attempt to create a baseline for CGE models projecting base year data using techniques based on input-output literature[i]. PIRAMID presents three advantages compared to conventional baselines:

i. Flexibility. It allows for a better control of the variables we are interested in and for introducing structural changes.
ii. Reproducibility. The projected IOTs are not model dependant and can be used by other modelling teams.
iii. Transparency. The baseline is built based on transparent macroeconomic assumptions and sector specific data. In case of discrepancy, these assumptions can be corrected and improved.

The rest of this report is organized as follows. Section 2 presents the Data Base of PIRAMID which is used a baseline. Section 3 discusses the methodology used in projecting IOTs. Section 4 explains how PIRAMID can be used as a baseline for CGE models. Section 5 illustrates some data of PIRAMID. Section 6 concludes.

2. Structure of the Data Base

The Data Base of PIRAMID is produced projecting base year IOTs. The resulting data is used as a baseline for CGE models[j]. In this section we present the structure of the Data Base.

The Data Base is composed of 78 Input Output Tables, which represent the economic structure of 13 countries/regions in 6 future years (2025, 2030, 2035, 2040, 2045 and 2050). The IOTs distinguish 31 commodities and 4 production factors (skilled and unskilled labour, capital and natural reserves). Other main features of IOTs are also included: private and government consumption,

---

[i] Böhringer et al. (2009) highlight the importance of baseline projections and perform a sensitivity analysis on the implications of alternative baseline projections. However, they do not project IOTs.

[j] The Data Base is available for download in Rey et al. (2019).

investment, exports and imports, transport margins and 5 types of taxes/subsidies[k].

The GTAP data base is the benchmark of our projections and, therefore, the countries/regions and the commodities represented in these tables have a direct mapping with GTAP sectoral and regional classification (see Tables 3 and 4 in the Appendix). Given that our main interest is in energy data, energy and energy intensive sectors are very similar to GTAP classification. We keep GTAP energy commodities: coal, gas, crude oil, oil products and electricity. Besides, similarly to GTAP-Power, we disaggregate the electricity sector and include 8 generation technologies: coal, oil, gas, nuclear, hydro, biomass, wind and solar. The remaining sectoral aggregation of the data base represents 3 agricultural sectors, 9 manufacturing sectors, construction, 3 transportation sectors and 2 service sectors. The regional/country classification represents the biggest economies (in terms of GDP) individually: China, United States, European Union, Russia, India and Japan.

Although the IOTs are presented individually, they are consistent among them geographically and over time. The geographical consistency is achieved through the bilateral export and import representation. The IOTs, which show total exports and imports, are complemented with the bilateral trade flows in separate tables. The temporal consistency is achieved through the capital investment link. As explained in the Annex, capital payments and investment values, which are represented in the IOTs, are consistent with the law of motion of the capital. Thus, the stock of capital is equal to the stock of capital (minus depreciation) plus the investment in the previous year.

Energy balances and GHG emissions

The IOTs are complemented with the Energy Balances which are consistent with the IOTs. The Energy Balances are built from the projections of partial equilibrium energy models (e.g. POLES and PRIMES) and show the quantities (in Mtoe) that are behind the values of the IOTS. They show the energy use, energy production and energy trade. As mentioned above, we distinguish between 5 energy products: coal, gas, crude oil, oil products and electricity, which is produced from 8 generation technologies. We report the energy use in the production of 31 commodities and in the final demand (i.e. households, government and investment). Energy trade shows bilateral exports and imports of the energy products.

In addition to the Energy Balances, we include GHG emissions. We show CO_2 emissions from fuel combustion by fuel and sector. Furthermore, we present CO_2

[k] Value added taxes, indirect taxes, duties, environmental taxes and subsidies.

process emissions and other GHG emissions: CH_4, N_2O and F-gases (HFC, PFC and SF_6).

3. Method

Building the Data Base of PIRAMID requires integrating and projecting data from different sources. It is useful to distinguish three types of information in the Data Base:

i. Exogenous inputs
ii. Preprocessed data
iii. Results of MR-GRAS procedure

The exogenous inputs are simply those taken from external sources. Second, the projection of the IOTs may require transforming the exogenous data into the appropriate format. This preprocessed data is usually based on ad hoc assumptions. Finally, the remaining information comes from projecting the IOTs using MR-GRAS procedure. It should be said, that this distinction is not very sharp and might depend on the available information.

In Appendix A we explain how exogenous data is processed before projecting base year IOTs. The remainder of this section summarizes the methodology to project the IOTs using the MR-GRAS procedure.

Table 1: Types of information that can be found in IO tables.

	Macroeconomic assumptions	Energy
1. Exogenous inputs	GDP, population	Total supply and use of fuels, fuels use in specific sectors, energy prices
2. Preprocessed data	Final demand, sectoral production	Bilateral trade of energy goods (physical units)
3. Results of MRGRAS	Bilateral trade flows	Energy use in non-specified sectors, cost structure of fuels production

This section summarizes the methodology used to project the IOTs. First, we argue why the projections are carried out in a Multi-Regional Input-Output (MRIO) setting. Second, we present the method to project most of MRIO components. Finally, the projection of the remaining MRIO components is explained.

The main philosophy of our projection approach is that one has to have the maximum possible control over the updating process. This implies that the practice of full automatization of the whole Multi-Regional Input-Output Tables (MRIOT) projection very often could lead to inferior results, as also our own experience shows. Instead, all the existing information for the framework at hand has to be fully utilized (which is usually different from database to database), while additional objectives/assumptions of the baseline or simulation projections have to be respected. An example of the last requirement in this project is the requirement of having the identical taxes and subsidies rates by sector for the baseline projections.

The fact that sectoral trade data between all regions have to be also estimated, the best way to proceed is using a (sort of) multi-regional input-output framework in the process of projections of the relevant parts of MRIOTs. This is especially true because of the nature of the GTAP data in that it does not represent the detailed sectoral trade of intermediates and final goods separately as is always the case within a full-edged multi-regional input-output setting. Given the importance of using a MRIO setting in the projections, we first present this setting relevant for the current project, which additionally provides a compact, macro overview of the global IOT.

Table 5 illustrates the basic structure of an MRIO framework of the world economy, which for simplicity is assumed to consist of only 3 regions and 3 sectors. It should be noted that it is not a fully-edged multi-regional input-output table since domestic intermediate uses are not separated from the imported uses of intermediate goods, and the trade data do not distinguish between trade in intermediates and in final goods. This feature of the GTAP database is, however, sufficient for CGE modelling purposes. Given this characteristics of the input data, the trade data have been included along the diagonal elements of the off-diagonal block-matrices of the MRIO inter-sectoral transactions table. All the trade data are illustrated in cyan in Table 1, which also includes the margins related to imports and exports (positioned, respectively, in the penultimate row and penultimate column). This again has to do with a specific way of treating international transportation margins in the GTAP model, which allows individual countries to export international transportation services to a "global transport sector" which subsequently satisfies demand for bilateral margins. Since only sector 2 is assumed to be a transport sector, only sector 2 can export transportation services. The sum of the imports margins exactly coincide with the total exports margins at the global level, but not (necessarily) at the levels of individual regions. Along the diagonal blocks of the MRIO table, the yellow blocks represent the total (i.e. domestic and imported) uses of intermediates by domestic sectors along the columns. Finally, the final demand categories (except for exports which are, by definition, part of the trade data) are given in orange, the gross value added (GVA) section in pink, and the taxes and subsidies (TxS) section in lime. The red margins are known exogenous data, corresponding total demand or, equivalently, total supply (Dem/Sup) figures and the global sum of GVA, TxS and final demand components. Uncoloured parts of the table refer to empty cells.

The procedure of projecting IOTs is implemented in two steps, which are described in what follows.

Step 1: Projection of all MRIO components, excluding those of GVA and TxS. This step is implemented using the MR-GRAS method, which allows for imposing aggregation constraints on final demand components (private consumption, government consumption and investment), and on intermediate demands of specific sectors[1] (e.g. energy sectors) as certain intermediate demand components can also be exogenously specified. When final demand vectors are exogenously given, these are not projected within the MR-GRAS framework and are subtracted from the relevant total demand figures. Additionally, the sectoral row sum restrictions need to account for the row sums of the fixed elements representing the sector-specific constraints. This adjustment is necessary because in the projections the fixed elements are nullified within the reference structure matrix and are added later to the main MRIO structure. On the other hand, the column sum for sectors are adjusted by the column-sums of the fixed data related to the sector-specific constraints. The same adjustments need to be done with respect to all the aggregation constraints as well.

To ensure that the total transportation margins on imports and the exports transportation margins at the world level match each other, we: (1) add a negative element (equal to the sum of the benchmark-year exports margins which is, by definition, also equal to the global sum of imports margins in the benchmark table) in the last row and last column of the MRIO table to be projected, corresponding to the crossing-point of these international trade-related transportation margins, (2) set their corresponding row sum to zero, which will give the estimate of the global import transportation margins of non-specific sectors, and (3) set the sum of the exports margin column to the total of the exogenously given imports margins of the specific sectors. It is not difficult to confirm that the last constraint ensures that the global export margins are equal to the sum of the global import margins of non-specific goods and those of specific goods.

Once we impose all the above-mentioned constraints, the trade balance restrictions will be automatically satisfied. That is, the net exports restrictions are, in fact, redundant within the MRIO setting projections. This is a sort of Walras Law, which within the MRIO setting is confirmed by the expenditure-side approach to GDP measurement stating that GDP equals the sum of private consumption, government consumption, investments, and net exports. As the

[1] In this paper, we refer to specific sectors to those in which exogenous assumptions are imposed.

MRIOT is a closed system at the global level, the net exports are projected endogenously and must match the exogenously given figures as long as all the relevant imposed restrictions satisfy the expenditure-side GDP constraint and there are no exogenously imposed inconsistencies between the overall basic structure constraints and the individual sector-specific constraints. For example, the value on total intermediate demand constraints on specific sectors adjusted for the sum of their fixed components cannot be negative by definition, while if there are positive cells to be endogenously estimated within the projection procedure along the specific sectors' columns, then the corresponding adjusted intermediate demand totals must be strictly positive. The same logical requirement applies to the adjusted totals of the final demand categories.

With respect to the question of constraints consistency, we would like to emphasize the following. Given that the sector-specific constraints come from a completely different modelling approach it is of crucial importance that their consistency with the relevant national aggregation constraints be checked. For example, when trade data of energy sectors is obtained from energy models such as POLES and PRIMES, one has to make sure that per region the projections of net exports of energy goods do not contradict the national net exports projections and the structure of net exports of non-energy sectors taken together.

Step 2: Projection of the components of TxS and GVA. One of the assumptions of this procedure is to have constant sectoral rates of taxes and subsidies over time for the baseline projections for all non-specific sectors (recall from the sector-specific constraints that the relevant data are already exogenous). The components of TxS thus have to respect the relevant benchmark rates, which imply the following formulas[m]:

(1) $VAT = \frac{VATrate}{1+VATrate} \times Con = \overline{VATcoef} \times Con,$

(2) $Sub = \overline{SubRate} \times (IntDem + TotGVA)$,

(3) $InTax = \overline{InTxRate} \times (IntDem + TotGVA + Imp - Exp - ExMrg),$

(4) $Duty = \sum_r Duty_r = \sum_r \left[\overline{DutyRate_r} \times \frac{Imp_r}{1+\overline{SubRate_r}+GHGrate_r}\right],$

(5) $TaxGHG = GHGrate \times (IntDem + TotGVA),$

where all the overlined terms refer to the base-year constant rates, and the following abbreviations are used: *VAT* = value added tax, *Con* = private consumption, *Sub* = subsidies, *IntDem* = total intermediate demand, *TotGVA*= total GVA (excluding *TxS*), Imp_r = total imports from the import partner *r*, *Imp* = total imports (i.e. $Imp = \Sigma_r Imp_r$), *InTax* = indirect taxes, *Exp* = total exports, *ExMrg*

[m] These formulas are obtained from the JRC-GEM-E3 model (REFERENCE), which was used for the IOTs projections. However, the projections are not model specific and different formulas could be used.

= exports of international transportation margins, $Duty_r$ = duties on imports obtained from the import partner (region) r, and *GHGrate* is the tax rate of other GHG emissions. Notice that subsidies and GHG taxes are derived in the same way, using the same base of domestic supply. However, the difference is that while subsidies rates are kept constant, while the GHG tax rates are endogenously derived (using equation 5) in order to allow direct incorporation of the GHG tax revenues *TaxGHG* that are exogenously specified for all the projection years. Therefore, all other rates are readily computed using the benchmark data and the relevant formulas given in (1)-(5). Also note from (4) that the base of duties is imports in volumes, where the prices of imports besides domestic prices also account for subsidies and GHG emissions taxes of the exporting regions.

Equation (1) implies that the value added tax (VAT) is straightforward to compute for all the projection years, since *Con* is already known from the first step discussed above. However, given that sectoral totals of GVA excluding taxes (*TotGVA*) are still unknown at this point, one cannot use (2)-(5) to obtain the values of subsidies, indirect taxes and duties. There is, however, a rather simple trick: use the input-side balance of the MRIO setting which states that for each sector the following identity should hold:

(6) $Sup = IntDem + Imp + ImMrg + TotTxS + TotGVA,$

where *ImMrg* stands for imports international transportation margins, and total taxes is, by definition, given by *TotTxS* = *VAT* + *Sub* + *IndTax* + *Duty* + *TaxCO2* + *TaxGHG*.

For the simplicity of the follow-up expressions, let us first introduce the following short-cut vector notations:

(7) $SR \equiv 1 + \overline{SubRate},$
(8) $ITR \equiv 1 + \overline{InTxRate},$
(9) $SITR \equiv 1 + \overline{SubRate} + \overline{InTxRate},$
(10) $DIM \equiv \overline{DutyRate} \cdot IMP,$
(11) $F \equiv Dem - \left[SITR \cdot IntDem + ITR \cdot Imp + ImMrg + VAT + TaxCO2 + TaxGHG - \overline{InTxRate} \cdot (Exp + ExMrg)\right],$

where *IMP* is the import (or trade) matrix consisting of Imp_r of all regions (thus excludes international transportation margins) and has the MRIO setting format as described/illustrated in Table 1 above, *DutyRate* is the matrix of the same structure as *IMP* but includes the benchmark rates of duties, and $\cdot$ denotes element-wise multiplication (Hadamard product). Then it can be easily shown that equations (2)-(11) jointly imply the following system of two (vector) equations:

(12) $TotGVA = \widehat{SITR}^{-1}(F - Duty)$,
(13) $Duty = DIM'\left(\widehat{SR} + \widehat{TaxGHG}\left[\widehat{IntDem} + \widehat{TotGVA}\right]^{-1}\right)^{-1}$,

where $\hat{x}$ refers to a diagonal matrix with the elements of vector x along its main diagonal and zero otherwise and a prime (') denotes matrix transposition. Notice that equation (13) is nothing else as the definition of duties in (4) written in a compact matrix form (using also the GHG taxes equation (5) in place of the GHG tax rates). Equations (12) and (13) are used within a loop format until the values of $TotGVA$ converge for each projection year[n]. Once $TotGVA$ is obtained, the remaining tax-related variables (i.e. subsidies, indirect taxes, duties, and GHG tax rates) are then straightforward to compute using their definitional equations.

Thus, for all the projection years and for all non-specific sectors, the taxes and subsidy rates (except for GHG tax rates) are exactly equal to their benchmark-year counterpart rates. If there is a need to change (some of) these rates, the respective benchmark rates have to be exogenously adjusted.

4. Calibration with PIRAMID

In this section we explain how the projected IOTs could be used as a baseline in a CGE model. We highlight the differences with conventional baseline building in most CGE models and the advantages of the proposed new methodology.

Figure 1 illustrates how baselines are generally built in most CGE models. Firstly, base year IOTs are used to calibrate the parameters of the model. Secondly, exogenous variables such as the endowment of production factors and key parameters such as technical progress are projected over time to meet exogenously determined targets, such as GDP, aggregate energy use, GHGs emissions, etc. Given that most CGE models have a large number of countries/regions and sectors, the number of exogenously determined targets are generally very limited. A high number of targets might lead to

[n] This is the simplest (and possibly most understandable) way of solving (12) and (13) for *TotGVA*; furthermore, the convergence speed is very fast (for any positive initial/starting values of *TotGVA*) requiring maximum up to 5 iterations for a very fine threshold level of 1.e-07. One, of course, can also substitute (13) in (12) and derive the corresponding matrix quadratic equation. This could then be readily solved using the relevant solution involving the use of eigenvectors and eigenvalues, which is arguably a somewhat more complicated approach.

strange/inconsistent parameter values. Consequently, this approach may not be able to reflect structural changes expected in the future. As argued by Ciarli and Savona (2019), CGE models address few aspects of structural changes and may not capture changes in consumption patterns, technology, labour market, etc.

In addition to this, the conventional methodology to build baselines is generally dependent on a specific model. Base year data is used to calibrate parameters which may be model specific. Then, baselines are built projecting key parameter of the model and, therefore, are dependent on the structure of that model. This makes it difficult for other modelling teams to replicate the same baseline.

Figure 1: Conventional baseline building methodology

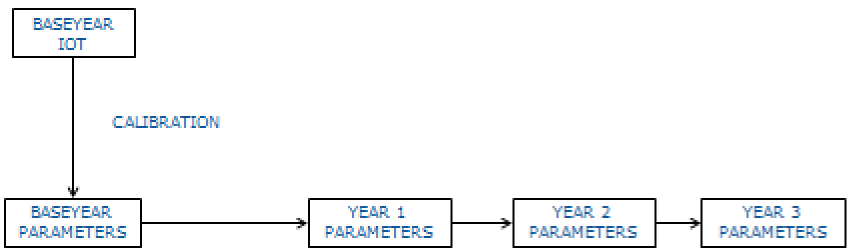
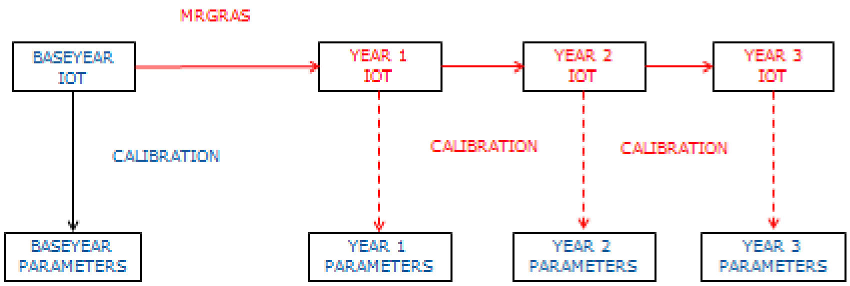


Figure 2 shows the structure of the proposed new methodology to build a baseline. The first step of our approach is to project base year IOTs over time. As explained in section 3, the projection of the IOTs is subject to exogenous constraints. In contrast to the conventional approach, our methodology allows for imposing a large number of exogenously determined targets. These targets can be both macroeconomic and sectoral. When additional information is not available, projected IOTs maintain a structure similar to the base year IOTs. However, PIRAMID is flexible enough to reflect structural changes predicted by exogenous models. Thus, projected IOTs can be subject to an exogenous assumption such as the development of a new technology or structural changes in consumption and trade patterns. Finally, projected IOTs do not have to be model dependent and other modelling teams could use them to replicate the same baseline.

Figure 2: PIRAMID



Once IOTs are projected, they are used to calibrate the parameters of the model for each year. Table 2 show how new methodology differs from conventional one using CES function example.

Table 2: Comparison of CES function (calibrated share form) calibration for two periods (T0 and T1).

	Conventional approach	PIRAMID (new approach)
T0	$\boldsymbol{y_{t0}} = \boldsymbol{\bar{y}_{t0}} \cdot [\sum_i (\boldsymbol{\theta_{i,t0}} \cdot (\frac{\boldsymbol{x_{i,t0}}}{\boldsymbol{\bar{x}_{i,t0}}})^\rho)]^{1/\rho}$	$\boldsymbol{y_{t0}} = \boldsymbol{\bar{y}_{t0}} \cdot [\sum_i (\boldsymbol{\theta_{i,t0}} \cdot (\frac{\boldsymbol{x_{i,t0}}}{\boldsymbol{\bar{x}_{i,t0}}})^\rho)]^{1/\rho}$
T1	$\boldsymbol{y_{t1}} = \boldsymbol{\alpha_{t1}} \cdot \boldsymbol{\bar{y}_{t0}} \cdot [\sum_i (\boldsymbol{\theta_{i,t0}} \cdot (\boldsymbol{\gamma_{i,t1}} \cdot \frac{\boldsymbol{x_{i,t1}}}{\boldsymbol{\bar{x}_{i,t0}}})^\rho)]^{1/\rho}$	$\boldsymbol{y_{t1}} = \boldsymbol{\bar{y}_{t1}} \cdot [\sum_i (\boldsymbol{\theta_{i,t1}} \cdot (\frac{\boldsymbol{x_{i,t1}}}{\boldsymbol{\bar{x}_{i,t1}}})^\rho)]^{1/\rho}$

As you can see, for T0 the calibration looks the same (taking initial Input-Output structure) for both approaches. Then the usual approach for baseline projection in time T1 is to use initial production function and adjust it by $\boldsymbol{\alpha_{t1}}$ and $\boldsymbol{\gamma_{i,t1}}$ that is 'TFP' and factor/inputs productivities respectively, to reflect observed technical progress (the most obvious example is decoupling of energy use and production). There can be two approaches for conventional method. In the first approach, parameter $\boldsymbol{\gamma_{i,t1}}$ is completely exogenous (usually based on some econometric estimation) and therefore production: $\boldsymbol{y_{t1}}$ and costs structure: $\boldsymbol{x_{i,t1}}$ is not known. The second (more sophisticated) approach tries to find $\boldsymbol{\gamma_{i,t1}}$ such that demand for input $\boldsymbol{x_{i,t1}}$ (and index i usually refers to energy goods) is consistent with (in case we want to replicate in the baseline energy/emission path) energy model. But in any case, for most cases, we do not control all relations in CES function as the baseline is in the fact solution of our CGE model in use.

The new approach seems to be much easier (in terms of controlling production levels and costs structure) as in principle we can control $\boldsymbol{y_{t1}}$ and $\boldsymbol{x_{i,t1}}$ (of course to the point we control IO tables we are projecting).

The problematic thing might be only what new parameters (shares) mean in terms of technological progress (and whether we can defend the results as realistic). The easiest way to determine technological improvement is to track evolution of input intensity $\frac{\boldsymbol{x^q_{i,t1}}}{\boldsymbol{y^q_{t1}}}$, where q superscript means that we have interpreted $\boldsymbol{y_{t1}}$ and $\boldsymbol{x_{i,t1}}$ 'values' from Input-Output tables in terms of prices and quantities.

5. Illustrative example of the Data Base

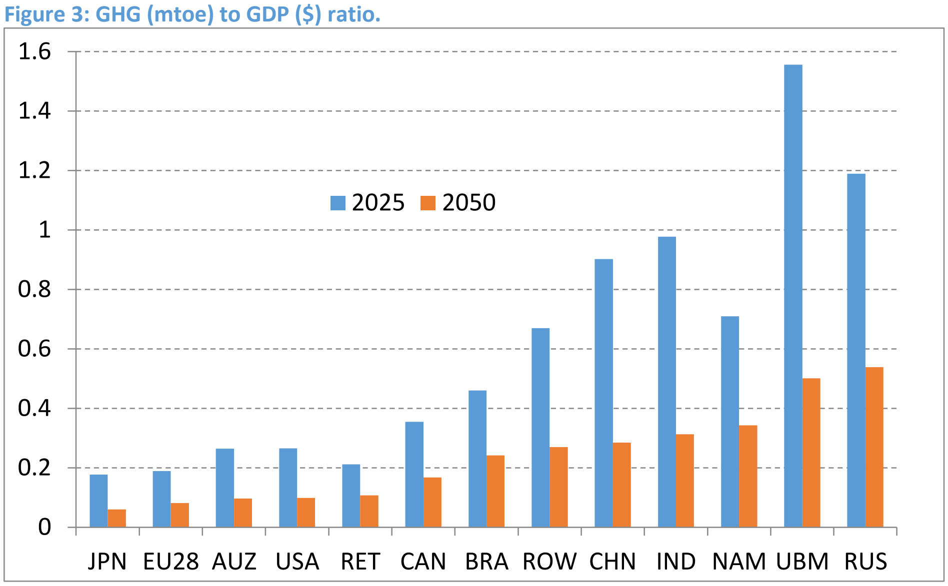
As an example of the practical use of PIRAMID method, we show some results from Global Energy Climate Outlook 2018 report (GECO), written in the JRC Seville. In the report Input-Output tables for 13 regions and 31 sectors are presented (see Appendix for description) in 5 years up to 2050.

Below we are not focusing on the inputs for GECO baseline (as they are taken from exogenous sources) rather we are trying to highlight what are consequences of putting them (from different sources) into one database Input-Output tables. And we believe this is strength of our approach – as one can relatively easily see whether they are mutually consistent.

First graph shows ratio of GHGs per one dollar of GDP in 2025 and 2050. What you can see is that for all regions GHGs intensity declines significantly, but 'developing' regions (starting from Brazil to the right) won't achieve in 2050 EU28 2025 intensity.

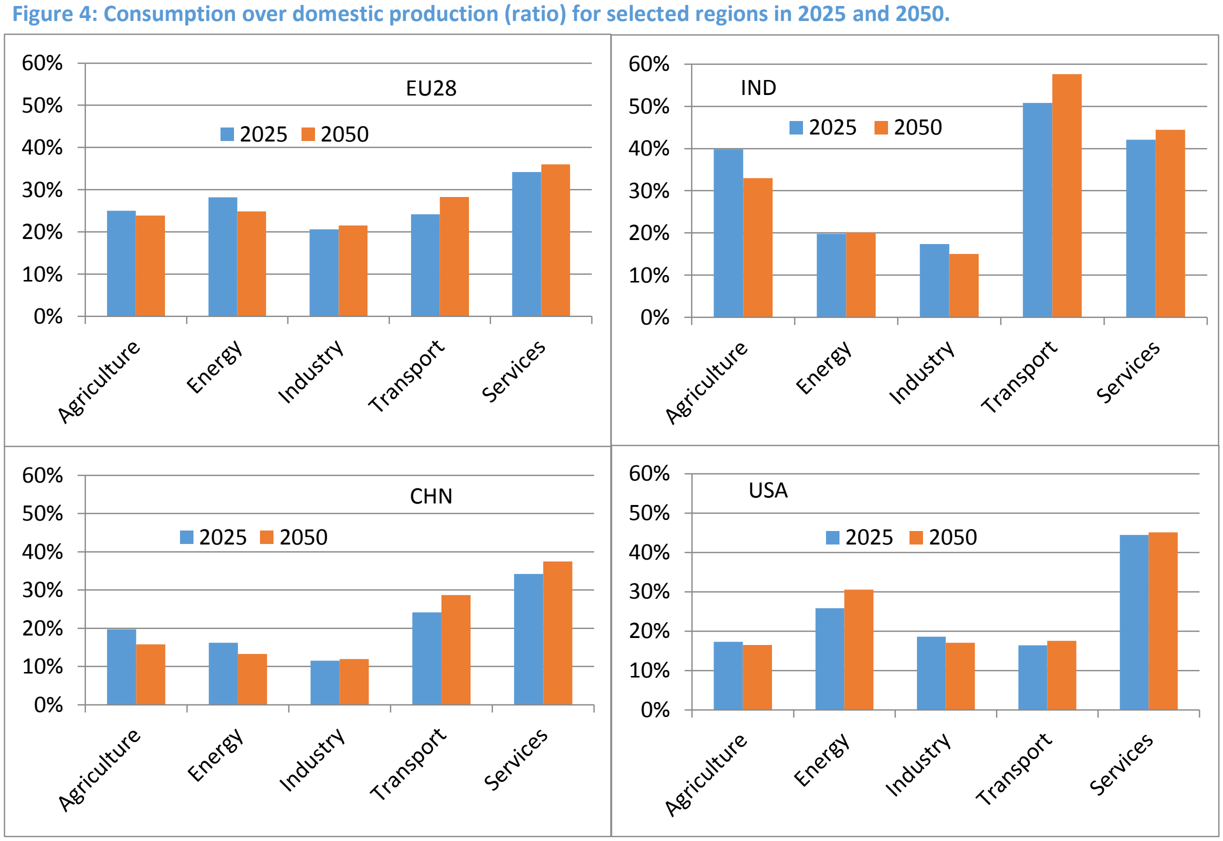
Next example of MRGRAS results, Figure 4, shows consumption to domestic production ratio for selected regions. This relation is driven by many factors: consumption structure over time (exogenous assumption), net trade assumptions, material intensity of economy and other factors. Therefore it is difficult to interpret results or to find general pattern.

Figure 3: GHG (mtoe) to GDP ($) ratio.



Source: Own calculations

Figure 4: Consumption over domestic production (ratio) for selected regions in 2025 and 2050.



Source: Own calculations

6. Conclusions

This paper presents a new method (PIRAMID) to build baselines for CGE models. In PIRAMID, baseline building is carried out in two main steps. In a first step, we project base year IOTs integrating consistently data from different external sources. As a result, we generate a data base composed of 78 IOTs, which represent the economic structure of 13 countries/regions in 6 future years. In a second step, we use the projected IOTs to calibrate the key parameters of the CGE model.

In our opinion, PIRAMID presents several advantages compared to conventional baseline building methods. First, the projected IOTs are consistent with external models and databases; the methodology allows for setting a large number of exogenous assumptions and reflecting potential structural changes. Second, baseline building is based on transparent assumptions which, in case of discrepancy, could be corrected and improved. Third, our baselines, based on the projected IOTs, can be used by other modelling teams.

There are however few caveats in the proposed method:

i. The structure of Input-Output tables is model (JRC-GEM-E3) dependent
ii. There are few elements of IO tables that are not fully controlled. For example we do not control the shares of domestic production and import in total supply
iii. There is a limited flexibility in terms of sectoral and regional aggregation

To overcome those imperfections, we believe that the best way would be to project GTAP database in its original format (as almost all global CGE models are starting with this format). This is work in progress.

Acknowledgements

Write acknowledgements here, including relevant grant numbers (if any). EC disclaimer to be put here!

References

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

Ciarli, T. and Savona, M., 2019. Modelling the evolution of economic structure and climate change: a review. Ecological Economics, 158, pp.51-64.

Golan, A., G. Judge, and S. Robinson. 1994. 'Information from incomplete or partial multisectoral economic data'. Review of Economics and Statistics 76, 541-549.

Golan, A. and S. J. Vogel. 2000. 'Estimation of non-stationary social accounting matrix coefficients with supply-side information'. Economic Systems Research 12, 447-471.

Harthoorn, R. and J. van Dalen. 1987. 'On the adjustment of tables with Lagrange multipliers'. NA-024. Central Bureau of Statistics, The Netherlands, National Accounts Research Division.

Kuroda, M. 1988. 'A method of estimation for the updating transaction matrix in the input-output relationships'. In: K. Uno and S. Shishido (eds.): Statistical Data Bank Systems. Socio-economic database and model building in Japan. Amsterdam: North Holland, pp. 43-56.

Le Treut, G., Ghersi, F., Combet, E. and Lefevre, J. 2014. Construction of hybrid input-output tables for E3 CGE model calibration and consequences on energy policy analysis. In Sustainable Energy Policy and Strategies for Europe, 14th IAEE European Conference, October 28-31, 2014. International Association for Energy Economics.

Leontief, W. W. 1941. The Structure of American Economy, 1919-1929: An Empirical Application of Equilibrium Analysis. Cambridge: Cambridge University Press.

Rey, L., K. Wojtowicz, B. Saveyn, M. Weitzel, T. Vandyck, M. Tamba, U. Temursho. 2019. Baseline GECO 11/2018. European Commission, Joint Research Centre (JRC) [Dataset] PID: http://data.europa.eu/89h/3ffc59a1-edff-491f-8894-3147d2202e42

Temurshoev, U., R. E. Miller, and M. C. Bouwmeester. 2013. 'A note on the GRAS method'. Economic Systems Research 25(3), 361-367.

Stone, R. 1961. Input-Output and National Accounts. Paris: Organization for European Economic Cooperation.

Appendix A. Processing exogenous data

Appendix A1. GDP decomposition

From the expenditure side, GDP is the sum of private consumption (*Con*), government consumption (*Gov*), investment (*Inv*) and net exports (*NetExp*):

$$GDP = Con + Gov + Inv + NetExp$$

We impose aggregate and sectoral values for each GDP component. The assumptions to project GDP components are the following:

- Government consumption: total government consumption is projected from external sources. We disaggregate total consumption at sectoral level based on the base year structure.
- Investment: total investment increases at the same rate as GDP in the following period. For example, if GDP increases by 2% in period t+1, total investment increases by 2% in period t. The sectoral disaggregation is based on capital accumulation.
- Net Exports: we assume that the trade balance of each country/region convergences to zero by 2300. Each sector contributes proportionally to the reduction of the trade surplus/deficit based on the following formula:

$$NetExp_{i,t+1} = NetExp_{i,t} + (NetExp_{t+1} - NetExp_t)\frac{|NetExp_{i,t}|}{\sum_i NetExp_{i,t}}$$

Where $NetExp_{i,t}$ is the net exports of sector i in the period t.

- Private Consumption: total private consumption is the residual, that is, the difference between GDP and government consumption, investment and net exports. By default, sectoral disaggregation of private consumption is based on base year structure. However, in some cases, we may deviate from base year structure. For instance, we might want to assume that the consumption structure of developing countries will be evolving towards the structure of more advanced economies and, thus, the share of agriculture products in total consumption will decrease over time (Engel's law). Similarly, we may want to reflect the vehicle stock projected in energy models (e.g. POLES and PRIMES) and, therefore, adjust the consumption of transport equipment.

Appendix A2. Capital Investment link

We are imposing the pathway of both total investment supply (Inv) and capital compensation (Capital) by sector. We need to make sure the consistency of both pathways and that the law of motion of capital is met over time. Consequently, the following steps are taken:

- First, we assume that the price of capital (*PK*) is constant over time and equal to one. Thus, capital compensation (*Capital*) is equal to capital quantity (*KAV*).
- Second, in order to meet the law of motion of capital, investment demand by sector follows:

$$INVdemQ_i = KAV_{i,t+1} - (1 - decl)KAV_i$$ [o]

Where $INVdemQ_i$ is investment demand in quantities in sector *i* and *decl* is the depreciation rate.

- Third, we adjust the price of investment demand (*PINV*) to make sure that total investment demand is equal to total investment supply :

$$PINV = \frac{\sum_j Inv_j}{\sum_i INVdemQ_i}$$

- Fourth, investment supply in quantities (invQ) is related with investment demand in quantities through an exogenous investment matrix (*tinvpv*)[p].

$$InvQ_j = tinvpv_{ij} INVdemQ_i$$

Appendix A3. Production Structure

In our approach, to project IOTs, we need rows and columns totals. We estimate those using Leontief model, i.e. we derive those from demand side, using well known formula:

$$X = (I - A)^{-1} \cdot FD$$

Where *FD* is final demand vector, *I* – identity matrix, *A* – technology coefficients matrix, and *X* – sectoral production estimates (total supply).

First however we need to calculate *FD* vector. This we do by decomposing its elements (*Con* – private consumption, *Inv* - investment, *Gov* – government consumption and *Exp* - export) as described in Appendix A1. PIRAMID method

---

[o] This can lead to negative investment demand. In those cases $KAV_{i,t+1}$ is increased until $INVV_i$ is zero.

[p] We use the investment matrix of the JRC-GEM-E3 model.

allows for flexibility here, so we can for example set consumption vector that reflects changes in the age of population, or include energy consumption that is consistent with energy model projections.

The important thing is that any demand structure we design; Leontief model should assure that the supply will be enough to satisfy this demand.

By construction, standard Leontief model does not assume technological change (as we are using initial coefficients – matrix *A*), but in principle they can be changed. And for the energy goods there is implicit efficiency improvement defined by the new relation between production and energy use (which is fixed in MRGRAS - see next section).

Appendix A4. Energy and emissions

Energy sectors in our procedure are treated in a different way than other economic sectors. There are two reasons for that:

- We want to track (also) energy balance in volumes (Mtoe) apart from the flows in the monetary terms;
- We want to create energy balance (in volumes) that is consistent with the projections from (external) energy model;

We use energy models (POLES and PRIMES) as a source of information about the supply and demand (at the sectoral level) of fuels and electricity. These data are provided in physical units (Mtoe). What is lacking in the energy model (but is needed for CGE model) is bilateral trade of the energy goods. This we create by finding trade matrix a_{ij} (for each energy good in each period) which is a solution of:

$$min \sum_{i}^{r} \sum_{j}^{s} a_{i,j} * (ln(a_{i,j}) - ln(\hat{a}_{i,j}))$$

$$s.t. \quad \sum_{j}^{s} a_{i,j} + \hat{s}_i = \sum_{j}^{s} a_{j,i} + \hat{u}_i \ , \quad \forall\, i \in r$$

$$and\ s.t. \quad \hat{s}_i \geq \sum_{j}^{s} a_{j,i} \ , \quad \forall\, i \in r$$

where:

$a_{i,j}$ - is an import of country i from country j,

$\hat{s}_i$ - is (given) domestic supply of energy good in country i,

$\hat{u}_i$ - is (given) use of energy good in country i.

The above procedure looks for a trade matrix $a_{i,j}$ that is as close as possible to the initial (GTAP 2011) trade matrix $\hat{a}_{i,j}$, at the same matching r (number of regions) constraints, where each constraint states that country`s supply of energy goods plus import is equal to its use plus export. In case we would want to reflect in the balance some exogenous information about the evolution of trade of energy good, we would need to fix some elements of matrix $a_{i,j}$.

Having determined supply and use (from energy model) and trade matrix (as described above), we have a balance for each energy good in physical terms (Mtoe). MRIOs are tables in monetary units however; therefore our next step is to create energy balances in values. To do so, first we take exogenous assumptions about the prices (costs of production) of energy goods. Then knowing what are the taxes (and subsidies) for different users of energy goods and knowing what is the use of energy goods in physical terms, we calculate the prices (and therefore values) for the different users[q].

As a result, we end with the following elements (in monetary values) of MRIO tables:

- Domestic supply
- Total use
- Trade matrix

These elements are fixed in the MRGRAS procedure.

As a supplementary balance to the economic one (Input-Output tables), we produce a baseline path of GHGs emissions. The major gas is CO_2 from fuels combustion and we derived those emissions from the fuels use and taking an assumption on the emission factor of each fuel (there are some exceptions as there is also non-energy use of fuels, which we should exclude when calculating emissions).

Non-CO_2 emissions (CH_4, N_2O, HFCs and PFCs) are taken from external sources (POLES and GAINS) and mapped into JRC-GEM-E3 sectors.

We also reflect economic flows related with GHGs emissions, i.e. if there is a tax (or permit scheme) in the baseline, we calculate appropriate revenues on the sectoral level and deduct this number from the Value Added defined as in the Section 3 above.

[q] By different users we mean here very broad categories: domestic use, export, import. In fact we know only **total value** of energy use in the economy. This value is then distributed across different sectors in the MRGRAS procedure (unless we fix the use in some sectors).

Appendix A5. GECO 2018 baseline sectoral and regional aggregation

Table 3: Sectoral aggregation of GECO 2018 baseline scenario (GTAP to GECO mapping):

PDR	Crops	PPP	Paper Products
WHT		P_C	Oil
GRO		CRP	Chemical Products
V_F		I_S	Ferrous metals
OSD		NFM	Non-ferrous metals
C_B		FMP	
PFB		MVH	Other Equipment Goods
OCR		OTN	
CTL	Livestock	OME	
OAP		OMF	
RMK		ELE	Electric Goods
WOL		ELY	Electricity supply
FSH		CNS	Construction
FRS	Forestry	OTP	Transport (Land)
OMN	Non-metallic minerals	WTP	Transport (Water)
NMM		ATP	Transport (Air)
CMT	Consumer Goods Industries	CMN	Market Services
OMT		OFI	
VOL		ISR	
MIL		OBS	
PCR		WTR	
SGR		TRD	
OFD		DWE	
B_T		ROS	Non Market Services
TEX		OSG	
WAP		COA	Coal
LEA		OIL	Crude oil
LUM		GAS	Gas
		GDT	

Table 4: Regional aggregation of GECO 2018 baseline scenario (GTAP to GECO mapping)

<table>
<tr><th colspan="3">GTAP</th><th>GECO</th><th>GTAP</th><th>GECO</th><th>GTAP</th><th>GECO</th></tr>
<tr><td>KOR</td><td>PER</td><td>OMN</td><td rowspan="28">ROW</td><td>ZAF</td><td rowspan="10">ROW</td><td>AUT</td><td rowspan="28">EU</td></tr>
<tr><td>MNG</td><td>URY</td><td>QAT</td><td>XSC</td><td>BEL</td></tr>
<tr><td>TWN</td><td>VEN</td><td>BEN</td><td>XTW</td><td>CYP</td></tr>
<tr><td>XEA</td><td>XSM</td><td>BFA</td><td>TUR</td><td>CZE</td></tr>
<tr><td>BRN</td><td>CRI</td><td>CMR</td><td>ARE</td><td>DNK</td></tr>
<tr><td>KHM</td><td>GTM</td><td>CIV</td><td>XWS</td><td>EST</td></tr>
<tr><td>IDN</td><td>HND</td><td>GHA</td><td>EGY</td><td>FIN</td></tr>
<tr><td>LAO</td><td>NIC</td><td>GIN</td><td>MAR</td><td>FRA</td></tr>
<tr><td>MYS</td><td>PAN</td><td>NGA</td><td>TUN</td><td>DEU</td></tr>
<tr><td>PHL</td><td>SLV</td><td>SEN</td><td>XNF</td><td>GRC</td></tr>
<tr><td>SGP</td><td>XCA</td><td>TGO</td><td>IND</td><td>IND</td><td>HUN</td></tr>
<tr><td>THA</td><td>DOM</td><td>XWF</td><td>CAN</td><td>CAN</td><td>IRL</td></tr>
<tr><td>VNM</td><td>JAM</td><td>XCF</td><td>USA</td><td>USA</td><td>ITA</td></tr>
<tr><td>XSE</td><td>PRI</td><td>XAC</td><td>BRA</td><td>BRA</td><td>LVA</td></tr>
<tr><td>BGD</td><td>TTO</td><td>ETH</td><td>AUS</td><td rowspan="2">AUZ</td><td>LTU</td></tr>
<tr><td>XOC</td><td>XCB</td><td>KEN</td><td>NZL</td><td>LUX</td></tr>
<tr><td>NPL</td><td>ALB</td><td>MDG</td><td>CHN</td><td rowspan="2">CHN</td><td>MLT</td></tr>
<tr><td>PAK</td><td>KAZ</td><td>MWI</td><td>HKG</td><td>NLD</td></tr>
<tr><td>LKA</td><td>KGZ</td><td>MUS</td><td>JPN</td><td>JPN</td><td>POL</td></tr>
<tr><td>XSA</td><td>XSU</td><td>MOZ</td><td>RUS</td><td>RUS</td><td>PRT</td></tr>
<tr><td>MEX</td><td>ARM</td><td>RWA</td><td>CHE</td><td rowspan="4">RET</td><td>SVK</td></tr>
<tr><td>XNA</td><td>AZE</td><td>TZA</td><td>NOR</td><td>SVN</td></tr>
<tr><td>ARG</td><td>GEO</td><td>UGA</td><td>XEF</td><td>ESP</td></tr>
<tr><td>BOL</td><td>BHR</td><td>ZMB</td><td>XER</td><td>SWE</td></tr>
<tr><td>CHL</td><td>IRN</td><td>ZWE</td><td>BLR</td><td rowspan="3">UBM</td><td>GBR</td></tr>
<tr><td>COL</td><td>ISR</td><td>XEC</td><td>UKR</td><td>BGR</td></tr>
<tr><td>ECU</td><td>JOR</td><td>BWA</td><td>XEE</td><td>HRV</td></tr>
<tr><td>PRY</td><td>KWT</td><td>NAM</td><td>SAU</td><td>SAU</td><td>ROU</td></tr>
</table>

Table 5: A hypothetical 3-region and 3-sector global (multi-regional) SAM/IOT

Sector	Sec1	Sec2	Sec3	Sec1	Sec2	Sect3	Sec1	Sec2	Sec3	Con	Gov	Inv	ExMrg	Dem
Sec1														
Sec2														
Sec3														
Sec1														
Sec2														
Sec3														
Sec1														
Sec2														
Sec3														
Value added Tax														
Subsidies														
Indirect taxes														
Duties														
Environmental taxes														
Capital														
Skilled labour														
Unskilled labour														
Reserves														
Imports Margins														
Total supply														