

THE ENVIRONMENTAL IMPACT AND SUSTAINABILITY  
APPLIED GENERAL EQUILIBRIUM (ENVISAGE) MODEL

Version 10.01

Dominique van der Mensbrugghe<sup>1</sup>  
The Center for Global Trade Analysis, Purdue University

March 8, 2019

<sup>1</sup> Email: [vandermd@purdue.edu](mailto:vandermd@purdue.edu)

## **Abstract**

This document's main purpose is to provide a full description of the ENVISAGE Model. ENVISAGE is a global recursive dynamic computable general equilibrium model developed originally at the World Bank. This current version reflects a complete re-coding of the model—initially done and supported by the Environment Directorate of the OECD. [Note that the model is still under development and does not yet contain all of the features of the original ENVISAGE model.] ENVISAGE has been developed to assess the interactions between economies and the global environment as affected by human-based emissions of greenhouse gases. At its core, ENVISAGE is a relatively standard recursive dynamic multi-sector multi-region CGE model. It has been complemented by an emissions and climate module [tbd] that links directly economic activities to changes in global mean temperature. And it incorporates a feedback loop that links changes in temperature to impacts on economic variables such as agricultural yields or damages created by sea level rise [tbd]. One of the overall objectives of the development of ENVISAGE has been to provide a greater focus on the economics of climate change for a more detailed set of developing countries as well as greater attention to the potential economic damages. The model remains a work in progress as there are several key features of the economics of climate change that are planned to be incorporated.

# Contents

<b>1</b>	<b>Introduction</b>	<b>1</b>
<b>2</b>	<b>The Envisage Model in a nutshell</b>	<b>4</b>
<b>3</b>	<b>Model specification</b>	<b>7</b>
3.1	Model dimensions . . . . .	7
3.2	Production block . . . . .	9
3.3	Commodity supply . . . . .	21
3.3.1	Non-electric goods . . . . .	21
3.3.2	Domestic supply of electricity . . . . .	22
3.4	Income block . . . . .	24
3.5	Final demand . . . . .	28
3.5.1	Household demand . . . . .	28
3.5.2	Other final demand . . . . .	35
3.6	International trade . . . . .	37
3.6.1	Import demand . . . . .	37
3.6.2	Export supply . . . . .	40
3.6.3	Bilateral trade prices . . . . .	41
3.6.4	Trade margins . . . . .	42
3.7	Goods market equilibrium . . . . .	42
3.8	Factor markets . . . . .	43
3.8.1	Labor market . . . . .	43
3.8.2	Capital market . . . . .	46
3.8.3	Land market . . . . .	48
3.8.4	Market for natural resources . . . . .	52
3.8.5	Water market . . . . .	53
3.8.6	Factor taxes and prices . . . . .	57
3.9	National accounts and model closure . . . . .	57
3.9.1	National accounts . . . . .	57
3.9.2	Closure . . . . .	58
3.9.3	Targeting prices . . . . .	64
3.9.4	Regional and global welfare . . . . .	64
3.10	Emissions . . . . .	66
3.10.1	Emissions . . . . .	66
3.10.2	Emission price regimes . . . . .	68

<b>4</b>	<b>Model dynamics</b>	<b>70</b>
4.1	Labor supply . . . . .	70
4.2	Capital supply . . . . .	71
4.3	Depletion module . . . . .	71
4.4	Knowledge module . . . . .	71
4.5	Technology shifts . . . . .	72
4.6	Preference shifters . . . . .	73
4.6.1	Armington twists . . . . .	73
4.6.2	Introducing cost curves . . . . .	74
4.7	Implementation . . . . .	76
4.8	Diagnostics . . . . .	77
4.8.1	Sectoral investment . . . . .	77
4.9	Exogenous variables . . . . .	77
4.9.1	Factors of production . . . . .	77
4.9.2	Macro closure . . . . .	78
4.9.3	Tax rates . . . . .	78
4.9.4	Technology variables . . . . .	79
4.9.5	Preference variables . . . . .	80
4.9.6	Emission variables . . . . .	80
<b>A</b>	<b>The CES and CET functions</b>	<b>81</b>
A.1	The CES function . . . . .	81
A.1.1	Basic formulas . . . . .	81
A.1.2	Special cases . . . . .	84
A.1.3	Calibration of the CES function . . . . .	86
A.1.4	Comparative statics . . . . .	87
A.1.5	Growth Accounting . . . . .	89
A.1.6	Parameter twists . . . . .	89
A.1.7	Summary . . . . .	92
A.2	The CET Function . . . . .	92
A.2.1	The basic formulation . . . . .	92
A.2.2	Converting to percent differences . . . . .	93
A.2.3	Twists and the CET . . . . .	94
A.3	Modified CES and CET that incorporate additivity . . . . .	94
A.3.1	The CET implementation . . . . .	94
A.3.2	ACET and twists . . . . .	96
A.3.3	The CES implementation . . . . .	96
A.3.4	Using twists with the adjusted CES . . . . .	97
<b>B</b>	<b>The demand systems</b>	<b>100</b>
B.1	The extended linear expenditure system (ELES) . . . . .	100
B.1.1	Basic formulation . . . . .	100
B.1.2	ELES elasticities . . . . .	101
B.1.3	Welfare . . . . .	101
B.1.4	Calibration . . . . .	102
B.2	An Implicitly Direct Additive Demand System (AIDADS) . . . . .	103
B.2.1	Basic formulation . . . . .	104
B.2.2	Elasticities . . . . .	105

B.2.3	Implementation . . . . .	107
B.2.4	Calibration . . . . .	108
B.3	Constant differences in elasticity (CDE) demand system . . . . .	109
B.3.1	General form . . . . .	109
B.3.2	Elasticities . . . . .	110
B.3.3	Calibration of the CDE . . . . .	112
B.3.4	CDE in first differences . . . . .	115
<b>C</b>	<b>Analytcs of dynamics</b>	<b>117</b>
C.1	Capital stock . . . . .	117
C.2	Knowledge stock . . . . .	118
<b>D</b>	<b>Analytical SAM and GTAP correspondence</b>	<b>122</b>
D.1	The analytical SAM . . . . .	122
D.2	Correspondence to GTAP database . . . . .	129
<b>E</b>	<b>Dimensions of the GTAP database, release 9.2</b>	<b>131</b>
<b>F</b>	<b>User Guide</b>	<b>138</b>
F.1	Data preparation facility . . . . .	138
F.1.1	Introduction . . . . .	138
F.1.2	Preliminaries . . . . .	139
F.1.3	Aggregation . . . . .	143
F.1.4	Filtering . . . . .	165
F.1.5	Altering a database . . . . .	168
F.1.6	Integrated command file . . . . .	170
F.2	Model Simulations . . . . .	171
F.2.1	Introduction . . . . .	171
F.2.2	Preliminaries . . . . .	171
F.2.3	Common options . . . . .	176
F.2.4	Comparative static simulations . . . . .	179
F.2.5	Dynamic simulations . . . . .	181
F.3	Post-simulation processing . . . . .	183
F.3.1	User inputs . . . . .	183
F.3.2	Extracting the indicators . . . . .	186
F.3.3	Standard indicators . . . . .	187
F.3.4	Calculating the indicators . . . . .	195
F.3.5	Creating the Excel files . . . . .	195
F.3.6	Creating graphs in R . . . . .	195

# List of Tables

3.1	Sets used in model definition . . . . .	8
3.2	Intermediate CES nests in production . . . . .	12
3.3	Example of a consumer transition matrix . . . . .	28
3.4	Aggregate water bundles . . . . .	54
3.5	Non-greenhouse gas emissions . . . . .	67
4.1	Exogenous factor variables . . . . .	77
4.2	Exogenous macro closure variables . . . . .	78
4.3	Exogenous tax rate . . . . .	79
4.4	Exogenous technology variables . . . . .	79
4.5	Exogenous preferences variables . . . . .	80
4.6	Emission variables . . . . .	80
A.1	Example of capital/labor twist . . . . .	91
A.2	Example of Armington import/domestic twist . . . . .	91
A.3	Key equations for CES implementation . . . . .	92
D.1	Analytical SAM . . . . .	124
D.2	Correspondence between GAMS model variables and GTAP database . . . . .	130
E.1	Regional dimension of GTAP database . . . . .	131
E.-2	Commodity dimension of GTAP database . . . . .	136

# List of Figures

1	Crop production nest . . . . .	11
2	Livestock production nest . . . . .	12
3	Default production nest . . . . .	15
4	KEF bundle nest . . . . .	16
5	Energy bundle nest . . . . .	19
6	CES nest for power bundle . . . . .	23
7	Consumer demand nest . . . . .	29
8	Energy bundle nest in household demand . . . . .	32
9	Labor market closure . . . . .	44
10	Capital supply regimes . . . . .	47
11	CET nest for land allocation . . . . .	50
12	Example of a kinked supply curve using the sigmoid function . . . . .	53
13	CET nest for water allocation . . . . .	55
14	Capital supply as a function of deviations in the rate of return . . . . .	61
15	Examples of cost curves . . . . .	75
16	Gamma distribution examples . . . . .	121
17	Possible directory structure . . . . .	139
18	Directory structure of output data . . . . .	143
19	Simulation flow chart . . . . .	174

# Chapter 1

## Introduction

The purpose of this document is to provide a complete specification of the equations of the ENVIRONMENTAL IMPACT AND SUSTAINABILITY APPLIED GENERAL EQUILIBRIUM (ENVISAGE) MODEL. The ENVISAGE Model is designed to analyze a variety of issues related to the economics of climate change:

- Baseline emissions of CO<sub>2</sub> and other greenhouse gases
- Impacts of climate change on the economy
- Adaptation by economic agents to climate change
- Greenhouse gas mitigation policies—taxes, caps and trade
- The role of land use in future emissions and mitigation
- The distributional consequences of climate change impacts, adaptation and mitigation—at both the national and household level.

ENVISAGE is intended to be flexible in terms of its dimensions. The core database—that includes energy volumes and CO<sub>2</sub> emissions—is the GTAP database, currently version 9.0 with a 2011 base year (see [Aguilar et al. \(2016\)](#)). The latter divides the world into 140 countries and regions, of which 120 are countries and the other region-based aggregations.<sup>1</sup> The database divides global production into 57 sectors—with extensive details for agriculture and food and energy (coal mining, crude oil production, natural gas production, refined oil, electricity, and distributed natural gas). Appendix E provides more detail. Due to numerical and algorithmic constraints, a typical model is limited to some 20-30 sectors and 20-30 regions.

This document describes the current version of ENVISAGE, which is still in a developmental stage. This current version includes the following:

- Capital vintage production technology that permits analysis of the flexibility of economies
- Two-level decomposition of domestic absorption into demand for domestic production and bilateral imports. This ranges from a simple version with a national agent optimizing each stage, to a full MRIO version with bilateral sourcing at the agent level.
- Partially endogenous technical change [tbd]

---

<sup>1</sup> The countries defined in GTAP cover well over 90 percent of global GDP and population.



- A detailed specification of energy demand in each economy
- Incorporation of a limited set of new energy technologies
- A resource depletion module for coal, oil and natural gas [tbd]
- The ability to introduce future alternative energy (or backstop) technologies [tbd]
- CO<sub>2</sub> emissions that are fuel and demand specific
- Incorporation of the main Kyoto greenhouse gases (methane, nitrous oxide and the fluoridated gases)
- A flexible system for incorporating any combination of carbon taxes, emission caps and tradable permits
- A simplified climate module that links greenhouse gas emissions to atmospheric concentrations combined with a carbon cycle that leads to radiative forcing and temperature changes. [tbd]
- A water module that incorporates water use for irrigation by crop, and aggregate water demand in other parts of the economy.
- [Added 21-Apr-2018] Government to government transfers have been incorporated. As more small developing countries are being included in the GTAP database and models, a more detailed accounting of the balance of payments is important to reflect macro balances.
- [Added 01-Dec-2018] The trade module now includes the *ad valorem* equivalent (AVE) of non-tariff measures (NTMs). Unlike tariff revenues, the revenues generated by the NTMs can be allocated to private households and government—either domestic or foreign.
- [Added 01-Dec-2018] The make module has a new technology parameter that can be used to capture changes in post-harvest (or production) losses.
- [Added 01-Dec-2018] Household demand includes a new waste module, with the possibility of taxing household waste.
- [Added 20-Jan-2019] A research and development (R&D) module has been added. R&D expenditures add to the knowledge stock, however with a distributed lag structure. The knowledge stock influences an endogenous component of labor productivity growth.
- [Added 20-Jan-2019] A global social welfare function has been introduced. This is still provisional.

The current work program includes the following tasks:

- Addition of marginal abatement cost curves for the non-CO<sub>2</sub> gases
- Adding a more detailed land-use module

The ENVISAGE Model is a descendant of a family of models that originated at the OECD in the late 1980s and early 1990s, WALRAS, RUNS and GREEN.<sup>2</sup> ENVISAGE was initially developed

---

<sup>2</sup> OECD (1989/1990), Burniaux (1987), Burniaux and van der Mensbrugghe (1994), Burniaux et al. (1992), and van der Mensbrugghe (1994). Note that WALRAS and RUNS were initially developed at Stanford University and the Université Libre de Bruxelles, respectively.

at the World Bank in 2007 and was a re-coded version of the World Bank’s LINKAGE model ([van der Mensbrugghe \(2011\)](#)), which had a trade focus. It is designed specifically to analyze climate change issues and thus incorporates a more developed energy sector, a climate module (that makes integrated assessment an option), and climate change impact feedbacks. ENVISAGE is coded using the GAMS/MCP package.

The current (draft) version is a complete re-coding of the original ENVISAGE Model largely undertaken at the behest of the OECD’s Environment Directorate, whose ENV-Linkages model also descended from the above-mentioned OECD models. [The current version has yet to incorporate all of the features of the original ENVISAGE Model, but work is ongoing to complete the task.] This version, developed and residing at the Center for Global Trade Analysis (GTAP), thus shares a core code with the OECD ENV-Linkages model, though further development of the individual models will diverge some over time due to differing priorities and preferences. The version developed for the OECD, ENV-Linkages, is a relatively faithful translation of their original specification ([Chateau et al. \(2014\)](#)), but with a clean interface, streamlined code and improvements in its numerical properties. The water module was developed specifically for the OECD. Note that the ENV-Linkages team at the OECD was a core contributor to the quantification of the Shared Socio-Economic Pathways that will frame the analysis of the economics of climate change for at least the next decade ([Dellink et al. \(2017\)](#)).

The following items have been added to this version of the model compared to the previous:

- A new capital flow module has been implemented, but is still undergoing testing and should be used with appropriate caution. It is inspired by the domestic capital allocation specification of the USAGE Model ([Dixon and Rimmer \(2005\)](#)).
- Preference shifters have been added to the power bundles.
- The model interface has been revamped to couple with the new GTAP database format. The latter is described in more detail in [Corong et al. \(2017\)](#). The standard GTAP model has undergone its first major facelift in over a decade and this has entailed a number of changes to the database—though mostly in terms of naming conventions as opposed to changes in functionality. The two most important changes include the introduction of a ‘make’ matrix, i.e. the explicit differentiation of activities from commodities; and the conversion of all taxes to be measured relative to base prices—in other words, all taxes are positive and subsidies negative. This affects output and export taxes.
- To hone more closely to the new GTAP standard model, two sets of taxes have been modified. The first relates to output taxes. These are now identified by two indices—source activity and supplied commodity. The current database, in the absence of additional tax information, assumes these are uniform across source activities, though need not be after aggregation. The second refers to taxes on factor income. These are now identified by source activity as well as by receiving factor. Thus taxes on capital (or any other factor) can differ depending on the activity where the factor is used.
- An MRIO version of the trade specification has been implemented. It has yet to be connected to an MRIO database. This allows for agent-based demand for imports by region of origin. In the standard model, a representative agent allocates import demand across regions of origin.

The next section provides a brief summary of the main model features. This is followed by a much longer section with a full-blown description of the ENVISAGE Model.

## Chapter 2

# The Envisage Model in a nutshell

The ENVISAGE Model at its core is a recursive dynamic and global computable general equilibrium model (CGE). It follows the circular flow of an economy paradigm. Firms purchase input factors (for example labor and capital) to produce goods and services. Households receive the factor income and in turn demand the goods and services produced by firms. And equality of supply and demand determine equilibrium prices for factors, goods and services. The model is solved as a sequence of comparative static equilibria where the factors of production are exogenous for each time period and linked between time periods with accumulation expressions.

Production is implemented as a series of nested constant-elasticity-of-substitution (CES) functions the aim of which is to capture the substitutability across all inputs. Three production archetypes are implemented. The first is for crops that reflects intensification of inputs versus land extensification. The second is for livestock that reflects range-fed versus ranch-fed production. The final, also referred to as the default, revolves largely around capital/labor substitutability. Some production activities highlight specific inputs (for example agricultural chemicals in crops and feed in livestock) and all activities include energy and its components as part of the cost minimization paradigm. Production is also identified by vintage—divided into *Old* and *New*—with typically lower substitution possibilities associated with *Old* capital.

Each production activity is allowed to produce more than one commodity—for example the ethanol sector can produce ethanol and distiller’s dried grains with solubles (DDGS). And commodities can be formed by the output of one or more activities (for example electricity). ENVISAGE therefore uses a different classification of activities and commodities.<sup>1</sup> One of the features of the model is that it integrates the new GTAP power data base that disaggregates GTAP’s electricity sector (‘ely’) into 11 different power sources plus electricity transmission and distribution. Though the database has both the supply and demand side for all 11 power sources, the aggregation facility permits the aggregation of electricity demand into a single commodity and the ‘make’ matrix specification combines the output from the different power activities into a single electricity commodity.

Income accrues from payments to factors of production and is allocated to households (after taxes). The government sector accrues all net tax payments and purchases goods and services. The model incorporates multiple utility functions for determining household demand. There is a set of three household demand functions linked to the ubiquitous linear expenditure system (LES): the standard LES, the extended LES (ELES) that incorporates household saving into the utility function, and ‘an implicitly directly additive demand system’ (AIDADS), that allows for non-linear Engel curves in the LES framework.<sup>2</sup> The fourth option uses the constant differences in elasticity

<sup>1</sup> Production activities are indexed with  $a$  and commodities are indexed with  $i$ .

<sup>2</sup> Users can also specify implementing a Cobb-Douglas (CD) utility function, which can be considered part of the

(CDE) utility function that is used in the core GTAP model (Hertel (1997) and Corong et al. (2017)). The ELES incorporates the decision to save in a top level utility function. The other demand systems assume savings is an exogenous proportion of disposable income in the default closure. The consumer utility function determines consumer demand bundles that are subsequently converted to produced goods using a consumer demand 'make' or transition matrix. Investment is savings driven and equal to domestic saving adjusted by net capital flows.

Trade is modeled using the so-called Armington specification that posits that demand for goods are differentiated by region of origin. The model allows for domestic/import sourcing at the aggregate level (after aggregating domestic absorption across all agents), or at the agent-level. In the standard specification, a second Armington nest allocates aggregate import demand across all exporting regions using a representative agent specification. Note that a newer, though minimally tested version, allows for sourcing imports by agent—also known as the MRIO specification. Exports are modeled in an analogous fashion using a nested constant-elasticity-of-transformation (CET) specification. The domestic supply of each commodity is supplied to the domestic market and an aggregate export bundle using a top-level CET function. The latter is allocated across regions of destination using a second-level CET function.<sup>3</sup> Each bilateral trade node is associated with four prices: 1) the producer price; 2) the export border price, also referred to as the free-on-board (FOB) price; 3) the import border price, also referred to as the cost, insurance and freight (CIF) price; and 4) the end-user price that includes all applicable trade taxes. The wedge between the producer price and the FOB price is represented by the export tax (or subsidy if negative) and the wedge between the CIF and end-user prices represents the import tariff (and perhaps other import related distortions). The wedge between the CIF and FOB prices represents the international trade and transport margins. These margins represent the use of real resources that are supplied by each region. The global international trade and transport sector purchases these services from each region so as to minimize the aggregate cost.

The model has two fundamental markets for goods and services. Domestically produced goods sold on the domestic market, and domestically produced goods sold by region of destination. All other goods and services are composite bundles of these goods. Two market equilibrium conditions are needed to clear these two markets.<sup>4</sup>

The model incorporates five types of production factors: 1) labor (of which there can be up to 5 types); 2) capital; 3) land; 4) a sector specific natural resource (such as fossil fuel energy reserves); and 5) water. The labor market is allowed to be segmented (though not required). The model allows for regime switching between full and partial wage flexibility. Capital is allocated across sectors so as to equalize rates of returns. If all sectors are expanding, *Old* capital is assumed to receive the economy-wide rate of return. In contracting sectors, *Old* capital is sold on secondary markets using an upward sloping supply curve. This implies that capital is only partially mobile across sectors. Aggregate land and water supply are specified using supply curves. Though there are several options, the preferred supply curve is a logistic function that has an upper bound. Water demand also includes exogenous components for environmental uses and groundwater recharge. Land and water are allocated across activities using a nested CET specification.<sup>5</sup> Natural resources are supplied to each sector using an iso-elastic supply function with the possibility differentiated

---

LES framework.

<sup>3</sup> The model allows for perfect transformation, which is the standard specification in the GTAP model.

<sup>4</sup> If there are  $N$  commodities and  $R$  regions, there will be  $R \times N$  market clearing conditions for domestic goods and  $R \times N \times R$  market clearing conditions for bilateral trade.

<sup>5</sup> Land is only implemented for agricultural activities. Water demand by activity is only present in irrigated crop sectors. Other water demand is based on aggregate demand functions with market clearing, but is not part of the cost structure.

elasticities depending on market conditions.

ENVISAGE incorporates the main greenhouse gases—carbon, methane, nitrous oxides and fluoridated gases. Emissions are generated by consumption of commodities (such as fuels), factor use (for example land in rice production and herds in livestock production) and there are also processed base emissions such as methane from landfills.<sup>6</sup> A number of carbon control regimes are available in the model. Carbon taxes can be imposed exogenously—potentially differentiated across regions. The incidence of the carbon tax allows for partial or full exemption by commodity and end-user. For example households can be exempted from the carbon tax on natural gas consumption. The model allows for emission caps in a flexible manner—where regions can be segmented into coalitions on a multi-regional or global basis. In addition to the standard cap system, a cap and trade system can be defined where each region within a coalition is assigned an initial emission quota.

Dynamics involves three elements. Labor supply (by skill level) grows at an exogenously determined rate. The aggregate capital supply evolves according to the standard stock/flow motion equation, i.e. the capital stock at the beginning of each period is equal to the previous period's capital stock, less depreciation, plus the previous period's level of investment. The third element is technological change. The standard version of the model assumes labor augmenting technical change—calibrated to given assumptions about GDP growth and inter-sectoral productivity differences. A portion of labor productivity growth is linked to R&D expenditures. In policy simulations, technology is typically assumed to be fixed at the calibrated levels, though can be influenced by changes in R&D expenditures.

---

<sup>6</sup> The current version of the model does not include carbon emissions from deforestation—an important source of global carbon emissions.

## Chapter 3

# Model specification

The next section describes the key model dimensions. Subsequent sections describe the various blocks or modules of the model using the traditional circular flow scheme of economics, i.e. starting with production and factor incomes, income distribution, demand, trade, and macro closure. At the end, there is a discussion on the model dynamics.

### 3.1 Model dimensions

The model specification in GAMS is mostly as general as possible, i.e. virtually all variables and equations are specified using set-based indices that are defined at run-time when the data is read in. The model can thus be configured for any number of regions (up to 140 with the latest GTAP release) and any number of sectors (up to 57 with the standard GTAP database and 76 with the power/water database).

Table 3.1 lists the main indices used by the model. On the demand side the key dimension is the number of Armington agents (indexed by  $aa$ ). The Armington agents include all production activities (indexed by  $a$ ), and final demand agents (indexed by  $fd$ ). The latter are split into households ( $h$ ), government ( $gov$ ) and investment ( $gov$ ). Activities, i.e. production sectors, are separated from commodities. The former are indexed by  $a$  and the latter by  $i$  (with a possible alias of  $j$ ). Thus intermediate demand,  $XA_{i,a}$  represents the demand for commodity  $i$  by activity  $a$ .<sup>1</sup> The ENVISAGE model allows for a non-diagonal make matrix that is user-determined at the time of aggregating the GTAP database. This has proven useful, for example, for the power sector. The new power database has 11 power generating technologies with a complementary transmission and distribution activity. These can be aggregated to a single electricity commodity using the make matrix (with a nested CES specification).

The model incorporates three different production structures—crops, livestock and other. These will be indexed by  $acr$ ,  $alv$  and  $ax$ , respectively—all subsets of  $a$ . Activities are also sub-divided into two 'geographic' zones—rural and urban. Typically, agricultural activities are assigned to the rural zone and all other activities to the urban zone. The model allows for segmented labor markets by geographic zone with a Harris-Todaro-type migration function (Harris and Todaro (1970)). Most of the other indices are relatively self-explanatory and will be further elaborated on in the model description.

Three key indices are not part of the table—an index for power ( $pb$ ), land ( $lb$ ) and water bundles ( $wbnd$ ). These indices are user-defined and are used to aggregate power supply in the

---

<sup>1</sup> In the 'classic' GTAP database there is a one-to-one mapping between activities and commodities. The new 'standard' GTAP database differentiates activities and commodities.

power module, allocate land across sectors in the land supply module and allocate water across uses in the water supply module. They will be further described in their respective sections.

Table 3.1: **Sets used in model definition**

Set	Description
<i>aa</i>	Armington agents
<i>a</i>	Activities (a subset of <i>aa</i> )
<i>acr</i>	Crop activities (a subset of <i>a</i> )
<i>alv</i>	Livestock activities (a subset of <i>a</i> )
<i>ax</i>	All other activities (a subset of <i>a</i> )
<i>elya</i>	Power activities (a subset of <i>a</i> )
<i>z</i>	Set of zones—just rural ( <b>rur</b> ) and urban ( <b>urb</b> )
<i>i</i>	Produced (or supplied) goods
<i>inum</i>	Set of manufacturing sectors (subset of <i>i</i> , used in definition of numéraire)
<i>fp</i>	Factors of production
<i>l</i>	Labor categories (subset of <i>fp</i> )
<i>ul</i>	Unskilled labor categories (subset of <i>l</i> )
<i>sl</i>	Skilled labor categories (subset of <i>l</i> )
<i>cap</i>	Capital account (subset of <i>fp</i> )
<i>lnd</i>	Land account (subset of <i>fp</i> )
<i>nrs</i>	Natural resource account (subset of <i>fp</i> )
<i>wat</i>	Water account (subset of <i>fp</i> )
<i>k</i>	Consumed commodities
<i>nrg(k)</i>	The energy bundle in consumed commodities
<i>fd</i>	Final demand (subset of <i>aa</i> )
<i>fdc</i>	Final demand excluding households (subset of <i>fd</i> )
<i>h</i>	Households (subset of <i>fd</i> )
<i>gov</i>	Government account (subset of <i>fd</i> )
<i>inv</i>	Investment account (subset of <i>fd</i> )
<i>gy</i>	Government revenue accounts
<i>itax</i>	Indirect taxes (subset of <i>gy</i> )
<i>ptax</i>	Production tax account (subset of <i>gy</i> )
<i>mtax</i>	Import tariff account (subset of <i>gy</i> )
<i>etax</i>	Export tax account (subset of <i>gy</i> )
<i>vtax</i>	Tax on factors of production account (subset of <i>gy</i> )
<i>ptax</i>	Production tax account (subset of <i>gy</i> )
<i>ctax</i>	Carbon tax (subset of <i>gy</i> )
<i>dtax</i>	Direct taxes (subset of <i>gy</i> )
<i>r</i>	Regions
<i>s, d</i>	Aliases with <i>r</i> (for <i>source</i> and <i>destination</i> regions)
<i>rnum</i>	Set of regions used in definition of numéraire (subset of <i>r</i> )
<i>rres</i>	Residual region (subset of <i>r</i> , must be of single dimension)
<i>em</i>	Emission types



## 3.2 Production block

Production is implemented using a nested CES structure—rather standard in many applied general equilibrium models.<sup>2</sup> Three production structure prototypes are introduced—crops, livestock and all other (the default production structure). Figures 1, 2 and 3 depict the production structures for each of the three prototypes. These figures are supplemented by the nesting for the so-called *KEF* and energy bundles in respectively figures 4 and 5. This section will provide a full description of all the nests starting from the top nest. At the end of each of the terminal nodes will be the derived demand for the basic components (i.e. inputs) of production—intermediate goods and factors of production.

Each nest is re-produced for each of the possible vintages. In the comparative static version of the model there is usually only a single vintage. The default in the dynamic version of the model is to have two vintages—so-called *Old* vintage (or installed capital), and *New* that represents the most recent supply of capital. The former is meant to be only partially mobile across sectors, whereas the latter is fully mobile—detailed further below.

Production for each vintage is thus associated with a unit (or marginal) cost of production, represented by  $uc$ . The post tax unit cost of production,  $PXv$ , is equal to the tax-adjusted pre-tax unit cost of production, equation (P-2), where  $\tau^{uc}$  is the tax on the cost of production. The aggregate unit cost of production across vintages,  $PX$ , is given in equation (P-1) and is the weighted sum of the vintage-specific costs of production, with weights given by the production volume shares where  $XPv$  and  $XP$  represent respectively output by vintage and aggregate output.<sup>3</sup>

$$PX_{r,a}XP_{r,a} = \sum_v PXv_{r,a,v}XPv_{r,a,v} \quad (\text{P-1})$$

$$PXv_{r,a,v} = UC_{r,a,v} (1 + \tau_{r,a,v}^{uc}) \quad (\text{P-2})$$

The remainder of this section describes the various CES nests that represent the production structure by vintage. The purpose of these CES nests is to replicate the substitution and complementarity relations across all of the inputs. The top level CES nest represents the combination of output,  $XPX$ , with a bundle of non-CO<sub>2</sub> greenhouse gas (GHG) emissions,  $XGHG$ . The latter is a special bundle that is used to simulate the marginal cost of abating non-CO<sub>2</sub> greenhouse gases. As the price of this bundle increases (e.g. from a tax on emissions), producers substitute away from this relatively more costly input.

Equations (P-3) and (P-4) represent the derived demands for the output and GHG bundles respectively. These are the standard CES demand functions where  $PXP$  and  $PXGHG$  represent the price of the component bundles and  $UC$  is the price of the composite (or aggregate bundle). In this case,  $UC$ , is the unit cost of production, including the price associated with the GHG emissions. The parameters  $\alpha^{xp}$  and  $\alpha^{ghg}$  represent the standard CES share parameters,  $A^{xpv}$  is a tech-neutral shift in the production nest and  $\sigma^{xp}$  is the elasticity of substitution between production and the emissions of GHGs. The production nest also allows for input specific technological change as represented by the parameters  $\lambda^{xp}$  and  $\lambda^{ghg}$ . These are typically exogenous. Equation (P-5) defines the component price of the CES bundle,  $UC$ . The expression uses the CES dual price formula and

<sup>2</sup> The main alternative is to implement some form of a single-level flexible functional form such as the translog (see for example Jorgenson et al. (2013)).

<sup>3</sup> An earlier version of the model had an output price that was equal to the unit cost of production plus a (fixed) markup and adjusted by a production tax. The output tax has been moved to the supply side, see below, and is now both activity- and commodity specific. The markup has been eliminated for the moment.



could be replaced by the zero-profit condition.<sup>4</sup>

$$XPX_{r,a,v} = \alpha_{r,a,v}^{xp} (A_{r,a,v}^{xpv} \lambda_{r,a,v}^{xp})^{\sigma_{r,a,v}^{xp}-1} \left( \frac{UC_{r,a,v}}{PXP_{r,a,v}} \right)^{\sigma_{r,a,v}^{xp}} XPv_{r,a,v} \quad (P-3)$$

$$XGHG_{r,a,v} = \alpha_{r,a,v}^{ghg} (A_{r,a,v}^{xpv} \lambda_{r,a,v}^{ghg})^{\sigma_{r,a,v}^{xp}-1} \left( \frac{UC_{r,a,v}}{PXGHG_{r,a,v}} \right)^{\sigma_{r,a,v}^{xp}} XPv_{r,a,v} \quad (P-4)$$

$$UC_{r,a,v} = \frac{1}{A_{r,a,v}^{xpv}} \left[ \alpha_{r,a,v}^{xp} \left( \frac{PXP_{r,a,v}}{\lambda_{r,a,v}^{xp}} \right)^{1-\sigma_{r,a,v}^{xp}} + \alpha_{r,a,v}^{ghg} \left( \frac{PXGHG_{r,a,v}}{\lambda_{r,a,v}^{ghg}} \right)^{1-\sigma_{r,a,v}^{xp}} \right]^{\frac{1}{1-\sigma_{r,a,v}^{xp}}} \quad (P-5)$$

The second level nest decomposes aggregate production net of the GHG bundle into two bundles,  $ND^1$  and  $VA$ . The first bundle includes all intermediate goods with the exception of energy goods and other intermediate goods that are treated specially in a given activity. In the default configuration of the model, activity-specific intermediate goods include fertilizers for crop activities and feed for livestock activities.<sup>5</sup> The  $VA$  contains all factors of production, the energy goods and activity-specific goods where applicable. Equation (P-6) determines the demand for the top level intermediate demand bundle,  $ND^1$ . Equation (P-7) determines the demand for the  $VA$  bundle. The bundle prices are respectively  $PND^1$  and  $PVA$  and the substitution elasticity is given by  $\sigma^p$ . It is worth noting that the equation for  $ND^1$  is summed over all vintages. This is because the further decomposition of the  $ND^1$  bundle is assumed to be independent of the vintage, whereas the decomposition of the  $VA$  is vintage specific as the substitution elasticities further down the nest are allowed to vary by vintage. Equation (P-8) determines the price of  $XPX$ ,  $PXP$ .

$$ND_{r,a}^1 = \sum_v \alpha_{r,a,v}^{nd1} \left( \frac{PXP_{r,a,v}}{PND_{r,a}^1} \right)^{\sigma_{r,a,v}^p} XPX_{r,a,v} \quad (P-6)$$

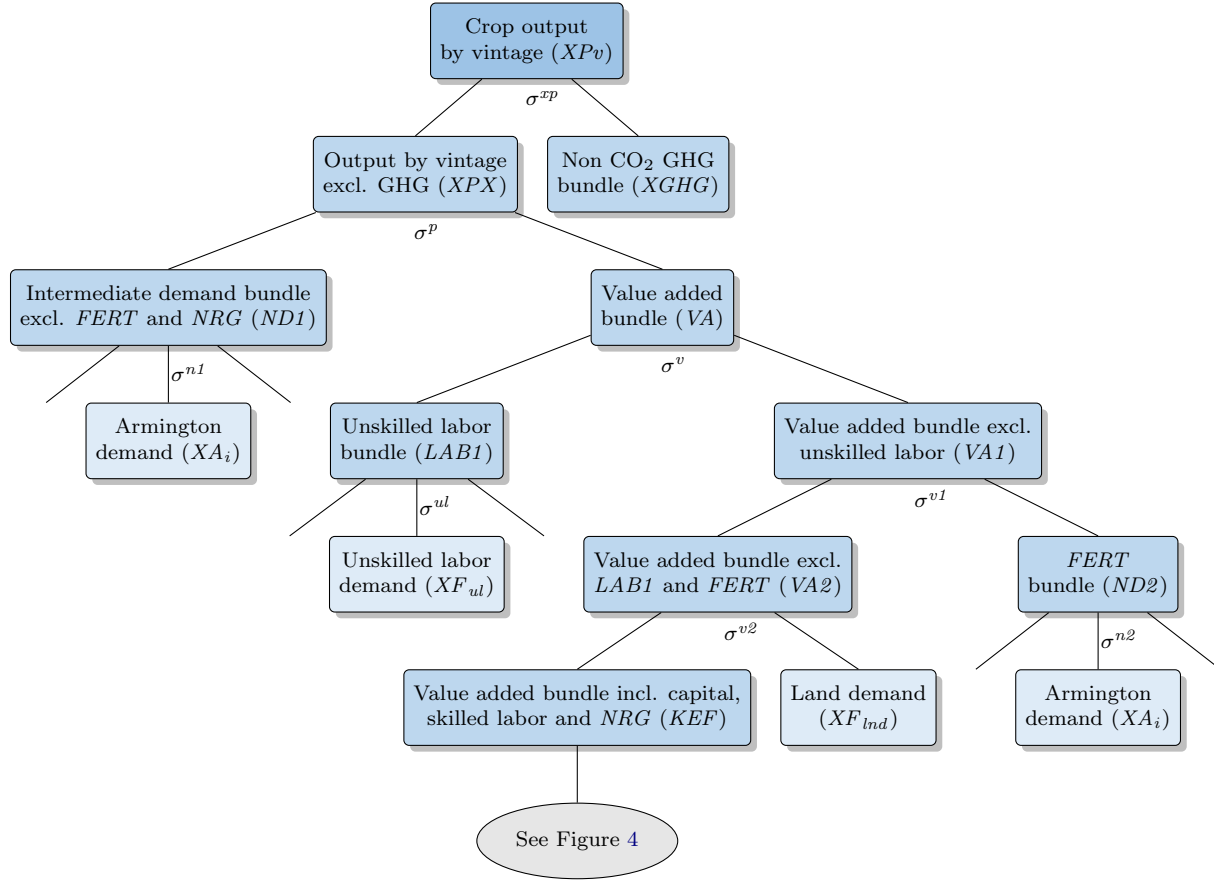
$$VA_{r,a,v} = \alpha_{r,a,v}^{va} \left( \frac{PXP_{r,a,v}}{PVA_{r,a,v}} \right)^{\sigma_{r,a,v}^p} XPX_{r,a,v} \quad (P-7)$$

$$PXP_{r,a,v} = \left[ \alpha_{r,a,v}^{nd1} (PND_{r,a}^1)^{1-\sigma_{r,a,v}^p} + \alpha_{r,a,v}^{va} (PVA_{r,a,v})^{1-\sigma_{r,a,v}^p} \right]^{\frac{1}{1-\sigma_{r,a,v}^p}} \quad (P-8)$$

The middle CES nests are activity specific and reflect the three production structure prototypes described above—crops, livestock and all other activities. The crop production structure is designed to capture production characterized by the choice between intensification versus extensification, i.e. where land is abundant and cheap, production will tend to expand using more land, and vice versa if land is scarce and expensive. Livestock production is characterized by feed versus land substitution. In the default production structure, the main characteristic of production is the standard capital and labor substitution.

<sup>4</sup> The zero profit condition (without indices) is  $UC.XPv = PXP.XPX + PXGHG.XGHG$ .

<sup>5</sup> The standard GTAP database does not include fertilizers. These are associated with the 'crp' good, i.e. chemicals, rubber, and plastics. The standard feed goods are wheat, coarse grains and oil seeds, respectively 'wht', 'gro' and 'osd' in the GTAP database.


 Figure 1: **Crop production nest**

These different structures will be captured by a set of intermediate bundles, represented by  $VA^1$  and  $VA^2$  that represent different composition of factors and activity-specific intermediate goods where appropriate. The top of the nest is represented by the bundle  $VA^1$ . The bottom of this set of nests are the bundles  $LAB^1$ ,  $XF_{Ind}$ ,  $KEF$  and  $ND^2$ , where appropriate. The bundle  $LAB^1$  represents the demand for the unskilled labor bundle.<sup>6</sup> The variable  $XF_{Ind}$  represents the activity's demand for the land factor. The bundle  $KEF$  represents the nested combination of capital, skilled labor, energy and the natural resource factor, where appropriate. The bundle  $ND^2$  represents the activity-specific demand for intermediates—fertilizers in the case of crops and feed in the case of livestock activities. The decomposition of these three bundles<sup>7</sup> is the same for all activities and described below. Table 3.2 summarily describes the composition of these middle nests for the three production prototypes.

Figures 1, 2 and 3 provide an illustrative description of the production nests.

<sup>6</sup> The user determines what labor types are considered 'unskilled' labor. The user can decide to have all labor types in this bundle, in which case the skilled labor bundle,  $LAB^2$ , will be empty.

<sup>7</sup>  $XF_{Ind}$  is final factor.

Table 3.2: Intermediate CES nests in production

Activity	Bundle composition
<b>Crops</b>	
VA	$CES (LAB^1, VA^1)$
VA <sup>1</sup>	$CES (ND^2, VA^2)$
VA <sup>2</sup>	$CES (XF_{lnd}, KEF)$
<b>Livestock</b>	
VA	$CES (VA^1, VA^2)$
VA <sup>1</sup>	$CES (LAB^1, KEF)$
VA <sup>2</sup>	$CES (XF_{lnd}, ND^2)$
<b>Default</b>	
VA	$CES (LAB^1, VA^1)$
VA <sup>1</sup>	$CES (XF_{lnd}, KEF)$

The equations for the intermediate nests will be described according to the demand for the individual component bundles, with the price equations described subsequently. The first set of equations determines the bundles  $VA^1$  and  $VA^2$ . The former is a share of  $VA$  for all activities. The latter is a share of  $VA^1$  in the case of crops and of  $VA$  in the case of livestock. The default production structure does not require the  $VA^2$  bundle.

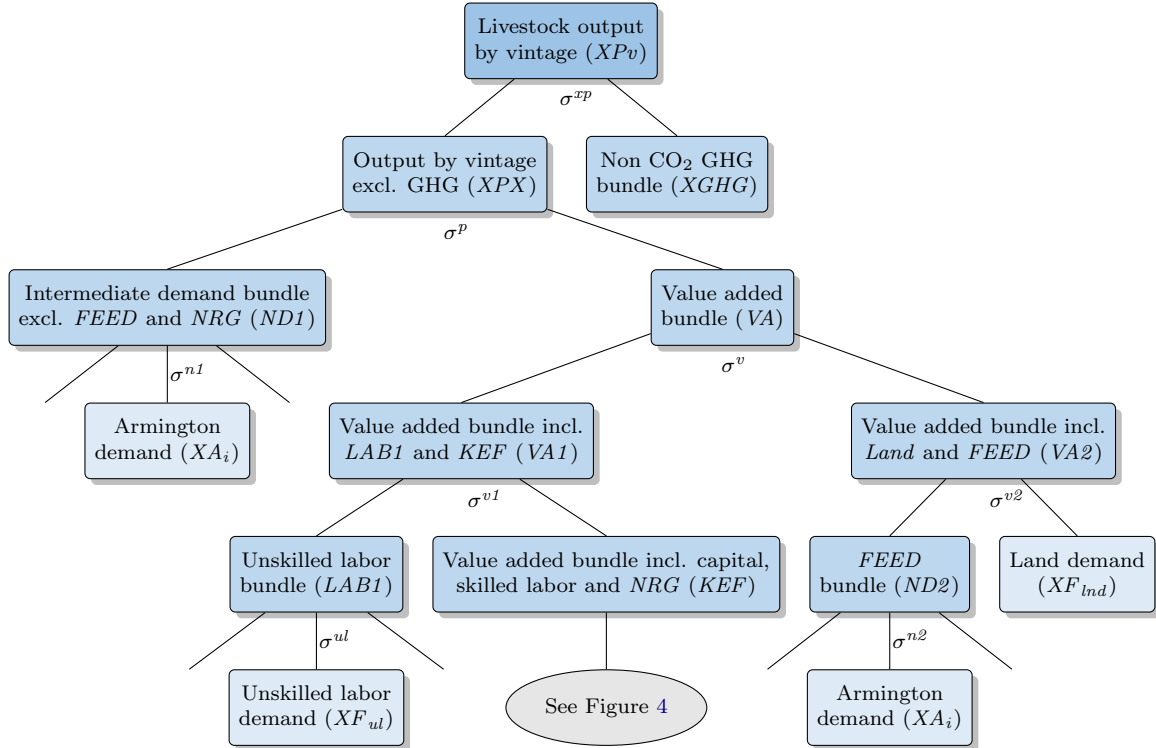


Figure 2: Livestock production nest

$$VA_{r,a,v}^1 = \alpha_{r,a,v}^{va1} \left( \frac{PVA_{r,a,v}}{PVA_{r,a,v}^1} \right)^{\sigma_{r,a,v}^v} VA_{r,a,v} \quad (\text{P-9})$$

$$VA_{r,a,v}^2 = \begin{cases} \alpha_{r,a,v}^{va2} \left( \frac{PVA_{r,a,v}^1}{PVA_{r,a,v}^2} \right)^{\sigma_{r,a,v}^{v1}} VA_{r,a,v}^1 & \text{if } a \in \{Crops\} \\ \alpha_{r,a,v}^{va2} \left( \frac{PVA_{r,a,v}}{PVA_{r,a,v}^2} \right)^{\sigma_{r,a,v}^v} VA_{r,a,v} & \text{if } a \in \{Livestock\} \end{cases} \quad (\text{P-10})$$

The next set of equations determine the bundles  $LAB^1$ ,  $KEF$  and  $ND^2$ . The subsequent decomposition of these three bundles will be identical for all activities. The  $LAB^1$  bundle is a share of  $VA^1$  for *Livestock* and  $VA$  for all other activities. The  $KEF$  bundle is a share of  $VA^2$  for *Crops* and  $VA^1$  for all other activities. The  $ND^2$  bundle is a share of  $VA^1$  for *Crops* and  $VA^2$  for *Livestock*. The  $ND^2$  bundle is not used in the *Default* production structure.

$$LAB_{r,a}^1 = \begin{cases} \sum_v \alpha_{r,a,v}^{l1} \left( \frac{PVA_{r,a,v}}{PLAB_{r,a}^1} \right)^{\sigma_{r,a,v}^v} VA_{r,a,v} & \text{if } a \in \{Default\} \\ \sum_v \alpha_{r,a,v}^{l1} \left( \frac{PVA_{r,a,v}^1}{PLAB_{r,a}^1} \right)^{\sigma_{r,a,v}^{v1}} VA_{r,a,v}^1 & \text{if } a \in \{Livestock\} \end{cases} \quad (\text{P-11})$$

$$KEF_{r,a,v} = \begin{cases} \alpha_{r,a,v}^{kef} \left( \frac{PVA_{r,a,v}^2}{PKEF_{r,a,v}} \right)^{\sigma_{r,a,v}^{v2}} VA_{r,a,v}^2 & \text{if } a \in \{Crops\} \\ \alpha_{r,a,v}^{kef} \left( \frac{PVA_{r,a,v}^1}{PKEF_{r,a,v}} \right)^{\sigma_{r,a,v}^{v1}} VA_{r,a,v}^1 & \text{if } a \in \{Default\} \end{cases} \quad (\text{P-12})$$

$$ND_{r,a}^2 = \begin{cases} \sum_v \alpha_{r,a,v}^{nd2} \left( \frac{PVA_{r,a,v}^1}{PND_{r,a}^2} \right)^{\sigma_{r,a,v}^{v1}} VA_{r,a,v}^1 & \text{if } a \in \{Crops\} \\ \sum_v \alpha_{r,a,v}^{nd2} \left( \frac{PVA_{r,a,v}^2}{PND_{r,a}^2} \right)^{\sigma_{r,a,v}^{v2}} VA_{r,a,v}^2 & \text{if } a \in \{Livestock\} \end{cases} \quad (\text{P-13})$$

The final demand equation in this set of intermediate nests determines the demand for the land factor,  $XF_{lnd}$ . It is a share of  $VA^2$  for *Crops* and *Livestock* and a share of  $VA^1$  for the *Default* activities. The parameter  $\lambda_{lnd}^f$  allows for efficiency improvement in the use of land. The relevant price of land is  $XF_{lnd}^p$  that represents the user (or agent) price of land. It is equal to the market price of land adjusted for an activity-specific tax or subsidy.

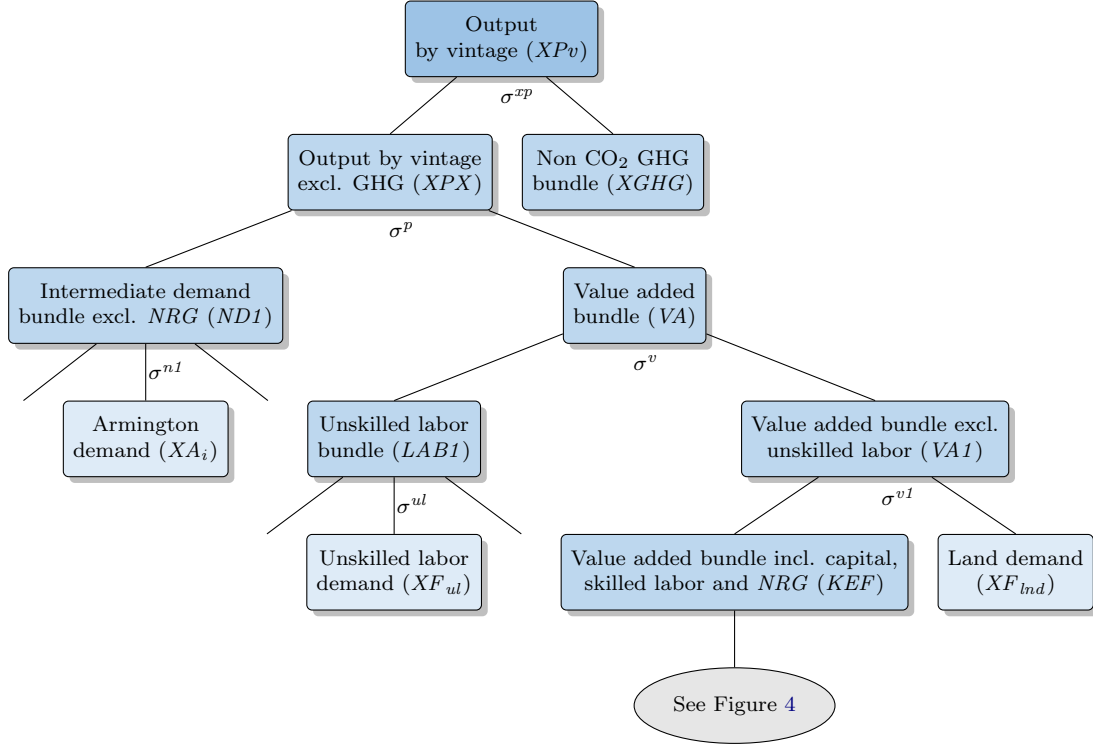
$$XF_{r,lnd,a} = \begin{cases} \sum_v \alpha_{r,a,v}^{land} \left( \frac{\lambda_{r,lnd,a}^f PVA_{r,a,v}^2}{PF_{r,lnd,a}^p} \right)^{\sigma_{r,a,v}^{v2}} \frac{VA_{r,a,v}^2}{\lambda_{r,lnd,a}^f} & \text{if } a \in \{Crops\} \\ \sum_v \alpha_{r,a,v}^{land} \left( \frac{\lambda_{r,lnd,a}^f PVA_{r,a,v}^2}{PF_{r,lnd,a}^p} \right)^{\sigma_{r,a,v}^{v2}} \frac{VA_{r,a,v}^2}{\lambda_{r,lnd,a}^f} & \text{if } a \in \{Livestock\} \\ \sum_v \alpha_{r,a,v}^{land} \left( \frac{\lambda_{r,lnd,a}^f PVA_{r,a,v}^1}{PF_{r,lnd,a}^p} \right)^{\sigma_{r,a,v}^{v1}} \frac{VA_{r,a,v}^1}{\lambda_{r,lnd,a}^f} & \text{if } a \in \{Default\} \end{cases} \quad (\text{P-14})$$

The description of the intermediate nests is finished with the determination of the prices of the intermediate bundles  $VA$ ,  $VA^1$  and  $VA^2$ . Equation (P-15) determines the price of the  $VA$  bundle,  $PVA$ . Equation (P-16) determines the price of the  $VA^1$  bundle,  $PVA^1$ . Equation (P-16) determines the price of the  $VA^2$  bundle,  $PVA^2$ .

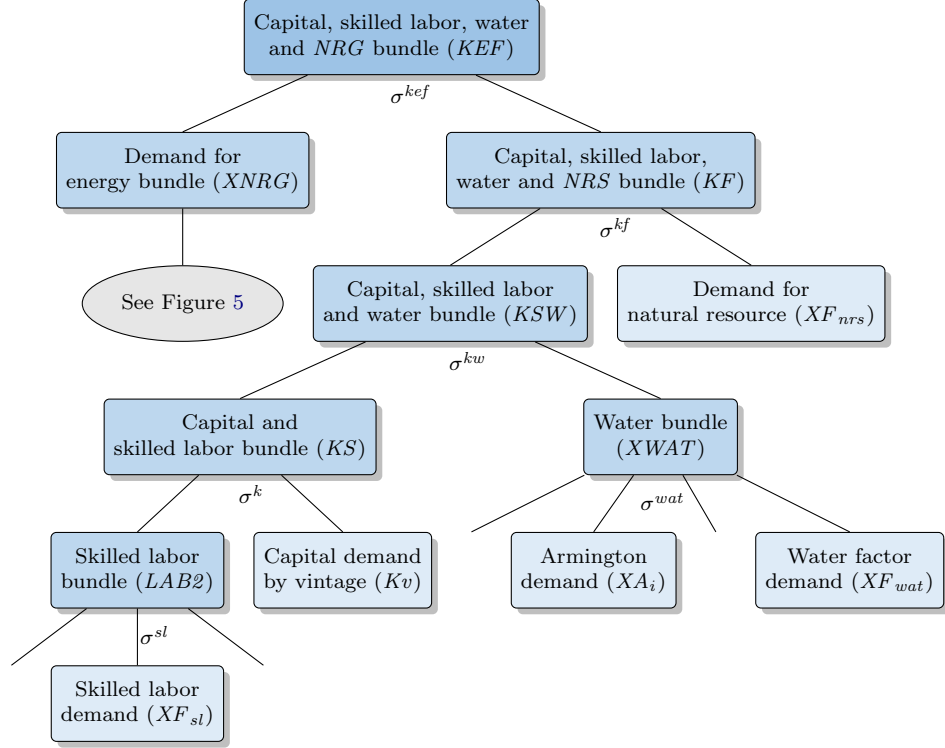
$$PVA_{r,a,v} = \begin{cases} \left[ \alpha_{r,a,v}^{ll} (PLAB_{r,a}^1)^{1-\sigma_{r,a,v}^v} + \alpha_{r,a,v}^{va1} (PVA_{r,a,v}^1)^{1-\sigma_{r,a,v}^v} \right]^{\frac{1}{1-\sigma_{r,a,v}^v}} & \text{if } a \in \{Default\} \\ \left[ \alpha_{r,a,v}^{va1} (PVA_{r,a,v}^1)^{1-\sigma_{r,a,v}^v} + \alpha_{r,a,v}^{va2} (PVA_{r,a,v}^2)^{1-\sigma_{r,a,v}^v} \right]^{\frac{1}{1-\sigma_{r,a,v}^v}} & \text{if } a \in \{Livestock\} \end{cases} \quad (P-15)$$

$$PVA_{r,a,v}^1 = \begin{cases} \left[ \alpha_{r,a,v}^{nd2} (PND_{r,a}^2)^{1-\sigma_{r,a,v}^{v1}} + \alpha_{r,a,v}^{va2} (PVA_{r,a,v}^2)^{1-\sigma_{r,a,v}^{v1}} \right]^{\frac{1}{1-\sigma_{r,a,v}^{v1}}} & \text{if } a \in \{Crops\} \\ \left[ \alpha_{r,a,v}^{ll} (PLAB_{r,a}^1)^{1-\sigma_{r,a,v}^{v1}} + \alpha_{r,a,v}^{kef} (PKEF_{r,a,v})^{1-\sigma_{r,a,v}^{v1}} \right]^{\frac{1}{1-\sigma_{r,a,v}^{v1}}} & \text{if } a \in \{Livestock\} \\ \left[ \alpha_{r,a,v}^{land} \left( \frac{PF_{r,land,a}^p}{\lambda_{r,land,a}^f} \right)^{1-\sigma_{r,a,v}^{v1}} + \alpha_{r,a,v}^{kef} (PKEF_{r,a,v})^{1-\sigma_{r,a,v}^{v1}} \right]^{\frac{1}{1-\sigma_{r,a,v}^{v1}}} & \text{if } a \in \{Default\} \end{cases} \quad (P-16)$$

$$PVA_{r,a,v}^2 = \begin{cases} \left[ \alpha_{r,a,v}^{land} \left( \frac{PF_{r,land,a}^p}{\lambda_{r,land,a}^f} \right)^{1-\sigma_{r,a,v}^{v2}} + \alpha_{r,a,v}^{kef} (PKEF_{r,a,v})^{1-\sigma_{r,a,v}^{v2}} \right]^{\frac{1}{1-\sigma_{r,a,v}^{v2}}} & \text{if } a \in \{Crops\} \\ \left[ \alpha_{r,a,v}^{land} \left( \frac{PF_{r,land,a}^p}{\lambda_{r,land,a}^f} \right)^{1-\sigma_{r,a,v}^{v2}} + \alpha_{r,a,v}^{nd2} (PND_{r,a}^2)^{1-\sigma_{r,a,v}^{v2}} \right]^{\frac{1}{1-\sigma_{r,a,v}^{v2}}} & \text{if } a \in \{Livestock\} \end{cases} \quad (P-17)$$


 Figure 3: **Default production nest**

The subsequent production nests are identical for all three production prototypes. The  $KEF$  bundle is composed of the  $KF$  bundle, a composition of capital, skilled labor, water and natural resources, and the energy bundle,  $XNRG$ . Equations (P-18) and (P-19) represent the derived demands for the  $KF$  and  $XNRG$  bundles respectively, with the bundle prices represented by  $PKF$  and  $PNRG$ . The main substitution elasticity is  $\sigma^{kef}$ . Equation (P-20) represents the CES dual price expression for the price of the  $KEF$  bundle,  $PKEF$ .


 Figure 4: **KEF bundle nest**

$$KF_{r,a,v} = \alpha_{r,a,v}^{kf} \left( \frac{PKEF_{r,a,v}}{PKF_{r,a,v}} \right)^{\sigma_{r,a,v}^{kef}} KEF_{r,a,v} \quad (\text{P-18})$$

$$XNRG_{r,a,v} = \alpha_{r,a,v}^e \left( \frac{PKEF_{r,a,v}}{PNRG_{r,a,v}} \right)^{\sigma_{r,a,v}^{kef}} KEF_{r,a,v} \quad (\text{P-19})$$

$$PKEF_{r,a,v} = \left[ \alpha_{r,a,v}^{kf} (PKF_{r,a,v})^{1-\sigma_{r,a,v}^{kef}} + \alpha_{r,a,v}^e (PNRG_{r,a,v})^{1-\sigma_{r,a,v}^{kef}} \right]^{\frac{1}{1-\sigma_{r,a,v}^{kef}}} \quad (\text{P-20})$$

The  $KF$  bundle is composed of the  $KSW$  bundle, a composition of capital, skilled labor and water and the sector-specific natural resource,  $XF_{nrs}$ . Equations (P-21) and (P-22) represent the derived demands for the  $KSW$  bundle and the sector-specific natural resource,  $XF_{nrs}$  (when present), respectively, with the respective prices represented by  $PKSW$  and  $PF_{nrs}^p$ . The main substitution elasticity is  $\sigma^{kf}$ . The natural resource factor has an efficiency factor,  $\lambda_{nrs}^f$ , that is typically exogenous. Equation (P-23) represents the CES dual price expression for the price of the  $KF$  bundle,  $PKF$ .

$$KSW_{r,a,v} = \alpha_{r,a,v}^{ksw} \left( \frac{PKF_{r,a,v}}{PKSW_{r,a,v}} \right)^{\sigma_{r,a,v}^{kf}} KF_{r,a,v} \quad (\text{P-21})$$

$$XF_{r,nrs,a} = \sum_v \alpha_{r,a,v}^{nrs} \left( \frac{\lambda_{r,nrs,a}^f PKF_{r,a,v}}{PF_{r,nrs,a}^p} \right)^{\sigma_{r,a,v}^{kf}} \frac{KF_{r,a,v}}{\lambda_{r,nrs,a}^f} \quad (\text{P-22})$$

$$PKF_{r,a,v} = \left[ \alpha_{r,a,v}^{ksw} (PKSW_{r,a,v})^{1-\sigma_{r,a,v}^{kf}} + \alpha_{r,a,v}^{nrs} \left( \frac{PF_{r,nrs,a}^p}{\lambda_{r,nrs,a}^f} \right)^{1-\sigma_{r,a,v}^{kf}} \right]^{\frac{1}{1-\sigma_{r,a,v}^{kf}}} \quad (\text{P-23})$$

The  $KSW$  bundle is composed of the  $KS$  bundle, a composition of capital and skilled labor and the water bundle,  $XWAT$ . Equations (P-24) and (P-25) represent the derived demands for the  $KS$  bundle and the water bundle, respectively, with the respective prices represented by  $PKS$  and  $PWAT$ . The main substitution elasticity is  $\sigma^{kw}$ . Equation (P-26) represents the CES dual price expression for the price of the  $KSW$  bundle,  $PKSW$ .

$$KS_{r,a,v} = \alpha_{r,a,v}^{ks} \left( \frac{PKSW_{r,a,v}}{PKS_{r,a,v}} \right)^{\sigma_{r,a,v}^{kw}} KSW_{r,a,v} \quad (\text{P-24})$$

$$XWAT_{r,a} = \sum_v \alpha_{r,a,v}^{wat} \left( \frac{PKSW_{r,a,v}}{PWAT_{r,a}} \right)^{\sigma_{r,a,v}^{kw}} KSW_{r,a,v} \quad (\text{P-25})$$

$$PKSW_{r,a,v} = \left[ \alpha_{r,a,v}^{ks} (PKS_{r,a,v})^{1-\sigma_{r,a,v}^{kw}} + \alpha_{r,a,v}^{wat} (PWAT_{r,a})^{1-\sigma_{r,a,v}^{kw}} \right]^{\frac{1}{1-\sigma_{r,a,v}^{kw}}} \quad (\text{P-26})$$

The  $KS$  bundle is composed of capital (by vintage),  $K^v$ , and the skilled labor bundle  $LAB^2$ . Equations (P-27) and (P-28) represent the derived demands for capital and the skilled labor bundle, respectively, with the respective prices represented by  $PK^p$  and  $PLAB^2$ . The main substitution elasticity is  $\sigma^k$ . The capital factor has an efficiency factor,  $\lambda_{cap}^f$ , that is typically exogenous. Equation (P-29) represents the CES dual price expression for the price of the  $KS$  bundle,  $PKS$ .

$$K_{r,a,v}^v = \alpha_{r,a,v}^k \left( \frac{\lambda_{r,cap,a}^f PKS_{r,a,v}}{PK_{r,a,v}^p} \right)^{\sigma_{r,a,v}^k} \frac{KS_{r,a,v}}{\lambda_{r,cap,a}^f} \quad (\text{P-27})$$

$$LAB_{r,a}^2 = \sum_v \alpha_{r,a,v}^{l2} \left( \frac{PKS_{r,a,v}}{PLAB_{r,a}^2} \right)^{\sigma_{r,a,v}^k} KS_{r,a,v} \quad (\text{P-28})$$

$$PKS_{r,a,v} = \left[ \alpha_{r,a,v}^k \left( \frac{PK_{r,a,v}^p}{\lambda_{r,cap,a}^f} \right)^{1-\sigma_{r,a,v}^k} + \alpha_{r,a,v}^{l2} (PLAB_{r,a}^2)^{1-\sigma_{r,a,v}^k} \right]^{\frac{1}{1-\sigma_{r,a,v}^k}} \quad (\text{P-29})$$

The next set of CES nests decomposes the two labor bundles, unskilled and skilled labor, respectively given by  $LAB^1$  and  $LAB^2$ . Equation (P-30) provides the decomposition of the bundles where the key substitution elasticities are  $\sigma^{ul}$  and  $\sigma^{sl}$ , the producer cost of labor is given by  $PF_l^p$ , and  $\lambda_l^f$  represents labor efficiency. Equation (P-31) determines the price of the unskilled labor bundle,  $PLAB^1$ , and equation (P-32) determines the price of the skilled labor bundle,  $PLAB^2$ . The user determines the composition of the labor bundles by mapping the specific skill types to



either the unskilled or skilled labor bundles. All skill levels can be mapped to one of the two bundles, in which case the other bundle would be empty and not included in the model definition.

$$XF_{r,l,a} = \begin{cases} \alpha_{r,l,a}^f \left( \frac{\lambda_{r,l,a}^f PLAB_{r,a}^1}{PF_{r,l,a}^p} \right)^{\sigma_{r,a}^{ul}} \frac{LAB_{r,a}^1}{\lambda_{r,l,a}^f} & \text{if } l \in \{Unskilled\} \\ \alpha_{r,l,a}^f \left( \frac{\lambda_{r,l,a}^f PLAB_{r,a}^2}{PF_{r,l,a}^p} \right)^{\sigma_{r,a}^{sl}} \frac{LAB_{r,a}^2}{\lambda_{r,l,a}^f} & \text{if } l \in \{Skilled\} \end{cases} \quad (\text{P-30})$$

$$PLAB_{r,a}^1 = \left[ \sum_{ul} \alpha_{r,ul,a}^f \left( \frac{PF_{r,ul,a}^p}{\lambda_{r,ul,a}^f} \right)^{1-\sigma_{r,a}^{ul}} \right]^{\frac{1}{1-\sigma_{r,a}^{ul}}} \quad (\text{P-31})$$

$$PLAB_{r,a}^2 = \left[ \sum_{sl} \alpha_{r,sl,a}^f \left( \frac{PF_{r,sl,a}^p}{\lambda_{r,sl,a}^l} \right)^{1-\sigma_{r,a}^{sl}} \right]^{\frac{1}{1-\sigma_{r,a}^{sl}}} \quad (\text{P-32})$$

Similar to the labor bundles, the next set of CES nests decomposes the three intermediate demand bundles,  $ND^1$ ,  $ND^2$  and  $XWAT$ . Recall that  $ND^2$  contains activity-specific inputs such as fertilizers in crops and feed in livestock. The  $ND^1$  bundle contains all of the other intermediate goods—with the exception of water and energy goods. The  $XWAT$  bundle contains all designated water commodities from intermediate demand as well as the water factor in some sectors—such as irrigated agriculture. Equation (P-33) provides the decomposition of the bundles where the key substitution elasticities are  $\sigma^{nd1}$ ,  $\sigma^{nd2}$  and  $\sigma^{wat}$ , the producer cost of intermediate goods is given by  $PA^a$ , and  $\lambda^{io}$  represents an efficiency factor for the use of intermediate goods. Equation (P-34) determines the price of the  $ND^1$  bundle,  $PND^1$  bundle, and equation (P-35) determines the price of the  $ND^2$  bundle,  $PND^2$ . The user determines the composition of the intermediate demand bundles by mapping the specific intermediate commodities to one of the two bundles.<sup>8</sup>

$$XA_{r,i,a} = \begin{cases} \alpha_{r,i,a}^{io} \left( \frac{\lambda_{r,i,a}^{io} PND_{r,a}^1}{PA_{r,i,a}^a} \right)^{\sigma_{r,a}^{nd1}} \frac{ND_{r,a}^1}{\lambda_{r,i,a}^{io}} & \text{if } i \in \{ND1\} \\ \alpha_{r,i,a}^{io} \left( \frac{\lambda_{r,i,a}^{io} PND_{r,a}^2}{PA_{r,i,a}^a} \right)^{\sigma_{r,a}^{nd2}} \frac{ND_{r,a}^2}{\lambda_{r,i,a}^{io}} & \text{if } i \in \{ND2\} \\ \alpha_{r,i,a}^{io} \left( \frac{\lambda_{r,i,a}^{io} PWAT_{r,a}}{PA_{r,i,a}^a} \right)^{\sigma_{r,a}^{wat}} \frac{XWAT_{r,a}}{\lambda_{r,i,a}^{io}} & \text{if } i \in \{iw\} \end{cases} \quad (\text{P-33})$$

$$PND_{r,a}^1 = \left[ \sum_{i \in \{ND1\}} \alpha_{r,i,a}^{io} \left( \frac{PA_{r,i,a}^a}{\lambda_{r,i,a}^{io}} \right)^{1-\sigma_{r,a}^{nd1}} \right]^{\frac{1}{1-\sigma_{r,a}^{nd1}}} \quad (\text{P-34})$$

$$PND_{r,a}^2 = \left[ \sum_{i \in \{ND2\}} \alpha_{r,i,a}^{io} \left( \frac{PA_{r,i,a}^a}{\lambda_{r,i,a}^{io}} \right)^{1-\sigma_{r,a}^{nd2}} \right]^{\frac{1}{1-\sigma_{r,a}^{nd2}}} \quad (\text{P-35})$$

<sup>8</sup> In the GAMS code the two mappings are driven by the sets `map1` and `map2`.

Equation (P-36) determines the demand for the water factor.<sup>9</sup> Equation (P-37) determines the price of the  $XWAT$  bundle, where the subset  $iw$  spans the set of water commodities.<sup>10</sup>

$$XF_{r,wat,a} = \alpha_{r,wat,a}^f \left( \frac{\lambda_{r,wat,a}^f PWAT_{r,a}}{PF_{r,wat,a}^p} \right)^{\sigma_{r,a}^{wat}} \frac{XWAT_{r,a}}{\lambda_{r,wat,a}^f} \quad (P-36)$$

$$PWAT_{r,a} = \left[ \sum_{i \in \{iw\}} \alpha_{r,i,a}^{io} \left( \frac{PA_{r,i,a}^a}{\lambda_{r,i,a}^{io}} \right)^{1-\sigma_{r,a}^{wat}} + \alpha_{r,wat,a}^f \left( \frac{PF_{r,wat,a}^p}{\lambda_{r,wat,a}^f} \right)^{1-\sigma_{r,a}^{wat}} \right]^{\frac{1}{1-\sigma_{r,a}^{wat}}} \quad (P-37)$$

The final set of nests in production concern the energy bundle,  $XNRG$ . It will be decomposed into demand for the energy commodities, figure 5. The energy bundle is first decomposed into electric and non-electric bundles. The latter is then decomposed into a coal bundle and a non-coal bundle (or the oil & gas bundle). The oil & gas bundle is then split into a gas bundle and an oil bundle. The four remaining bundles—electric, coal, oil and gas—represent a combination of existing or future energy sources. In the case of the original GTAP database, the electric bundle would hold the 'ely' commodity, the coal bundle would hold the 'coa' commodity, the oil bundle would hold the 'oil' and 'p\_c' commodities and the gas bundle would hold the 'gas' and 'gdt' commodities. Non-GTAP commodities would be mapped to one of the existing bundles. For example, 'p\_c' could be split into gasoline and diesel, and/or could include ethanol or bio-diesel.

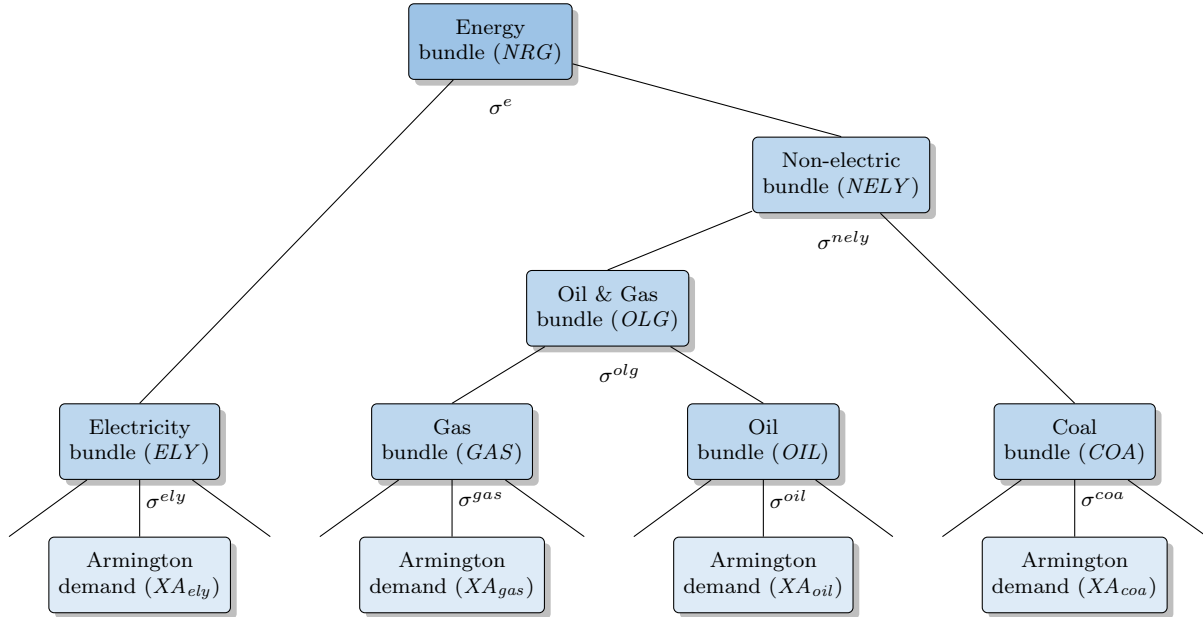


Figure 5: **Energy bundle nest**

Equation (P-38) determines the demand for the electric bundle,  $XA^{ely}$ .<sup>11</sup> Equation (P-39) de-

<sup>9</sup> At the moment, only irrigated crops have any water demand.

<sup>10</sup> In a standard aggregation of the GTAP database, the subset  $iw$  would be composed of the GTAP commodity labeled **wtr**.

<sup>11</sup> In the GAMS code, the demand for the four energy bundles ('ely', 'gas', 'oil' and 'coa') are represented by the variable **xaNRG** that has an additional dimension representing the four main energy carriers. This simplifies the decomposition of these bundles that only requires one set of equations rather than four. In principle, it also makes it somewhat easier to increase the number of energy bundles.

termines the demand for the non-electric bundle,  $XNELY$ . In both equations, the key substitution elasticity is given by  $\sigma^e$ . Equation (P-40) then describes the aggregate price of energy,  $PNRG$ .

$$XA_{r,a,v}^{ely} = \alpha_{r,a,v}^{ely} \left( \frac{PNRG_{r,a,v}}{PA_{r,a,v}^{ely}} \right)^{\sigma_{r,a,v}^e} XNRG_{r,a,v} \quad (P-38)$$

$$XNELY_{r,a,v} = \alpha_{r,a,v}^{nely} \left( \frac{PNRG_{r,a,v}}{PNELY_{r,a,v}} \right)^{\sigma_{r,a,v}^e} XNRG_{r,a,v} \quad (P-39)$$

$$PNRG_{r,a,v} = \left[ \alpha_{r,a,v}^{ely} \left( PA_{r,a,v}^{ely} \right)^{1-\sigma_{r,a,v}^e} + \alpha_{r,a,v}^{nely} \left( PNELY_{r,a,v} \right)^{1-\sigma_{r,a,v}^e} \right]^{\frac{1}{1-\sigma_{r,a,v}^e}} \quad (P-40)$$

Equation (P-41) determines the demand for the coal bundle,  $XA^{coa}$ . Equation (P-42) determines the demand for the oil & gas bundle,  $XOLG$ . In both equations, the key substitution elasticity is given by  $\sigma^{nely}$ . Equation (P-43) then describes the aggregate price of the non-electric bundle,  $PNELY$ .

$$XA_{r,a,v}^{coa} = \alpha_{r,a,v}^{coa} \left( \frac{PNELY_{r,a,v}}{PA_{r,a,v}^{coa}} \right)^{\sigma_{r,a,v}^{nely}} XNELY_{r,a,v} \quad (P-41)$$

$$XOLG_{r,a,v} = \alpha_{r,a,v}^{olg} \left( \frac{PNELY_{r,a,v}}{POLG_{r,a,v}} \right)^{\sigma_{r,a,v}^{nely}} XNELY_{r,a,v} \quad (P-42)$$

$$PNELY_{r,a,v} = \left[ \alpha_{r,a,v}^{coa} \left( PA_{r,a,v}^{coa} \right)^{1-\sigma_{r,a,v}^{nely}} + \alpha_{r,a,v}^{olg} \left( POLG_{r,a,v} \right)^{1-\sigma_{r,a,v}^{nely}} \right]^{\frac{1}{1-\sigma_{r,a,v}^{nely}}} \quad (P-43)$$

The remaining two energy bundles are oil and gas and emanate from the  $XOLG$  bundle. Equation (P-44) determines the demand for the oil bundle,  $XA^{oil}$ . Equation (P-45) determines the demand for the gas bundle,  $XA^{gas}$ . In both equations, the key substitution elasticity is given by  $\sigma^{OLG}$ . Equation (P-46) then describes the aggregate price of the oil & gas bundle,  $POLG$ .

$$XA_{r,a,v}^{oil} = \alpha_{r,a,v}^{oil} \left( \frac{POLG_{r,a,v}}{PA_{r,a,v}^{oil}} \right)^{\sigma_{r,a,v}^{olg}} XOLG_{r,a,v} \quad (P-44)$$

$$XA_{r,a,v}^{gas} = \alpha_{r,a,v}^{gas} \left( \frac{POLG_{r,a,v}}{PA_{r,a,v}^{gas}} \right)^{\sigma_{r,a,v}^{olg}} XOLG_{r,a,v} \quad (P-45)$$

$$POLG_{r,a,v} = \left[ \alpha_{r,a,v}^{oil} \left( PA_{r,a,v}^{oil} \right)^{1-\sigma_{r,a,v}^{olg}} + \alpha_{r,a,v}^{gas} \left( PA_{r,a,v}^{gas} \right)^{1-\sigma_{r,a,v}^{olg}} \right]^{\frac{1}{1-\sigma_{r,a,v}^{olg}}} \quad (P-46)$$

The final nest in the energy bundle is to decompose the four aggregate energy bundles into their constituent parts that represent the Armington demand for the energy commodities. Equation (P-47) reflects the Armington demand for energy commodity  $e$ ,  $XA$ , where the cost to producers is given by  $PA^a$ . The key substitution elasticity for each energy bundle is given by  $\sigma^{NRG}$ . Equation (P-48) represents the price of the aggregate energy bundles,  $PA^{NRG}$ .

$$XA_{r,e,a} = \sum_v \alpha_{r,e,a,v}^{eio} \left( \frac{\lambda_{r,e,a,v}^e PA_{r,a,v}^{NRG}}{PA_{r,e,a}^a} \right)^{\sigma_{r,a,v}^{NRG}} \frac{XA_{r,a,v}^{NRG}}{\lambda_{r,e,a,v}^e} \quad \text{if } e \in \{NRG\} \quad (P-47)$$

$$PA_{r,a,v}^{NRG} = \left[ \sum_{e \in \{NRG\}} \alpha_{r,e,a,v}^{eio} \left( \frac{PA_{r,e,a}^a}{\lambda_{r,e,a,v}^e} \right)^{1-\sigma_{r,a,v}^{NRG}} \right]^{\frac{1}{1-\sigma_{r,a,v}^{NRG}}} \quad (\text{P-48})$$

### 3.3 Commodity supply

The model allows for each activity  $a$  to produce one or more commodities (indexed by  $i$ ). For example, the ethanol industry could produce both ethanol as well as distiller's dried grain with solubles (or DDGS). Similarly, a single commodity can be produced by one or more activities. For example, the electricity commodity can be produced by multiple generation activities—thermal, nuclear, hydro, renewables, etc. Joint production is captured by a CET transformation function, with the possibility of perfect transformation. The aggregation of output from multiple activities is captured with a CES preference function, again with the possibility of commodity homogeneity, in which case the law-of-one-price holds.<sup>12</sup>

As an option to the standard model, power generation can be treated in a special fashion compared to other goods and services. The aggregation of production across activities is normally determined using a single CES nest. For the power activities, aggregation to the single electricity commodity uses a nested CES structure that is explained in the subsequent section.

The latest version of ENVISAGE hones to the new GTAP standard model, see [Corong \(2017\)](#), that applies the output tax to the 'make' commodities, i.e. the output tax is now both activity- and commodity-specific. This introduces a wedge between the basic price of commodity  $X$ ,  $P$ , and the tax-inclusive price,  $PP$ .<sup>13</sup>

#### 3.3.1 Non-electric goods

This section describes the make matrix for all commodities with the exception of electricity that is described subsequently. The equations thus hold for all commodities indexed by  $i$  except for the electricity commodity. The allocation of output,  $XP_{r,a}$ , from activity  $a$ , to supply one or more commodity  $i$ , is described in equation (S-1) that is the standard CET supply allocation expression. The variable  $X$  represents the supply of commodity  $i$  by activity  $a$ . The transformation elasticity is given by  $\omega^s$ . The model allows for perfect transformation, in which case the law-of-one-price holds. Equation (S-2) is in essence an equilibrium condition that determines the aggregate output of activity  $a$ ,  $XP$ . In the case of perfect transformation, it simply aggregates the individual supplies. The CET expression incorporates an efficiency parameter,  $\lambda^s$  that allows for changes between the quantity produced and the quantity to reach the market.<sup>14</sup> This can be used, for example, to reduce post-harvest losses.

$$\begin{cases} X_{r,a,i} = \gamma_{r,a,i}^p \left( \frac{1}{\lambda_{r,a,i}^s} \right)^{1+\omega_{r,a}^s} \left( \frac{P_{r,a,i}}{PX_{r,a}} \right)^{\omega_{r,a}^s} XP_{r,a} & \text{if } \omega_{r,a}^s \neq \infty \\ P_{r,a,i} = \lambda_{r,a,i}^s PX_{r,a} & \text{if } \omega_{r,a}^s = \infty \end{cases} \quad (\text{S-1})$$

<sup>12</sup> In the standard GTAP database, there is a one-to-one correspondence between activities and commodities, i.e. the make matrix is diagonal. The aggregation facility for the ENVISAGE model allows the user to determine the mapping from activities to commodities. For example, it is possible to have a single agricultural production function that produces a number of agricultural commodities. Similarly, it is possible to have a number of agricultural production activities that produce a single agricultural commodity.

<sup>13</sup> Note that the variable  $PP$  is substituted out of all equations in the model specification.

<sup>14</sup> [NEW: 4-OCT-2018]

$$PX_{r,a}XP_{r,a} = \sum_i P_{r,a,i}X_{r,a,i} \quad (\text{S-2})$$

The output tax,  $\tau^p$ , is applied to price  $P$  and generates the supply price of commodity  $i$  produced by activity  $a$ , equation (S-3). The supply of commodity  $i$  is the (CES) aggregation of output of one or more activities  $a$ . A CES preference function is used for the aggregation. Equation (S-4) determines the 'demand' for output  $a$  to compose commodity  $i$ ,  $X$ .<sup>15</sup> The substitution elasticity is given by  $\sigma^s$ . The model allows for perfect substitution, in which case the law-of-one-price holds. Equation (S-5) determines the (market) price of commodity  $i$ ,  $PS$ .

$$PP_{r,a,i} = \left(1 + \tau_{r,a,i}^p\right) P_{r,a,i} \quad (\text{S-3})$$

$$\begin{cases} X_{r,a,i} = \alpha_{r,a,i}^s \left(\frac{PS_{r,i}}{PP_{r,a,i}}\right)^{\sigma_{r,i}^s} XS_{r,i} & \text{if } \sigma_{r,i}^s \neq \infty \\ PP_{r,a,i} = PS_{r,i} & \text{if } \sigma_{r,i}^s = \infty \end{cases} \quad (\text{S-4})$$

$$PS_{r,i}XS_{r,i} = \sum_a PP_{r,a,i}X_{r,a,i} \quad (\text{S-5})$$

### 3.3.2 Domestic supply of electricity

The bundling of electricity uses a nested CES structure instead of a single nest, see figure 6. The top nest combines aggregate power supply with distribution and transmission services to form aggregate domestic electric supply. The power nest combines a number of different (and user-determined) power bundles. Subsequently, each of these power bundles are formed by the different power activities that are segmented into the different power bundles (under user-mapping). For example, the power bundles may be composed of coal-, oil- and gas-generation, nuclear, and all other. Using the GTAP power database, base load coal could be mapped to the coal power bundle, base and peak load oil could be mapped to the oil power bundle, base and peak load gas could be mapped to the gas power bundle, base load nuclear would be mapped to the nuclear power bundle, and all other power activities (wind, solar, hydro and other) could be mapped to the other power bundle. The strategy for future technologies would be to bundle them in power bundles. For example coal capture and storage could be in the coal power bundle, and advanced nuclear could be incorporated in the nuclear power bundle.

Equation (S-6) determines the demand for electricity services—indexed by activities *etd*—used to produce one or more electric commodities—indexed by *ely*.<sup>16</sup> It is linked to the total supply of power,  $XS^{pow}$ , in a CES bundle. The normal specification assumes a Leontief technology, i.e. a substitution elasticity of zero. Equation (S-7) determines the demand for the power bundle, it is a bundle of all electricity generation, and excludes the transmission and distribution services. Equation (S-8) determines the supply price of aggregate electricity.

$$X_{r,etd,ely} = \alpha_{r,etd,ely}^s \left(\frac{PS_{r,ely}}{PP_{r,etd,ely}}\right)^{\sigma_{r,ely}^{el}} XS_{r,ely} \quad (\text{S-6})$$

<sup>15</sup> Equation (S-1) determines supply and equation (S-4) determines demand. The supply/demand equilibrium condition is substituted out.

<sup>16</sup> Typically there is a single transmission and distribution activity and a single electricity commodity.

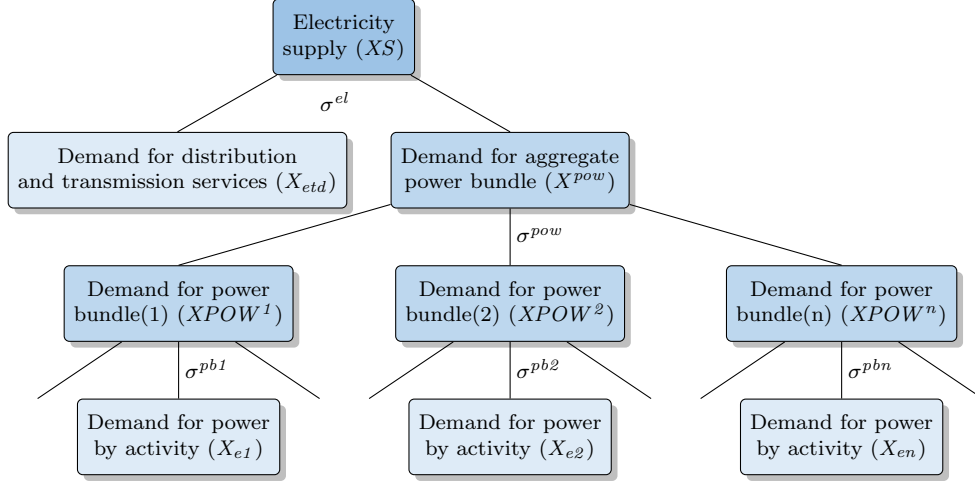


Figure 6: CES nest for power bundle

$$XPOW_{r,ely} = \alpha_{r,ely}^{pow} \left( \frac{PS_{r,ely}}{PPOW_{r,ely}} \right)^{\sigma_{r,ely}^{el}} XS_{r,ely} \quad (S-7)$$

$$PS_{r,ely} = \left[ \alpha_{r,etd,ely}^s (PP_{r,etd,ely})^{1-\sigma_{r,ely}^{el}} + \alpha_{r,ely}^{pow} (PPOW_{r,ely})^{1-\sigma_{r,ely}^{el}} \right]^{1/(1-\sigma_{r,ely}^{el})} \quad (S-8)$$

The following stage decomposes aggregate demand for power into a user-determined number of power bundles—indexed by  $pb$ . Equation (S-9) determines the demand for the power bundles. Notice that the aggregate price used in the expression is  $PPOWN$  and not  $PPOW$ . The latter is the aggregate, or average price of the power bundle. The former is a price index that is defined in equation (S-10). In the standard CES, the two price concepts are identical. The power decomposition uses the so-called adjusted CES, which preserves the additivity of the CES components (see Appendix A). The demand expressions in both versions of the CES are similar. However, the expression for the aggregate price index differs and the price index is not equal to the average price (as calculated using the zero profit condition). Thus, equation (S-11) is added that evaluates the average price,  $PPOW$ .

$$XPB_{r,pb,ely} = \alpha_{r,pb,ely}^{pb} \left( \lambda_{r,pb,ely}^{pow} \right)^{-\sigma_{r,ely}^{pow}} \left( \frac{PPOWN_{r,ely}}{PPB_{r,pb,ely}} \right)^{\sigma_{r,ely}^{pow}} XPOW_{r,ely} \quad (S-9)$$

$$PPOWN_{r,ely} = \left[ \sum_{pb} \alpha_{r,pb,ely}^{pb} \left( \lambda_{r,pb,ely}^{pow} PPB_{r,pb,ely} \right)^{-\sigma_{r,ely}^{pow}} \right]^{-1/\sigma_{r,ely}^{pow}} \quad (S-10)$$

$$PPOW_{r,ely} XPOW_{r,ely} = \sum_{pb} PPB_{r,pb,ely} XPB_{r,pb,ely} \quad (S-11)$$

The subsequent nest decomposes the various power bundles into component power activities. Each power activity is mapped to one of the aggregate power bundles. Equation (S-12) determines the demand for power generated by power activity  $elya$  that is mapped to power bundle  $pb$ . Equations (S-13) and (S-14) determine respectively the price index for the power bundle  $pb$ , as derived

from the adjusted CES price index expression, and the average price of the power bundle  $pb$  using the zero profit condition.

$$X_{r,elya,ely} = \alpha_{r,elya,ely}^s \left( \frac{PPBN_{r,pb,ely}}{\lambda_{r,elya,ely}^{pb} PP_{r,elya,ely}} \right)^{\sigma_{r,pb,ely}^{pb}} XPB_{r,pb,ely} \quad \text{if } elya \in pb \quad (\text{S-12})$$

$$PPBN_{r,pb,ely} = \left[ \sum_{elya \in pb} \alpha_{r,elya,ely}^s \left( \lambda_{r,elya,ely}^{pb} PP_{r,elya,ely} \right)^{-\sigma_{r,pb,ely}^{pb}} \right]^{-1/\sigma_{r,pb,ely}^{pb}} \quad (\text{S-13})$$

$$PPB_{r,pb,ely} XPB_{r,pb,ely} = \sum_{elya \in pb} PP_{r,elya,ely} X_{r,elya,ely} \quad (\text{S-14})$$

### 3.4 Income block

There are three domestic final demand agents—households ( $h$ ), an aggregate government sector ( $gov$ ) and an aggregate investment sector ( $inv$ ). All factor income, net of taxes, accrues to the private household, government revenues are generated by the various indirect taxes in the economy as well as direct taxes, and investment income is equated to the sum of domestic and foreign savings. A portion of capital income flows to a 'global' holder of equity that then portions out profits from the global fund. Remittances are also incorporated and are fully bilateral.

The depreciation allowance is calculated as the replacement cost of the estimated depreciation, equation Y-1. The parameter  $\delta^f$  is allowed to differ from the physical rate of depreciation,  $\delta$ , though in most cases it will be identical. The variable  $PPD_{inv}$  is the unit cost of investment and  $K^s$  is the non-normalized level of the aggregate capital stock. The normalized level of the capital stock, further described below, is scaled to the initial aggregate remuneration of capital, i.e. its price in the base year is 1. The non-normalized level is needed for calculating the depreciation allowance and in the dynamic equation for updating the aggregate capital stock.

$$DeprY_r = \delta_r^f PPD_{r,inv} K_r^s \quad (\text{Y-1})$$

The income variable  $ntmY$  represents the total income generated by the *ad valorem* equivalent (AVE) of non-tariff measures (NTMs). The latter is measured by the trade instrument  $\tau^{ntm}$ , which is an additive component to the standard bilateral tariff. Equation Y-2 defines the total NTM income—that can be distributed to both domestic and foreign agents.<sup>17</sup> The standard GTAP database does not incorporate NTMs. These have been introduced using a version of Altermat (Malcolm (1998)) using AVE estimates from Kee et al. (2009).

$$ntmY_r = \begin{cases} \sum_s \sum_i \tau_{s,i,r}^{ntm} \lambda_{s,i,r}^w PWM_{s,i,r} XW_{s,i,r} & \text{if default specification} \\ \sum_s \sum_i \sum_{aa} \tau_{s,i,r,aa}^{ntma} PWM_{s,i,r} XW_{s,i,r,aa}^a & \text{if MRIO specification} \end{cases} \quad (\text{Y-2})$$

The model incorporates some level of international income flow from profits. A portion of each region's profits (net of taxes) flows to a global equity fund that disburses its aggregate income across regions. Equation Y-3 represents the flow of a region's profits net of taxes,  $YQTF$ , to the global equity fund. Total income for the global equity fund,  $TrustY$ , is provided by equation Y-4. Foreign profit inflows,  $YQHT$ , are represented by equation Y-5. Equation Y-6 determines remittances from

<sup>17</sup> [NEW: 4-OCT-2018] The MRIO version of the model has not yet been modified to handle NTMs.

country  $r$  to country  $s$  for labor of skill  $l$ .<sup>18</sup> Remittances are calculated relative to after tax labor income.

$$YQTF_r = \chi_r^f \sum_a \left(1 - \kappa_{r, cap, a}^f\right) PF_{r, cap, a} XF_{r, cap, a} \quad (Y-3)$$

$$TrustY = \sum_r YQTF_r \quad (Y-4)$$

$$YQHT_r = \chi_r^h TrustY \quad (Y-5)$$

$$Remit_{s, l, r} = \chi_{s, l, r}^r \sum_a \left(1 - \kappa_{r, l, a}^f\right) PF_{r, l, a} XF_{r, l, a} \quad (Y-6)$$

The model incorporates government to government transfers.<sup>19</sup> These have been labeled as Official Development Assistance (ODA), but in fact cover a fairly broad range of inter-government transfers. Equation (Y-7) determines outward ODA. The basic assumption is that outward ODA grows at the same rate as GDP, where the parameter  $\chi^{OO}$  represents the initial outward ODA to GDP ratio. The equation allows for the ratio to increase with the growth in per capita income, where  $\eta^{ODA}$  represents the elasticity of ODA relative to GDP with respect to per capita income growth.<sup>20</sup> Equation (Y-8) represents aggregate ODA, i.e. the sum of ODA across all regions. Global ODA is allocated in fixed proportions across all regions, i.e. Equation (Y-9) determines inward ODA.

$$ODAOut_r = \chi_r^{OO} GDPMP_r \left( \frac{RGDPpc_r}{RGDPpc_{r,0}} \right)^{\eta^{ODA}} \quad (Y-7)$$

$$ODAGbl = \sum_r ODAOut_r \quad (Y-8)$$

$$ODAI n_r = \chi_r^{OI} ODAGbl \quad (Y-9)$$

Equation (Y-10) describes household income,  $YH$ . It is the sum across all activities of factor income, at market prices and net of taxes and depreciation. Household income also includes net foreign capital income, net remittances and some portion of NTM-related income from all modeled regions.<sup>21</sup> Factor returns at the price producers pay have a superscript  $p$ , which is absent from the factor returns at market prices. Equation (Y-11) describes disposable income,  $YD$ , where  $\kappa^h$  is the marginal (and average) rate of tax on household income. Macro closure is discussed below.

$$\begin{aligned} YH_r = & \sum_f \sum_a \left(1 - \kappa_{r, f, a}^f\right) PF_{r, f, a} XF_{r, f, a} + YQHT_r - YQTF_r \\ & + \sum_d \sum_l Remit_{r, l, d} - \sum_s \sum_l Remit_{s, l, r} - DeprY_r + \sum_s \chi_{s, r}^{hNTM} ntmY_s \end{aligned} \quad (Y-10)$$

$$YD_r = \left(1 - \kappa_r^h\right) YH_r \quad (Y-11)$$

<sup>18</sup> In the GMig model, remittances are linked to the labor income of migrants by region of origin, i.e.  $s$ . Herein, the remittances are linked to the entire wage bill (by skill), not the wage bill of migrants from region  $s$ .

<sup>19</sup> This feature was added 21-Apr-2018.

<sup>20</sup> The default value is 0.

<sup>21</sup> [NEW: 4-OCT-2018] The relevant revenue shares related to NTM income must add to 1.



The next set of equations describe government revenues, contained in the variable  $YGOV$ , which has an additional index for the different revenue streams ( $gy$ ). Equation (Y-12) describes revenues from production and cost taxes. The production tax is applied on commodity  $i$  produced by activity  $a$ . The revenue index is given by  $ptx$ . Equation (Y-13) describes revenues generated on the factors of production—labor, capital, land and natural resources. The relevant factor prices are the market prices, i.e. prices received by the factors themselves. The revenue index is given by  $vtx$ .

$$YGOV_{r,ptx} = \sum_a \left[ \sum_i \tau_{r,a,i}^p P_{r,a,i} X_{r,a,i} + \sum_v \tau_{r,a,v}^{uc} UC_{r,a,v} XP_{r,a,v} \right] \quad (Y-12)$$

$$YGOV_{r,vtx} = \sum_a \sum_f \tau_{r,f,a}^v PF_{r,f,a} XF_{r,f,a} \quad (Y-13)$$

Equation (Y-14) describes revenues generated by consumption of goods, essentially a sales tax. The sum is over all domestic agents (indexed by  $aa$ ). The relevant price is the market price of good  $i$ . The equation incorporates the two different Armington options. In the first case, the sourcing of goods is made at the national (aggregate) level and thus all users face a common Armington price,  $PAT$ , that is then adjusted by the end-user sales tax.<sup>22</sup> The second option assumes a top level Armington sourcing by agent in which case the domestic sales tax is differentiated by source. Sales taxes on import consumption is based on agents' prices of imports, which is the tariff-inclusive price of imports given by the variable  $PM^a$ . If bilateral imports are determined at the national level, the standard model assumption, agents' prices are uniform and represented by the variable  $PMT_i$ , which represents the average tariff-inclusive price of imports. In the MRIO version of the model, agents' import prices are agent-specific and represent potentially differences in both the preference structure and tariffs. The revenue index is given by  $itx$ .

$$YGOV_{r,itx} = \begin{cases} \sum_{aa} \sum_i \tau_{r,i,aa}^a \gamma_{r,i,aa}^{eda} PAT_{r,i} XA_{r,i,aa} & \text{if } ArmFlag = 0 \\ \sum_{aa} \sum_i \tau_{r,i,aa}^{ad} \gamma_{r,i,aa}^{edd} PDT_{r,i} XD_{r,i,aa} \\ \quad + \tau_{r,i,aa}^{am} \gamma_{r,i,aa}^{edm} PM_{r,i,aa}^a XM_{r,i,aa} & \text{if } ArmFlag \neq 0 \end{cases} \quad (Y-14)$$

Equation (Y-15) describes revenues generated by import tariffs. This equation has two versions depending on whether or not the MRIO version is implemented. In the standard specification, tariffs are uniform across all agents and the tax collection is done at the border relative to the aggregate level of bilateral tariffs and summed across all source countries ( $s$ ), where the first regional index is always the exporting region and the second regional index is always the destination (or importing) region. The tariffs are applied to the border (or CIF) price of imports,  $PWM$ .<sup>23</sup> In the MRIO version, tariffs are agent specific,  $\tau^{ma}$ , and applied on the variable  $XW^a$ , which represents the volume of imports by agent  $aa$  sourced in region  $s$ . It is assumed that all agents pay the same CIF

<sup>22</sup> Goods at this level are assumed to be additive though allowances are made for different prices captured by the  $\gamma^e$  coefficients.

<sup>23</sup> The model allows for iceberg trade costs using the parameter  $\lambda^w$ . This generates a wedge between the volume exported and the volume imported. The import volume is given by  $XW^d$  and the export volume by  $XW^s$ . At equilibrium, the following expression holds:  $XW^d = \lambda^w XW^s$  and this expression is used to substitute out  $XW^d$ . Thus in the model implementation, the import tax revenue expression also contains the iceberg parameter as the model only carries the variable  $XW$  without a superscript and it represents export supply.

price for their imports. The revenue index is given by  $mtx$ .

$$YGOV_{r,mtx} = \begin{cases} \sum_s \sum_i \tau_{s,i,r}^m \lambda_{s,i,r}^w PWM_{s,i,r} XW_{s,i,r}^d & \text{if default specification} \\ \sum_s \sum_i \sum_{aa} \tau_{s,i,r,aa}^{ma} PWM_{s,i,r} XW_{s,i,r,aa}^a & \text{if MRIO specification} \end{cases} \quad (Y-15)$$

Equation (Y-16) describes revenues generated by export taxes and subsidies. These are summed over all destination countries ( $d$ ). They are applied to the producer price of exports,  $PE$ . The revenue index is given by  $etx$ .

$$YGOV_{r,etx} = \sum_d \sum_i \tau_{r,i,d}^e PE_{r,i,d} XW_{r,i,d}^s \quad (Y-16)$$

Equation (Y-17) describes revenues generated by a tax on waste generated by household consumption.<sup>24</sup> The tax instrument is  $\tau^w$  and  $XA^w$  represents the level of household waste. The revenue index is given by  $wtx$ .

$$YGOV_{r,wtx} = \sum_h \sum_i \tau_{r,i,h}^w PA_{r,i,h} XA_{r,i,h}^w \quad (Y-17)$$

Equation (Y-18) describes revenues generated by carbon taxes. This equation holds for either Armington specification. Details on the carbon tax are described further below. The revenue index is given by  $ctx$ .

$$YGOV_{r,ctx} = \begin{cases} \sum_{em} \sum_i \sum_{aa} \chi_{em}^{Emi} \rho_{r,em,i,aa}^{Emi} \phi_{r,em,i,aa}^{Emi} \tau_{r,em,aa}^{Emi} XA_{r,i,aa} & \text{if } ArmFlag = 0 \\ \sum_{em} \sum_i \sum_{aa} \chi_{em}^{Emi} \rho_{r,em,i,aa}^{Emi,d} \phi_{r,em,i,aa}^{Emi} \tau_{r,em,aa}^{Emi} XD_{r,i,aa} \\ + \chi_{em}^{Emi} \rho_{r,em,i,aa}^{Emi,m} \phi_{r,em,i,aa}^{Emi} \tau_{r,em,aa}^{Emi} XM_{r,i,aa} & \text{if } ArmFlag \neq 0 \end{cases} \quad (Y-18)$$

Finally, direct taxes are described in equation (Y-19), where the revenue index is given by  $dtx$ . Direct taxes are imposed on specific factor incomes and there is a net direct tax (that can be negative) on total household income after factor taxes that balances the government account (for a given deficit level).

$$YGOV_{r,dtx} = \sum_f \sum_a \kappa_{r,f,a}^f PF_{r,f,a} XF_{r,f,a} + \kappa_r^h YH_r \quad (Y-19)$$

Equation (Y-20) describes the financing of gross investment. The variable  $YFD$  represents final demand expenditures, in value terms, for the final demand agents, indexed by  $fd$  that takes on values of  $h$ ,  $gov$  and  $inv$  respectively for households, government and investment.<sup>25</sup> Gross investment is equated to the sum of all savings, domestic from households ( $S^h$ ) and government ( $S^g$ ), and foreign ( $S^f$ ), where the latter is evaluated using a global price index,  $PW^{sav}$ , and the depreciation allowance. Macro closure defines what variable this equation determines. In the default closure, investment is savings driven and therefore this equation determines the nominal level of investment. If investment is fixed, then this equation could determine either household or public savings. Macro closure is further discussed below. This equation is defined for all regions,

<sup>24</sup> [NEW 4-OCT-2018] The model also includes the possibility of an excise tax on waste as well as the *ad valorem* tax.

<sup>25</sup> With the introduction of the R & D module, it will also include  $r\_d$  for R & D expenditures.

but one should be dropped due to Walras' Law. In fact, the equation includes an additive shifter labeled *Walras*, that is added to the equation for a residual region. It should evaluate to zero.<sup>26</sup>

$$YFD_{r,inv} = S_r^h + S_r^g + PW^{sav} S_r^f + Depr Y_r \quad (Y-20)$$

## 3.5 Final demand

### 3.5.1 Household demand

The commodity space for household demand has a different classification than the product commodity space (indexed by  $i$ ). Household demand for commodities is indexed by  $k$ . A transition matrix approach is used to convert the household commodities into the produced commodities.<sup>27</sup> In value terms, the transition matrix is an  $i \times k$  matrix. The sum across the columns equals the value of consumption for commodity  $i$  and must line up with the values from the consumer demand vector of the SAM. The sum down the columns equals the value of consumption of commodity  $k$ . Table 3.3 provides an example from the GREEN database.<sup>28</sup>

Table 3.3: **Example of a consumer transition matrix**

	Food&Bev	Energy	Trp&Comm	Other	Total
Agric	15.9	0.1	0.0	5.2	21.3
Coal	0.0	0.3	0.0	0.0	0.3
Oil	0.0	0.0	0.0	0.0	0.0
Gas	0.0	12.8	0.0	0.0	12.8
Energy int manu	0.0	0.0	0.1	50.8	50.9
Refined oil	0.0	17.5	50.0	0.0	67.5
Electricity	0.0	73.1	0.0	22.9	96.0
Other	347.9	0.1	366.2	1,644.0	2,358.3
Total	363.9	103.8	416.3	1,723.0	2,607.0

There are several advantages to the transition matrix approach.

1. Potentially, it provides a more accurate picture of consumer demand. For example, it is possible to connect different energy bundles with different sources of demand. The demand for transport is likely to be dominated by liquid fuels, whereas the demand for energy in households is likely to be a mix of electricity, gas and coal.
2. It is likely to provide a better link with household surveys.
3. It may be easier to generate more plausible demand behavior. For example, in the standard LES demand system, a rise in the price of energy (due to a carbon tax for example) will normally lead to a rise in demand for automobiles. If fuel and cars are bundled in a transport bundle, the carbon tax will lead to a drop in the demand for the transport bundle and thus also a drop in the demand for cars.

<sup>26</sup> In the model implementation, the residual region is a subset of  $r$  and is designated as **rres**.

<sup>27</sup> The consumer demand transition matrix approach was utilized in the OECD GREEN model, see [Burniaux et al. \(1992\)](#) and [van der Mensbrugghe \(1994\)](#). See also [Cardenete et al. \(2012\)](#), section 5.3, for a more recent description of the approach in GE modeling.

<sup>28</sup> The data is from the GREEN model database and represents the transition matrix for the USA in \$1985 million at the Armington level of aggregation.

With the original GTAP database, the transition matrix approach is largely limited to mapping specific supplied commodities to consumer commodities. At a minimum there would be a one-to-one mapping of all non-energy commodities, i.e. a diagonal transition matrix for the non-energy commodities, and all energy commodities mapped to a single energy consumer good.<sup>29</sup> However, with appropriate mapping it is possible to limit the number of consumer goods. For example, the user can map the GTAP commodities `tex`, `wap` and `lea` to a single 'clothing' consumer good and yet retain the production of all three commodities individually. In a similar vein, certain agriculture and food products can be aggregated together for the purposes of the top level consumer utility function, yet continue to be produced individually. The consumer demand nesting is then used to decompose the aggregate demand, for example for food or clothing, into their components.

The consumer demand system is thus implemented as a nested structure that starts with disposable income, see figure 7. A top nest allocates disposable income between savings and consumer commodities (indexed by  $k$ ). Each commodity  $k$  is decomposed into demand for the various product commodities (indexed by  $i$ ) using a series of nested CES functions. The top CES nest decomposes demand for commodity  $k$  into a non-energy aggregate bundle and an energy bundle. The former is decomposed into demand for non-energy Armington goods with a single nest (and added across all consumer goods  $k$ ). The latter is decomposed using the same energy nesting as in production with energy demand similarly aggregated across all consumer goods  $k$  to generate total energy demand by carrier by households.

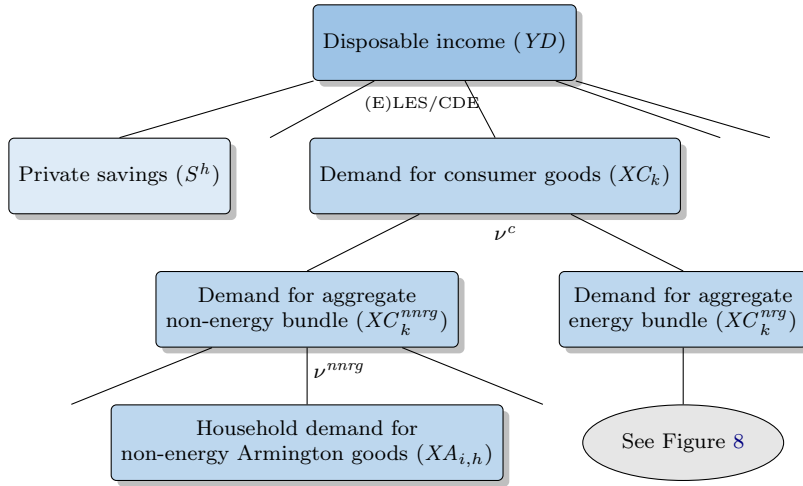


Figure 7: **Consumer demand nest**

The latest version of the ENVISAGE Model has four possible top-level consumer demand utility functions. Each can determine household demand for goods and services (at the level of consumer goods, i.e. indexed by  $k$ ). Three are in a class of models linked to the oft-used linear expenditure system (LES)—one of which, the ELES, integrates the savings decision in the utility function. The fourth utility function is the *constant differences in elasticity* (CDE) utility function—widely popularized in the GTAP model and its variants.<sup>30</sup>

<sup>29</sup> The user provides the relevant mappings between the GTAP goods (after aggregation) and the user-defined consumer goods. This mapping is part of the bridge file used by the aggregation facility.

<sup>30</sup> See Hertel (1997) and Corong et al. (2017).

### The Linear Expenditure System and its offshoots

This section describes use of the linear expenditure system (LES) and two of its offshoots. The first offshoot is the extended linear expenditure system (ELES) developed by [Lluch \(1973\)](#) and [Howe \(1975\)](#). The ELES integrates the savings decision into top level utility, i.e. the savings decision is handled with demand for the other goods and services. The second offshoot is the so-called "An Implicitly Direct Additive Demand System", or AIDADS, that leaves the possibility of non-linear Engel curves, unlike the LES—see [Rimmer and Powell \(1992a\)](#) and [Rimmer and Powell \(1996\)](#). All three demand systems share the same underlying specification with some deviations.

Under the LES-type demand system, household demand is the sum of two components—the so-called subsistence minimum, or floor consumption, and a marginal share of supernumerary income. The latter is equal to income less expenditures on the subsistence minima. Equation (D-1) defines supernumerary income on a per capita basis,  $Y^{sup}$ .<sup>31</sup> The first term is per capita income. In the case of the ELES, it is disposable income. In the case of the LES and AIDADS, it is disposable income less savings. The second term is the total amount spent on the subsistence minima where  $\gamma$  represents the per capita subsistence minimum and expenditures are evaluated at consumer prices, all taxes included,  $PC$ . Equation (D-2) determines consumer demand for good  $k$ , where  $\mu^c$  represents the marginal budget share. Utility and parameters of the LES are defined in per capita terms. Hence, per capita demand is multiplied by population to get aggregate household demand. Equation (D-3) defines the utility level for the LES-type demand specification. In the case of the ELES, the variable  $P^{sav}$  represents the perceived 'price' of savings. Given the overall myopic expectations, it is equated to the consumer price index as it represents future claims on consumption. The parameter  $\mu^s$  represents the marginal propensity to save out of disposable income.

$$Y_{r,h}^{sup} = \begin{cases} \frac{YD_r}{Pop_r} - \sum_k PC_{r,k,h} \gamma_{r,k,h} & \text{if ELES} \\ \frac{YD_r - S_{r,h}^h}{Pop_r} - \sum_k PC_{r,k,h} \gamma_{r,k,h} & \text{if LES or AIDADS} \end{cases} \quad (D-1)$$

$$XC_{r,k,h} = Pop_r \left( \gamma_{r,k,h} + \frac{\beta_{r,h}^c \mu_{r,k,h}^c}{PC_{r,k,h}} Y_{r,h}^{sup} \right) \quad (D-2)$$

$$u_{r,h} = \begin{cases} \frac{Y_{r,h}^{sup}}{\left( \frac{P_{r,h}^{sav}}{\mu_{r,h}^s} \right)^{\mu_{r,h}^s} \prod_k \left( \frac{PC_{r,k,h}}{\beta_{r,h}^c \mu_{r,k,h}^c} \right)^{\beta_{r,h}^c \mu_{r,k,h}^c}} & \text{if ELES} \\ -1 - \ln(A_{r,h}^{ad}) + \sum_k \mu_{r,k,h}^c \ln \left[ \frac{XC_{r,k,h}}{Pop_r} - \gamma_{r,k,h} \right] & \text{if LES or AIDADS} \end{cases} \quad (D-3)$$

The key difference between AIDADS and the standard LES is that the marginal budget share parameter,  $\mu^c$ , is a function of utility in the case of AIDADS, but is constant in the case of the LES. This allows for non-linear Engel curves as income expands. Equation (D-4) defines the marginal

<sup>31</sup> The consumer demand system is written to allow for multiple households. However, at the moment, disposable income and population are only provided at an aggregate level. To convert the model to multiple households would require the development of a more complete income distribution specification.

budget share specification. With  $\alpha^{ad}$  equal to  $\beta^{ad}$ , the function reverts back to the LES with constant marginal budget shares.

$$\mu_{r,k,h}^c = \frac{\alpha_{r,k,h}^{ad} + \beta_{r,k,h}^{ad} e^{u_{r,h}}}{1 + e^{u_{r,h}}} \quad (\text{D-4})$$

### The Constant Differences in Elasticity demand system

The CDE demand system is composed of four equations. The first equation, (D-5), evaluates an auxiliary variable,  $ZC$ , that is used in subsequent formulas and allows for some simplification in implementation of the CDE. The  $e^h$  parameters are referred to as the expansion parameters and are closely linked to the income elasticities (see Appendix B). The  $b^h$  parameters are referred to as the substitution parameters and are linked to the price elasticities. The variable  $u$  is the utility level and is required for the CDE demand system. The  $\alpha^h$  parameters are calibrated shifters. Equation (D-6) determines the budget shares for the CDE. Equation (D-7) implicitly determines utility that is required to calculate the budget shares. And equation (D-8) is used to back out consumption levels using the budget shares. Note that equation (D-8) is also used in the LES-type demand systems to calculate the budget shares—though these are not strictly needed.<sup>32</sup>

$$ZC_{r,k,h} = \alpha_{r,k,h}^h b_{r,k,h}^h u_{r,h}^{e_{r,k,h}^h b_{r,k,h}^h} \left( \frac{PC_{r,k,h}}{(YD_r - S_r^h)/Pop_r} \right)^{b_{r,k,h}^h} \quad (\text{D-5})$$

$$s_{r,k,h}^h = \frac{ZC_{r,k,h}}{\sum_{k'} ZC_{r,k',h}} \quad (\text{D-6})$$

$$\sum_k \frac{ZC_{r,k,h}}{b_{r,k,h}^h} \equiv 1 \quad (\text{D-7})$$

$$s_{r,k,h}^h = \frac{PC_{r,k,h} X C_{r,k,h}}{YD_r - S_r^h} \quad (\text{D-8})$$

### Conversion of consumer goods to producer goods

The subsequent nests decompose consumer demand  $XC$  into Armington demand and the latter are summed across all consumer categories  $k$ . The first nest allocates  $XC$  between a non-energy bundle and an energy bundle. Equations (D-9) and (D-10) describe respectively the demand for the aggregate non-energy and energy bundles,  $XC^{nnrg}$  and  $XC^{nrg}$ , with a substitution elasticity given by  $\nu^c$ . Equation (D-11) represents the price of the consumer good defined by index  $k$ ,  $PC$ .

$$XC_{r,k,h}^{nnrg} = \alpha_{r,k,h}^{cnrg} \left( \frac{PC_{r,k,h}}{PC_{r,k,h}^{nnrg}} \right)^{\nu_{r,k,h}^c} XC_{r,k,h} \quad (\text{D-9})$$

$$XC_{r,k,h}^{nrg} = \alpha_{r,k,h}^{cnrg} \left( \frac{PC_{r,k,h}}{PC_{r,k,h}^{nrg}} \right)^{\nu_{r,k,h}^c} XC_{r,k,h} \quad (\text{D-10})$$

$$PC_{r,k,h} = \left[ \alpha_{r,k,h}^{cnrg} \left( PC_{r,k,h}^{nnrg} \right)^{1-\nu_{r,k,h}^c} + \alpha_{r,k,h}^{cnrg} \left( PC_{r,k,h}^{nrg} \right)^{1-\nu_{r,k,h}^c} \right]^{1/(1-\nu_{r,k,h}^c)} \quad (\text{D-11})$$

<sup>32</sup> Note that the budget shares reflect the shares net of savings. In the case of the ELES, the shares needed for calibration are relative to disposable income, i.e. inclusive of savings.

Aggregate non-energy demand is decomposed next into Armington demand and summed across all consumer goods. Equation (D-12) represents demand for non-energy Armington goods by private households, where the across-goods substitution is given by  $\nu^{nnrg}$ .<sup>33</sup> The price of the non-energy demand bundle is given by equation (D-13).<sup>34</sup>

$$XA_{r,in,h} = \sum_k \left[ \alpha_{r,in,k,h}^c \left( \frac{PC_{r,k,h}^{nnrg}}{PA_{r,in,h}^h} \right)^{\nu_{r,k,h}^{nnrg}} XC_{r,k,h}^{nnrg} \right] \quad (D-12)$$

$$PC_{r,k,h}^{nnrg} = \left[ \sum_{in} \alpha_{r,in,k,h}^c \left( PA_{r,in,h}^h \right)^{1-\nu_{r,k,h}^{nnrg}} \right]^{1/(1-\nu_{r,k,h}^{nnrg})} \quad (D-13)$$

The decomposition of the energy bundle is similar to the energy nest in production, see figure 8. The energy bundle is first decomposed into electric and non-electric bundles. The latter is then decomposed into a coal bundle and a non-coal bundle (or the oil & gas bundle). The oil & gas bundle is then split into a gas bundle and an oil bundle. The four remaining bundles—electric, coal, oil and gas—represent a combination of existing or future energy sources. In the case of the original GTAP database, the electric bundle would hold the 'ely' commodity, the coal bundle would hold the 'coa' commodity, the oil bundle would hold the 'oil' and 'p\_c' commodities and the gas bundle would hold the 'gas' and 'gdt' commodities. Non-GTAP commodities would be mapped to one of the existing bundles. For example, 'p\_c' could be split into gasoline and diesel, and/or could include ethanol or bio-diesel.

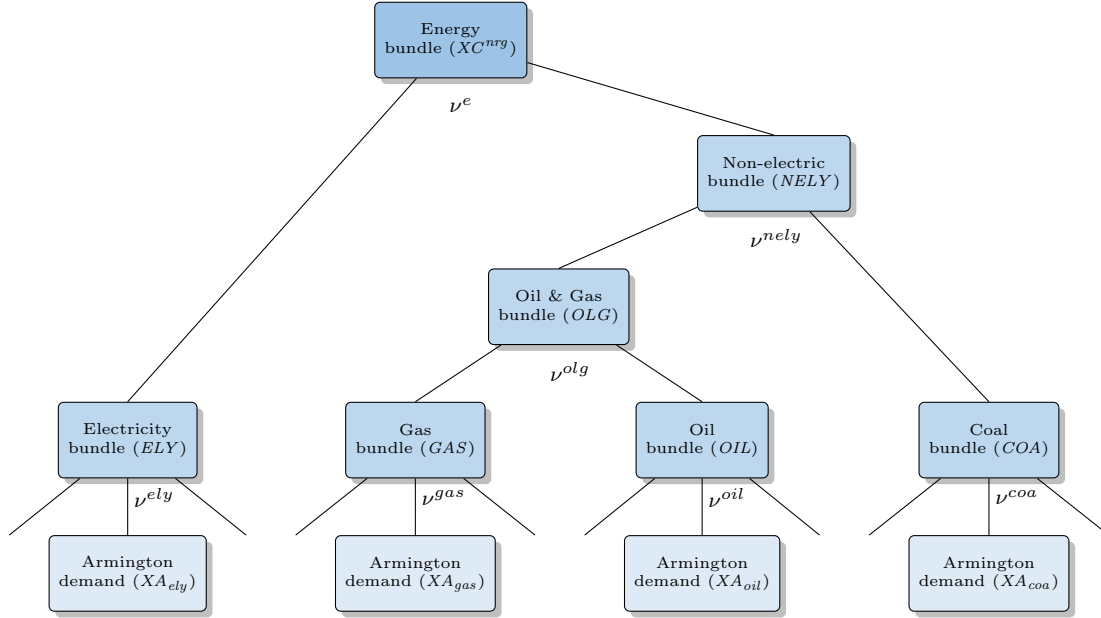


Figure 8: **Energy bundle nest in household demand**

Equation (D-14) determines the demand for the electric bundle,  $XC^{ely}$ .<sup>35</sup> Equation (D-15)

<sup>33</sup> The index  $in$  covers the non-energy commodities and is a subset of  $i$ .

<sup>34</sup> [NEW: 4-OCT-2018] Consumer prices are now represented by the variable  $PA^h$  instead of  $PA^a$ . The difference between the two prices reflects the introduction of a waste tax.

<sup>35</sup> In the GAMS code, the demand for the four energy bundles ('ely', 'gas', 'oil' and 'coa') are represented by the variable `xacNRG` that has an additional dimension representing the four main energy carriers. This simplifies the



determines the demand for the non-electric bundle,  $XC^{nely}$ . In both equations, the key substitution elasticity is given by  $\nu^e$ . Equation (D-16) then describes the aggregate price of energy,  $PC^{nrg}$ .

$$XC_{r,k,h}^{ely} = \alpha_{r,k,h}^{ely} \left( \frac{PC_{r,k,h}^{nrg}}{PC_{r,k,h}^{ely}} \right)^{\nu_{r,k,h}^e} XC_{r,k,h}^{nrg} \quad (D-14)$$

$$XC_{r,k,h}^{nely} = \alpha_{r,k,h}^{cnely} \left( \frac{PC_{r,k,h}^{nrg}}{PC_{r,k,h}^{nely}} \right)^{\nu_{r,k,h}^e} XC_{r,k,h}^{nrg} \quad (D-15)$$

$$PC_{r,k,h}^{nrg} = \left[ \alpha_{r,k,h}^{ely} \left( PC_{r,k,h}^{ely} \right)^{1-\nu_{r,k,h}^e} + \alpha_{r,k,h}^{cnely} \left( PC_{r,k,h}^{nely} \right)^{1-\nu_{r,k,h}^e} \right]^{\frac{1}{1-\nu_{r,k,h}^e}} \quad (D-16)$$

Equation (D-17) determines the demand for the coal bundle,  $XC^{coa}$ . Equation (D-18) determines the demand for the oil & gas bundle,  $XC^{olg}$ . In both equations, the key substitution elasticity is given by  $\nu^{nely}$ . Equation (D-19) then describes the aggregate price of the non-electric bundle,  $PC^{nely}$ .

$$XC_{r,k,h}^{coa} = \alpha_{r,k,h}^{cco} \left( \frac{PC_{r,k,h}^{nely}}{PC_{r,k,h}^{coa}} \right)^{\nu_{r,k,h}^{nely}} XC_{r,k,h}^{nely} \quad (D-17)$$

$$XC_{r,k,h}^{olg} = \alpha_{r,k,h}^{colg} \left( \frac{PC_{r,k,h}^{nely}}{PC_{r,k,h}^{olg}} \right)^{\nu_{r,k,h}^{nely}} XC_{r,k,h}^{nely} \quad (D-18)$$

$$PC_{r,k,h}^{nely} = \left[ \alpha_{r,k,h}^{cco} \left( PC_{r,k,h}^{coa} \right)^{1-\nu_{r,k,h}^{nely}} + \alpha_{r,k,h}^{colg} \left( PC_{r,k,h}^{olg} \right)^{1-\nu_{r,k,h}^{nely}} \right]^{\frac{1}{1-\nu_{r,k,h}^{nely}}} \quad (D-19)$$

The remaining two energy bundles are oil and gas and emanate from the  $XC^{olg}$  bundle. Equation (D-20) determines the demand for the oil bundle,  $XC^{oil}$ . Equation (D-21) determines the demand for the gas bundle,  $XC^{gas}$ . In both equations, the key substitution elasticity is given by  $\nu^{olg}$ . Equation (D-22) then describes the aggregate price of the oil & gas bundle,  $PC^{olg}$ .

$$XC_{r,k,h}^{oil} = \alpha_{r,k,h}^{coil} \left( \frac{PC_{r,k,h}^{olg}}{PC_{r,k,h}^{oil}} \right)^{\nu_{r,k,h}^{olg}} XC_{r,k,h}^{olg} \quad (D-20)$$

$$XC_{r,k,h}^{gas} = \alpha_{r,k,h}^{cgas} \left( \frac{PC_{r,k,h}^{olg}}{PC_{r,k,h}^{gas}} \right)^{\nu_{r,k,h}^{olg}} XC_{r,k,h}^{olg} \quad (D-21)$$

$$PC_{r,k,h}^{olg} = \left[ \alpha_{r,k,h}^{coil} \left( PC_{r,k,h}^{oil} \right)^{1-\nu_{r,k,h}^{olg}} + \alpha_{r,k,h}^{cgas} \left( PC_{r,k,h}^{gas} \right)^{1-\nu_{r,k,h}^{olg}} \right]^{\frac{1}{1-\nu_{r,k,h}^{olg}}} \quad (D-22)$$

The final nest in the energy bundle is to decompose the four aggregate energy bundles into their constituent parts that represent the Armington demand for the energy commodities for each commodity  $k$  and are then summed across all consumer commodities. Equation (D-23) reflects the Armington demand for energy commodity  $e$ ,  $XA_h$ , where the cost to consumers is given by

---

decomposition of these bundles that only requires one set of equations rather than four. In principle, it also makes it somewhat easier to increase the number of energy bundles.



$PA^h$ .<sup>36</sup> The key substitution elasticity for each energy bundle is given by  $\nu^{NRG}$ . Equation (D-24) represents the price of the aggregate energy bundles,  $PC^{NRG}$ .

$$XA_{r,e,h} = \sum_k \alpha_{r,e,k,h}^c \left( \frac{\lambda_{r,e,k,h}^{ce} PC_{r,k,h}^{NRG}}{PA_{r,e,h}^h} \right)^{\nu_{r,k,h}^{NRG}} \frac{XC_{r,k,h}^{NRG}}{\lambda_{r,e,k,h}^{ce}} \quad \text{if } e \in \{NRG\} \quad (\text{D-23})$$

$$PC_{r,k,h}^{NRG} = \left[ \sum_{e \in \{NRG\}} \alpha_{r,e,k,h}^c \left( \frac{PA_{r,e,h}^h}{\lambda_{r,e,k,h}^{ce}} \right)^{1-\nu_{r,k,h}^{NRG}} \right]^{\frac{1}{1-\nu_{r,k,h}^{NRG}}} \quad (\text{D-24})$$

### Household savings

Household savings is treated differently across the demand systems. In the case of the ELES, savings is incorporated in the utility function. In the case of the other demand systems, a separate assumption is made about the savings propensity and savings is subtracted from disposable income with the residual allocated to consumer goods.

Equation (D-25) determines savings for all utility function with the exception of the ELES. Savings is simply a proportion of household disposable income. The initial savings rate is given by  $aps$  and the savings rate shifter variable,  $\chi^s$  is equal, to set to 1 initially. Depending on closure, the variable  $\chi^s$  can be endogenous to target some other objective, for example, the level of investment relative to GDP. In other closures, it could be exogenous, in which case investment is savings driven. In the case of the ELES, the equation in essence determines  $\chi^s$ , i.e. the average shifter of the savings rate, since savings is determined simultaneously with consumption. In the case of the ELES, equation (D-26) determines household savings as the difference between disposable income and total expenditures on goods and services.<sup>37</sup> In the case of the other demand systems, the equation is dropped.

$$S_{r,h}^h = \chi_r^s aps_r YD_r \quad (\text{D-25})$$

$$S_{r,h}^h = YD_r - \sum_k PC_{r,k,h} XC_{r,k,h} \quad (\text{D-26})$$

### Consumer waste module

[4-OCT-2018: This section is provisional and is being tested.] Total household consumption of commodity  $i$ ,  $XA_{r,i,h}$ , is split into actual consumption,  $XA_{r,i,h}^c$ , and waste,  $XA_{r,i,h}^w$ , using the additive form of the CES function (ACES). Equations (D-27) and (D-28) determine the respective demand for actual consumption and waste, where the substitution elasticity is given by  $\sigma^{ac}$ . The demand functions allow for preference shifters given by the  $\lambda$  parameters. Equation (D-29) determines the composite price, which represents the price of utility.

$$XA_{r,i,h}^c = \alpha_{r,i,h}^{ac} XA_{r,i,h} \left( \frac{PACC_{r,i,h}}{\lambda_{r,i,h}^{ac} PA_{r,i,h}^c} \right)^{\sigma_{r,i,h}^{ac}} \quad (\text{D-27})$$

<sup>36</sup> [NEW: 4-OCT-2018] Consumer prices are given by  $PA^h$  instead of  $PA^a$  to reflect the introduction of the tax on waste.

<sup>37</sup> To target investment, in the case of the ELES it is possible to target a given level of investment (as a share of GDP for example) and to endogenize the  $\beta^c$  parameter that in effect will find a level of savings consistent with targeted investment.

$$XA_{r,i,h}^w = \alpha_{r,i,h}^{aw} XA_{r,i,h} \left( \frac{PACC_{r,i,h}}{\lambda_{r,i,h}^{aw} PA_{r,i,h}^w} \right)^{\sigma_{r,i,h}^{ac}} \quad (D-28)$$

$$PACC_{r,i,h} = \left[ \alpha_{r,i,h}^{ac} (\lambda_{r,i,h}^{ac} PA_{r,i,h}^c)^{-\sigma_{r,i,h}^{ac}} + \alpha_{r,i,h}^{aw} (\lambda_{r,i,h}^{aw} PA_{r,i,h}^w)^{-\sigma_{r,i,h}^{ac}} \right]^{-1/\sigma_{r,i,h}^{ac}} \quad (D-29)$$

It is assumed that both actual consumption and waste are both priced at the market price,  $PA_{r,i,h}$ . However, the price of waste is allowed to deviate using a 'waste' tax.<sup>38</sup> The government revenue equation has been modified to incorporate the waste tax. The final equation defines the final consumer price of total consumption,  $PA_{r,i,h}^h$ . In the absence of a waste tax, it is equal to the market price of consumption.

$$PA_{r,i,h}^c = PA_{r,i,h} \quad (D-30)$$

$$PA_{r,i,h}^w = PA_{r,i,h} (1 + \tau_{r,i,h}^w) \quad (D-31)$$

$$PA_{r,i,h}^h XA_{r,i,h} = PA_{r,i,h}^c XA_{r,i,h}^c + PA_{r,i,h}^w XA_{r,i,h}^w \quad (D-32)$$

### 3.5.2 Other final demand

In the standard model there are two other domestic agents—government and investment. The model allows for a number of other domestic final demand agents, though this typically requires supplemental information as the standard database only has public and investment expenditures. One option, for example, is to include expenditures on research and development. The same expenditure function is assumed, albeit with a different cost structure and substitution elasticity. The demand for goods and services for each is governed by a CES expenditure function. Equation (D-33) represents the demand, therefore, for the remaining final demand agents, indexed by  $fdc$ . The substitution elasticity across goods is given by  $\sigma^{fd}$  and will typically be either 0 or 1, i.e. either constant volume shares or constant value shares. Equation (D-34) determines the other final demand aggregate price deflator,  $PFD_{fdc}$ , using the ubiquitous CES dual price expression.

$$XA_{r,i,fdc} = \alpha_{r,i,fdc}^{fd} \left( \frac{PFD_{r,fdc}}{PA_{r,i,fdc}^a} \right)^{\sigma_{r,fdc}^{fd}} XFD_{r,fdc} \quad (D-33)$$

$$PFD_{r,fdc} = \left[ \sum_i \alpha_{r,i,fdc}^{fd} (PA_{r,i,fdc}^a)^{1-\sigma_{r,fdc}^{fd}} \right]^{1/(1-\sigma_{r,fdc}^{fd})} \quad (D-34)$$

The expenditure price deflator for government and investment expenditures is well-defined as the CES dual price. There is no similar expression of household consumption. A CPI index is created using the Fisher price index. More formally, for a price index over set  $i$ , the Fisher price index is equal to the following:

$$F^p = \sqrt{\frac{\sum_i P_{i,t} Q_{i,t-1}}{\sum_i P_{i,t-1} Q_{i,t-1}} \cdot \frac{\sum_i P_{i,t} Q_{i,t}}{\sum_i P_{i,t-1} Q_{i,t}}} = \sqrt{L^p \cdot P^p}$$

<sup>38</sup> The model also allows for an excise tax on waste instead of the *ad valorem* tax.

The Fisher price index represents the geometric mean of the Laspeyres ( $L^p$ ) and Paasche price indices ( $P^p$ ), with the former using lagged volume weights and the latter using current volume weights.

$$L^p = \frac{\sum_i P_{i,t} Q_{i,t-1}}{\sum_i P_{i,t-1} Q_{i,t-1}} \quad P^p = \frac{\sum_i P_{i,t} Q_{i,t}}{\sum_i P_{i,t-1} Q_{i,t}}$$

These indices represent percent change relative to the previous period and thus the price indices are chain weighted. The price index itself is thus defined as:

$$PI_t = PI_{t-1} F_t^p$$

where typically the price index will be set to 1 or 100 in some base period.

The Fisher volume index has a similar definition using different price weights:

$$F^q = \sqrt{\frac{\sum_i P_{i,t-1} Q_{i,t}}{\sum_i P_{i,t-1} Q_{i,t-1}} \cdot \frac{\sum_i P_{i,t} Q_{i,t}}{\sum_i P_{i,t} Q_{i,t-1}}} = \sqrt{L^q \cdot P^q}$$

The Fisher volume index represents the geometric mean of the Laspeyres ( $L^q$ ) and Paasche volume indices ( $P^q$ ), with the former using lagged price weights and the latter using current price weights.

$$L^q = \frac{\sum_i P_{i,t-1} Q_{i,t}}{\sum_i P_{i,t-1} Q_{i,t-1}} \quad P^q = \frac{\sum_i P_{i,t} Q_{i,t}}{\sum_i P_{i,t} Q_{i,t-1}}$$

It is easy to show that the aggregate nominal value is equal to the product of the price and volume indices:

$$P_t \cdot Q_t = \sum_i P_{i,t} \cdot Q_{i,t} = F_t^p \cdot F_t^q \cdot P_{t-1} \cdot Q_{t-1}$$

Equation (D-35) describes the CPI using the Fisher price index where we define the indicator  $QFD_{fd,tp,tp}$  to be:<sup>39</sup>

$$QFD_{r,h,tp,tq} = \sum_i PA_{r,i,h,tp}^h XA_{r,i,h,tq}$$

In other words the indicator is the value of aggregate absorption (for agent  $fd$ ) using prices of period  $tp$  and quantities period  $tq$ . Equation (D-36) defines  $YFD$  for households, i.e. total nominal expenditure.

$$PFD_{r,h,t} = PFD_{r,h,t-1} \sqrt{\frac{QFD_{r,h,t,t-1}}{QFD_{r,h,t-1,t-1}} \cdot \frac{QFD_{r,h,t,t}}{QFD_{r,h,t-1,t}}} \quad (D-35)$$

$$YFD_{r,h} = PFD_{r,h} XFD_{r,h} \quad (D-36)$$

Equation (D-37) is an identity relating nominal expenditures to real expenditures. What equation (D-37) identifies will depend on macro closure. In the standard model, it provides a definition of the volume of consumer expenditures (where the price index is defined above), the nominal level

<sup>39</sup> [NEW: 4-OCT-2014] Use of the waste-inclusive consumer price,  $PA^h$  instead of  $PA^a$ .

of government expenditures where by default the volume of government expenditures is exogenous, and the volume of investment expenditures, where the nominal level is determined by the investment-savings balance equation, equation (Y-20) above.

$$YFD_{r,fd} = PFD_{r,fd}XFD_{r,fd} \quad (D-37)$$

## 3.6 International trade

### 3.6.1 Import demand

Similar to most CGE models, import demand uses the Armington specification that assumes product differentiation between domestic and imported goods. A nested CES structure first differentiates between domestic goods and aggregate imports. A second nest differentiates imports across source regions. The GTAP database and the standard GTAP model assume that the top level Armington nest is agent-specific, i.e. the allocation of Armington demand between domestic goods and aggregate imports. The ENVISAGE model allows for 'national' sourcing, i.e. that the top level Armington disaggregation is done after aggregating Armington demand across all agents. While the former is perhaps a more plausible assumption under certain circumstances, this formulation requires significantly more computation. The top level sourcing is governed by *ArmFlag*. If it is set to zero, national sourcing is assumed. Any other value implies agent-specific sourcing. The second level decision nest decomposes import demand by region of origin. The default specification assumes this is done by a representative national agent—in common with the GTAP model and consistent with the standard GTAP database. ENVISAGE incorporates agent-specific sourcing of imports, consistent with the multi-regional input-output (MRIO) model. If used with the standard GTAP database, the initial agent-specific share parameters are uniform across agents, as are the agent-specific import tax rates. The Armington elasticities could, in principle, be differentiated across agents, and simulations could assess the impacts of changes to agent-specific tariffs.

#### National sourcing of aggregate imports

This first section describes the top level Armington equations in the case of national sourcing. Equation (T-1) adds up Armington demand (or domestic absorption for the Armington good) across all Armington agents to measure aggregate Armington demand,  $XAT$ . The latter is then decomposed into demand for domestically produced goods,  $XDT^d$ , equation (T-2), and aggregate import demand,  $XMT$ , equation (T-3). The parameter  $\sigma^{mt}$  represents the key substitution elasticity and is often referred to as the Armington elasticity. Each region in the model provides international trade and transport services  $XTT$ , further described below. It is assumed that these services are only provided using domestically produced goods. Hence, equation (T-2) also includes the region's provision of these services, in addition to the demand for domestic goods coming from the other agents in the economy. Equation (T-4) determines the Armington price (of the aggregate Armington bundle),  $PAT$ , a CES combination of the domestic price and the aggregate import price.

$$XAT_{r,i} = \sum_{aa} \gamma_{r,i,aa}^{eda} XA_{r,i,aa} \quad (T-1)$$

$$XDT_{r,i}^d = \alpha_{r,i}^{dt} \left( \frac{PAT_{r,i}}{PDT_{r,i}} \right)^{\sigma_{r,i}^{mt}} XAT_{r,i} + XTT_{r,i} \quad (T-2)$$

$$XMT_{r,i} = \alpha_{r,i}^{mt} \left( \frac{PAT_{r,i}}{PMT_{r,i}} \right)^{\sigma_{r,i}^{mt}} XAT_{r,i} \quad (\text{T-3})$$

$$PAT_{r,i} = \left[ \alpha_{r,i}^{dt} (PDT_{r,i})^{1-\sigma_{r,i}^{mt}} + \alpha_{r,i}^{mt} (PMT_{r,i})^{1-\sigma_{r,i}^{mt}} \right]^{1/(1-\sigma_{r,i}^{mt})} \quad (\text{T-4})$$

Note that the zero profit condition is given by:

$$PAT_{r,i} XAT_{r,i} = PDT_{r,i} \left( XDT_{r,i}^d - XTT_{r,i} \right) + PMT_{r,i} XMT_{r,i}$$

Agents pay  $PA$  for Armington goods, as defined by equation (T-5), where  $\tau^a$  is the agent specific tax on Armington consumption. The economy-wide Armington price,  $PAT$ , is allowed to vary across end-users using a price wedge represented by  $\gamma^{eda}$ . The emissions tax is treated as a Pigouvian tax that is added to the after-tax Armington price. The parameter  $\rho^{Emi}$  represents the rate of emissions per unit of consumption and  $\tau^{Emi}$  is the emissions tax per unit of emissions. The parameter  $\chi^{Emi}$  is a global shifter on the emission rates. And the parameter  $\phi^{Emi}$  allows for partial or full exemption of the tax incidence. By default,  $\phi^{Emi}$  is set to 1. Any valid value is between 0 (full exemption) to 1 (zero exemption). For example specific activities and/or households can be exempted from the emissions tax and for specific fuels as well—for example natural gas.

$$PA_{r,i,aa} = (1 + \tau_{r,i,aa}^a) \gamma_{r,i,aa}^{eda} PA_{r,i} + \sum_{em} \chi_{em}^{Emi} \rho_{r,em,i,aa}^{Emi} \phi_{r,em,i,aa}^{Emi} \tau_{r,em,aa}^{Emi} \quad (\text{T-5})$$

This section determines  $XAT$ ,  $XDT$ ,  $XMT$ ,  $PAT$ , and  $PA$ . The variables  $XAT$  and  $PAT$  are specific to the national-sourcing assumption and are not part of the model specification in the case of agent-sourcing. The other variables are included in the agent-sourcing specification, but have an alternative specification.

### Agent sourcing of aggregate imports

In the agent-sourcing of imports, each agent has a unique preference for domestic goods and (aggregate) imports. This specification is more computationally intensive, but is likely to be a more plausible specification. Equations (T-6) and (T-7) determine respectively the agents' demand for domestic goods and the aggregate import bundle, respectively  $XD$  and  $XM$ . All relevant variables and parameters are indexed by  $aa$ , which represents the set of all Armington agents. Equation (T-8) defines the agent-specific Armington price, which corresponds to  $PA$  above, but where the agent-specificity was due simply to the end-user sales tax since all agents were faced with the same top-level Armington price.

$$XD_{r,i,aa} = \alpha_{r,i,aa}^d \left( \frac{PA_{r,i,aa}}{PD_{r,i,aa}} \right)^{\sigma_{r,i,aa}^m} XA_{r,i,aa} \quad (\text{T-6})$$

$$XM_{r,i,aa} = \alpha_{r,i,aa}^m \left( \frac{PA_{r,i,aa}}{PM_{r,i,aa}} \right)^{\sigma_{r,i,aa}^m} XA_{r,i,aa} \quad (\text{T-7})$$

$$PA_{r,i,aa} = \left[ \alpha_{r,i,aa}^d (PD_{r,i,aa})^{1-\sigma_{r,i,aa}^m} + \alpha_{r,i,aa}^m (PM_{r,i,aa})^{1-\sigma_{r,i,aa}^m} \right]^{1/(1-\sigma_{r,i,aa}^m)} \quad (\text{T-8})$$

The next two equations define the end-user, or agent-specific, sales price, equations (T-9) and (T-10). Agents are assumed to face a common price of domestic goods,  $PDT$ . Under the MRIO specification, agents could face distinct import prices—the preference structure of bilateral

imports can differ across agents and import tariffs could be agent-specific. The agents' tariff-inclusive price of imports is given by the variable  $PM^a$ . In the standard specification, this price is uniform across agents and is given by the variable  $PMT$ . Prices can be adjusted by exogenous wedges (to allow for differences in quality for example) and sales taxes differ between domestic and imported goods. The carbon tax is a Pigouvian tax added to the tax-inclusive sales price, where the emission coefficient is source-specific.

$$PD_{r,i,aa} = \left(1 + \tau_{r,i,aa}^{ad}\right) \gamma_{r,i,aa}^{edd} PDT_{r,i} + \sum_{em} \chi_{em}^{Emi} \rho_{r,em,i,aa}^{Emi,d} \phi_{r,em,i,aa}^{Emi} \tau_{r,em,aa}^{Emi} \quad (T-9)$$

$$PM_{r,i,aa} = \left(1 + \tau_{r,i,aa}^{am}\right) \gamma_{r,i,aa}^{edm} PM_{r,i,aa}^a + \sum_{em} \chi_{em}^{Emi} \rho_{r,em,i,aa}^{Emi,m} \phi_{r,em,i,aa}^{Emi} \tau_{r,em,aa}^{Emi} \quad (T-10)$$

The final two equations determine aggregate demand for domestic goods,  $XDT$ , and the import bundle,  $XMT$ . These correspond to equation (T-1) above and represent a simple aggregation of the agent-specific demand (adjusted by the agent-specific coefficients to allow for exogenous price relatives). Note that domestic supply of international trade and transport margins is added to the (Armington) agents' demand for domestic goods. The variable  $XMT$  is not calculated for the MRIO version of the model.<sup>40</sup>

$$XDT_{r,i}^d = \sum_{aa} \gamma_{r,i,aa}^{edd} XD_{r,i,aa} + XTT_{r,i} \quad (T-11)$$

$$XMT_{r,i} = \sum_{aa} \gamma_{r,i,aa}^{edm} XM_{r,i,aa} \quad (T-12)$$

This section determines  $XD$ ,  $XM$ ,  $PA$ ,  $PD$ ,  $PM$ ,  $XDT^d$  and  $XMT$ . The variables  $PA$ ,  $XDT^d$  and  $XMT$  are common to both specifications albeit with different expressions. The variables  $XD$ ,  $XM$ ,  $PD$  and  $PM$  are only included with the agent-sourcing specification. The variable  $XMT$  is not part of the MRIO model specification.

### Second level Armington nest—standard specification

The second CES nest decomposes the aggregate demand for imports,  $XMT$ , into demand for imports by source region,  $XW^d$ . Equation (T-13) describes demand for imports by region  $r$  for imports from region  $s$  for good  $i$ .<sup>41</sup> The variable  $PDM$  represents the end-user price of imports, i.e. it includes bilateral tariffs. The key substitution elasticity is given by  $\sigma^w$ . The aggregate import price,  $PMT$ , is the CES aggregation of the tariff-inclusive bilateral prices as described in equation (T-14).

$$XW_{s,i,r}^d = \alpha_{s,i,r}^w \left( \frac{PMT_{r,i}}{PDM_{s,i,r}} \right)^{\sigma_{r,i}^w} XMT_{r,i} \quad (T-13)$$

$$PMT_{r,i} = \left[ \sum_s \alpha_{s,i,r}^w (PDM_{s,i,r})^{1-\sigma_{r,i}^w} \right]^{1/(1-\sigma_{r,i}^w)} \quad (T-14)$$

<sup>40</sup> The MRIO model only operates if the top-level Armington specification is agent-specific.

<sup>41</sup> As mentioned above, the  $XW^d$  is substituted out and thus this expression carries the iceberg parameter in the model implementation.

### Second level Armington nest—MRIO specification

Under the MRIO specification, the allocation of agent-specific import demand,  $XM$ , across source regions is also agent-specific—unlike the standard specification where agents' import demand is summed across all agents, and then allocated across trading partners. Equation (T-15) represents the agent specific allocation of  $XM$  across trading partners,  $XW^a$ . The relevant prices include  $PDM^a$ , which is the tariff-inclusive price for agent  $aa$  for commodity  $i$  sourced in region  $s$ , and  $PM^a$ , which represents the average tariff inclusive price for agent  $aa$  for commodity  $i$ , averaged over all source countries. We allow for the substitution elasticity to be agent-specific, though there is currently little empirical evidence to support this. Equation (T-16) defines the agent-specific average price of imports using the zero-profit condition. Equation (T-17) defines the tariff-inclusive price of imports for agent  $aa$  sourced in region  $s$ . Total import demand, that must be reconciled with export supply, is provided in equation (T-18).

$$XW_{s,i,r,aa}^a = \alpha_{s,i,r,aa}^{wa} \left( \frac{PM_{r,i,aa}^a}{PDM_{s,i,r,aa}^a} \right)^{\sigma_{r,i,aa}^{wa}} XM_{r,i,aa} \quad (T-15)$$

$$PM_{r,i,aa}^a XM_{r,i,aa}^a = \sum_s PDM_{s,i,r,aa}^a XW_{s,i,r,aa}^a \quad (T-16)$$

$$PDM_{s,i,r,aa}^a = (1 + \tau_{s,i,r,aa}^{ma}) PWM_{s,i,r} \quad (T-17)$$

$$XW_{s,i,r}^d = \sum_{aa} XW_{s,i,r,aa}^a \quad (T-18)$$

### Summary

The standard specification requires variables  $PMT$  and  $PDM$  that represent respectively the average tariff-inclusive price of imports—uniform across all agents—, and the bilateral tariff-inclusive price of imports. In addition, the variable  $XMT$  represents aggregate import demand, summed across all agents. These three variables are not used in the MRIO version of the model. The MRIO version of the model uses instead  $PM^a$ ,  $PDM^a$  and  $XW^a$ . Both models require the variable  $XW^d$ . Note that the MRIO model significantly expands the size of the model—even allowing for some substitution. The 10x10 model, with no substitution has around 45,000 endogenous variables. This can be pared down to some 26,000 variables with judicious substitution. The model with the standard trade specification (and substitution) solves for some 13,000 variables. The agent- and source-specific import variable requires some 12,000 variables alone for the 10x10 model.

### 3.6.2 Export supply

Domestic supply of good  $i$  is given by  $XS$ . It will be allocated between domestic and export markets analogously to the Armington assumption on the demand side. At the top level, the national supplier decides the optimal allocation of total supply between the domestic market and aggregate exports. At the second level, aggregate exports are allocated across all regions of destination. The allocation functions are based on the CET transformation specification. The implementation does allow for perfect transformation, in which case the law-of-one-price holds.

Equation (T-19) determines the optimal supply of domestic production to be allocated to the domestic market,  $XDT^s$ . The transformation elasticity is given by  $\omega^x$ . In the case of perfect transformation, the market price of goods sold domestically is equal to the average supply price



adjusted by an exogenous price wedge,  $\gamma^{esd}$ . Equation (T-20) determines the optimal supply of aggregate exports,  $XET$ . Equation (T-21) represents the zero-profit condition and is in some sense the equilibrium condition for total supply,  $XS$ .<sup>42</sup>

$$\begin{cases} XDT_{r,i}^s = \gamma_{r,i}^d (\gamma_{r,i}^{esd})^{-\omega_{r,i}^x - 1} \left( \frac{PDT_{r,i}}{PS_{r,i}} \right)^{\omega_{r,i}^x} XS_{r,i} & \text{if } \omega_{r,i}^x \neq \infty \\ PDT_{r,i} = \gamma_{r,i}^{esd} PS_{r,i} & \text{if } \omega_{r,i}^x = \infty \end{cases} \quad (\text{T-19})$$

$$\begin{cases} XET_{r,i} = \gamma_{r,i}^e (\gamma_{r,i}^{ese})^{-\omega_{r,i}^x - 1} \left( \frac{PET_{r,i}}{PS_{r,i}} \right)^{\omega_{r,i}^x} XS_{r,i} & \text{if } \omega_{r,i}^x \neq \infty \\ PET_{r,i} = \gamma_{r,i}^{ese} PS_{r,i} & \text{if } \omega_{r,i}^x = \infty \end{cases} \quad (\text{T-20})$$

$$PS_{r,i} XS_{r,i} = PDT_{r,i} XDT_{r,i}^s + PET_{r,i} XET_{r,i} \quad (\text{T-21})$$

The second level nest is represented by equation (T-22), which determines the quantity of exports allocated to each of the destination markets ( $d$ ),  $XW^s$ . The key transformation elasticity is given by  $\omega^w$ . Equation (T-23) determines the aggregate export price,  $PET$ . Note that when the elasticity is infinite, the prices in this equation are all identical and thus the equation holds in both volume and value terms. Similar to the top-level CET nest, the second level nest allows or an exogenous wedge in the bilateral prices relative to the average, or aggregate, price.

$$\begin{cases} XW_{r,i,d}^s = \gamma_{r,i,d}^w (\gamma_{r,i,d}^{ew})^{-\omega_{r,i}^w - 1} \left( \frac{PE_{r,i,d}}{PET_{r,i}} \right)^{\omega_{r,i}^w} XET_{r,i} & \text{if } \omega_{r,i}^w \neq \infty \\ PE_{r,i,d} = \gamma_{r,i,d}^{ew} PET_{r,i} & \text{if } \omega_{r,i}^w = \infty \end{cases} \quad (\text{T-22})$$

$$PET_{r,i} XET_{r,i} = \sum_d PE_{r,i,d} XW_{r,i,d}^s \quad (\text{T-23})$$

### 3.6.3 Bilateral trade prices

Each bilateral trade node is associated with four prices.  $PE$  represents the price domestic producers in region  $r$  receive for goods exported to destination region  $d$ . An export tax or subsidy,  $\tau^e$ , is introduced between the producer price and the border (of FOB) price,  $PWE$ , equation (T-24). A transportation margin represents the wedge between the border (or FOB) price of the exporting region and the border (or CIF) price of the importing region,  $PWM$ . The variable  $\zeta^{mg}$  represents the per unit transportation margin that is valued at the average price of  $PWMG$ , equation (T-25). The CIF import price is adjusted by the iceberg parameter,  $\lambda^w$ . The final price of imports for the end-user is the border (or CIF) price plus trade protection measures of which there are two: bilateral import tariffs,  $\tau_{r,i,d}^m$ , and the *ad valorem* equivalent of non tariff measures,  $\tau^{ntm}$ .<sup>43</sup> This price is represented by  $PDM$ , equation (T-26).<sup>44</sup>

$$PWE_{r,i,d} = (1 + \tau_{r,i,d}^e) PE_{r,i,d} \quad (\text{T-24})$$

$$PWM_{r,i,d} = \left[ PWE_{r,i,d} + PWMG_{r,i,d} \zeta_{r,i,d}^{mg} \right] / \lambda_{r,i,d}^w \quad (\text{T-25})$$

<sup>42</sup> It could be replaced with either the CET primal or dual expressions.

<sup>43</sup> [NEW: 4-OCT-2018] NTMs are new to the model. See introduction above in the income section (3.4).

<sup>44</sup> Note that in the model implementation, these equations are typically substituted out of the model in order to reduce its overall size.



$$PDM_{r,i,d} = (1 + \tau_{r,i,d}^m + \tau_{r,i,d}^{ntm}) PWM_{r,i,d} \quad (\text{T-26})$$

### 3.6.4 Trade margins

International trade is associated with transport margins that capture the wedge between the price at the source port and the price at the destination port, i.e. the CIF/FOB price wedge. Thus trade engenders a demand for services. The supply of these services is assumed to emanate from a global supplier that chooses the lowest cost supply subject to a CES preference function.

Equation (T-27) provides the demand for international trade and transport services per bilateral node,  $XWMG$ , a simple linear technology. The bilateral demand for trade and transport services is allocated across margin commodities ( $m$ ) using a similar linear technology, equation (T-28). The specification allows for technological change as rendered by the variable  $\lambda^{mg}$ . The average price of transportation per bilateral node,  $PWMG$ , is given by equation (T-29).

$$XWMG_{r,i,d} = \zeta_{r,i,d}^{mg} XW_{r,i,d}^s \quad (\text{T-27})$$

$$XMG_{m,r,i,d}^m = \alpha_{m,r,i,d}^{mg} \frac{XWMG_{r,i,d}}{\lambda_{m,r,i,d}^{mg}} \quad (\text{T-28})$$

$$PWMG_{r,i,d} = \sum_m \alpha_{m,r,i,d}^{mg} \frac{PTMG_m}{\lambda_{m,r,i,d}^{mg}} \quad (\text{T-29})$$

Total global demand (and therefore supply) for trade and transport margins,  $XTMG$ , for each commodity  $m$ , is the sum across all potential bilateral nodes equation (T-30). The global supplier allocates this demand across potential suppliers using a CES preference function. Equation (T-31) determines region  $r$ 's supply of trade and transport services,  $XTT$  for commodity  $m$ . The global average price for trade and transport service  $m$ ,  $PTMG$ , is given by equation (T-32).

$$XTMG_m = \sum_r \sum_i \sum_d XMG_{m,r,i,d}^m \quad (\text{T-30})$$

$$XTT_{r,m} = \alpha_{r,m}^{tt} \left( \frac{PTMG_m}{PD_{r,m}} \right)^{\sigma_m^{mg}} XTMG_m \quad (\text{T-31})$$

$$PTMG_m XTMG_m = \sum_m PD_{r,m} XTT_{r,m} \quad (\text{T-32})$$

## 3.7 Goods market equilibrium

There are two sets of equilibrium conditions. The domestic market for goods produced domestically and bilateral trade. Equation (E-1) determines  $PD$ , the market or equilibrium price equating supply of domestically produced goods to the demand for domestic goods. Equation (E-2) determines  $PE$ , the producer price of exported goods by region of origin  $r$  destined for region  $d$ . The equilibrium condition allows for iceberg trade costs that are represented by the parameter  $\lambda^w$ . The parameter is initialized at 1. An improvement in trade, i.e. a reduction in iceberg costs, is represented by an increase from 1. Thus a 10 percent improvement is represented by setting  $\lambda^w$  to 1.1. In model implementation, these equilibrium conditions are substituted out of the model specification.

$$XDT_{r,i}^d = XDT_{r,i}^s \quad (\text{E-1})$$

$$XW_{r,i,d}^d = \lambda_{r,i,d}^w XW_{r,i,d}^s \quad (\text{E-2})$$

## 3.8 Factor markets

### 3.8.1 Labor market

Equilibrium on the labor markets allows for labor market segmentation—rural, equated with agricultural sectors, and urban—or full mobility across all sectors. In the case of labor market segmentation, the model also implements a migration function that allows for some adjustment in the respective labor supplies depending on expected relative wage rates. The two labor market segments are indexed by  $z$ , or zone. When market segmentation is assumed, the index  $z$  has two elements, **rur** and **urb**. For nationally integrated labor markets, the index  $z$  has a single element, **nsg** (i.e. not-segmented). Note that the model allows for mixed specifications. For example, market segmentation can be assumed for unskilled workers but skilled workers could be fully mobile.

Within each zone (in either the segmented or non-segmented specifications), clearance of the labor market allows for less than full employment. A reservation (or minimum) wage is defined that puts a lower bound on market wages. With sufficiently high demand the equilibrium wage will likely exceed the reservation wage and the wage will be determined by equality of demand and supply (allowing for 'natural' unemployment). If labor demand is low, the market wage would be determined by the reservation wage and labor supply will exceed labor demand generating unemployment.

Equation (F-1) defines labor demand within the relevant labor market zone.<sup>45</sup>

$$LD_{r,l,z}^z = \sum_{a \in z} XF_{r,l,a} \quad (\text{F-1})$$

Labor market closure allows for endogenous regime shifting between full employment with the equilibrium wage determined by the equality of supply and demand and an a regime of less than full employment. A reservation (or minimum) wage puts a floor on wages. If labor demand is low relative to supply, the market wage will be determined by the reservation wage. In the case of high labor demand, wages will be determined by equilibrium conditions. The two regimes are specified using a mixed complementarity expression that takes the following form:

$$(W^e - W^{res})(UE - UE^{min}) = 0$$

with

$$W^e \geq W^{res} \quad \text{and} \quad UE \geq UE^{min}$$

where  $W^e$  is the market wage,  $W^{res}$  is the reservation wage,  $UE$  is the rate of unemployment and  $UE^{min}$  is a floor for the unemployment rate. For the complementarity condition to hold either the unemployment rate is at the minimum and the market wage is greater than the reservation wage, or, the unemployment level is greater than the minimum level and the market wage is equal to the reservation wage.

Figure 9 provides an illustration of the two different market closures. The vertical line labeled  $L^s$  represents total labor supply, however effective labor supply is represented by the other vertical

<sup>45</sup> All equations relating to labor market zones are conditioned on a set flag represented by **lsFlag**. For segmented labor markets, the flag will be active for **rur** and **urb**. For non-segmented labor markets, the flag will only be active for **nsg**.

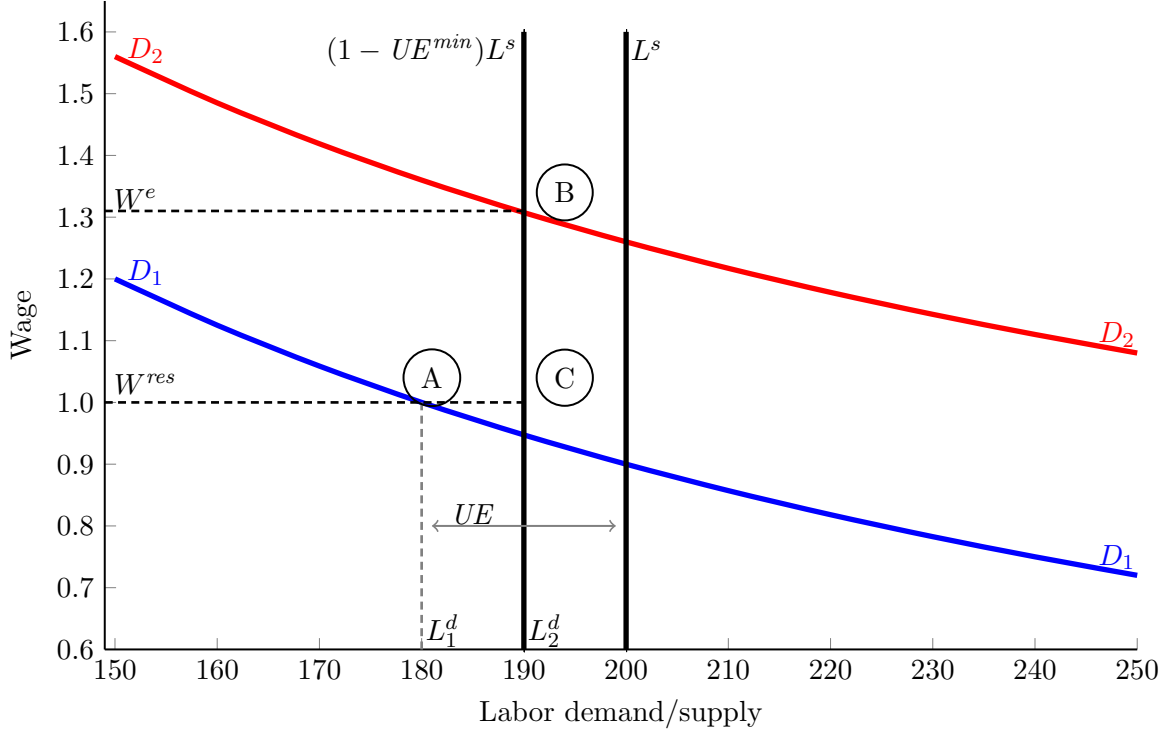


Figure 9: Labor market closure

line,  $(1 - UE^{min})L^s$ , and represents labor supply adjusted by the minimal (or natural) rate of unemployment (that is exogenous). The labor demand curves  $D_1D_1$  and  $D_2D_2$  represent two different situations with the former having relatively low labor demand and the latter relatively high labor demand. The  $D_1D_1$  curve crosses the effective labor supply curve below the reservation wage and hence the market wage will be given by the reservation wage as indicated at point A. The rate of unemployment is measured as the distance between total labor supply and effective labor demand ( $L_1^d$ ), of which some part reflects the minimum rate of unemployment. The  $D_2D_2$  curve, on the other hand, crosses the effective labor supply curve at a wage greater than the reservation wage, in which case the market wage is given by supply/demand equality at point B and with a market wage of  $W^e$ . Effective labor demand is  $L_2^d$  and equal to effective labor supply. As the demand curve  $D_1D_1$  shifts up and to the right, the market wage will be determined by the reservation wage until point C when the regime shift towards full employment occurs.

Equation (F-2) defines the reservation wage,  $W^{res}$ . It is a positive function of income growth per capita and the CPI. It is a negative function of the unemployment rate, i.e. if the rate of unemployment increases, this would tend to have a downward pressure on the reservation wage. The (positive) elasticities  $\omega^{rwg}$  and  $\omega^{rwp}$  reflect the sensitivity of the reservation wage with respect to income growth and the CPI, respectively. The (positive) elasticity  $\omega^{rue}$  reflects the sensitivity of the reservation wage with respect to the rate of unemployment.<sup>46</sup> The market wage,  $W^e$  is

<sup>46</sup> To deal with the possibility of zero unemployment, i.e. full employment, the reservation wage equation is formulated in terms of changes to 1 minus the unemployment rate, i.e. the rate of employment. The elasticity in this case should be positive, i.e. as we approach full employment, the reservation wage is likely to increase. The respective elasticities are linked by the formula:  $\eta = -\omega(1 - UE)/UE$ , where  $\omega$  is the elasticity with respect to the unemployment rate and  $\eta$  is the elasticity with respect to the employment rate. Note that this relation is not constant and depends on the rate of unemployment. For example, if the elasticity between the reservation wage and the unemployment rate is -0.5 and the initial unemployment rate is 10 percent, the elasticity with

determined by equation (F-3) reflecting the complementarity condition as the expression uses the '=g=' relational operator of GAMS. This equation will be paired with the complementarity condition on the unemployment level. Equation (F-4) defines the unemployment level.

$$W_{r,l,z,t}^{res} = \chi_{r,l,z}^{rw} (1 + g_{r,t}^y)^{\omega_{r,l,z}^{rwg}} \left( \frac{1 - UE_{r,l,z,t}^z}{1 - UE_{r,l,z,t-1}^z} \right)^{\omega_{r,l,z}^{rwue}} \left( \frac{PFD_{r,h,t}}{PFD_{r,h,t-1}} \right)^{\omega_{r,l,z}^{rwp}} \quad (F-2)$$

$$W_{r,l,z,t}^e \geq W_{r,l,z,t}^{res} \perp UE_{r,l,z,t}^z \geq UEMin_{r,l,z,t}^z \quad (F-3)$$

$$(1 - UE_{r,l,z,t}^z) LS_{r,l,z,t}^z = LD_{r,l,z,t}^z \quad (F-4)$$

The model allows for inter-sectoral wage differentials. These are represented by the  $\pi^w$  parameters that are calibrated using base year wage differentials and held fixed afterwards. This implies that all wages (for a particular skill and within a zone) move proportionately but allowing for wage differentials. Note that this assumption has potential implications for model results as any inter-sectoral shifts in labor may imply shifts in aggregate labor productivity (irrespective of technological change). Equation (F-5) sets the sectoral wage equal to the equilibrium wage (by zone) adjusted by the sector-specific wage differential.

$$PF_{r,l,a} = \pi_{r,l,a}^w W_{r,l,z}^e \quad \text{for } a \in z \quad (F-5)$$

If the rural to urban migration module is active, it depends on the so-called urban premium, i.e. the extent to which the urban wage exceeds the rural wage. Equation (F-6) defines the average wage in each of the zones,  $W^a$  (by skill). Equation (F-7) defines the urban premium,  $\pi^{Urb}$ . It is the ratio of the expected urban wage to the expected rural wage, where the actual average wage is adjusted by the probability of employment. Migration, which will be a function of the urban premium, is described below as part of the dynamics of the model. In the comparative static version of the model, it is assumed that labor supply is fixed (by zone).

$$W_{r,l,z}^a = \frac{\sum_{a \in z} PF_{r,l,a} XF_{r,l,a}}{\sum_{a \in z} XF_{r,l,a}} \quad (F-6)$$

$$\pi_{r,l}^{Urb} = \frac{(1 - UE_{r,l,Urb}^z) W_{r,l,Urb}^a}{(1 - UE_{r,l,Rur}^z) W_{r,l,Rur}^a} \quad (F-7)$$

Another potentially useful concept is the notion of the skill premium, i.e. the extent to which skilled wages exceed unskilled wages.<sup>47</sup> Equation (F-8) defines the average economy-wide wage for each skill type,  $W^t$ .<sup>48</sup> Equation (F-9) defines the so-called skill premium,  $\pi^s$ . It is the percentage by which the reference wage exceeds other wages, where the reference wage is equal to a weighted average of wages over a subset of labor skills (where the subset is indexed by  $lr$ )—the subset is typically the set of skilled labor types.

$$W_{r,l}^t = \frac{\sum_a PF_{r,l,a} XF_{r,l,a}}{\sum_a XF_{r,l,a}} \quad (F-8)$$

respect to the employment rate is 4.5 ( $= -(-0.5) \cdot 0.9/0.1$ ).

<sup>47</sup> In a dynamic calibration scenario, assumptions regarding the evolution of the skill premium can be used to calibrate the relative growth of skilled vs. unskilled labor or to calibrate productivity differentials across skills.

<sup>48</sup> In the case of non-segmented markets,  $W_{r,l}^t$  should be equal to  $W_{r,l,nsq}^a$ .

$$\pi_{r,l,t}^s = \frac{\sum_{lr} W_{r,lr,t}^t LAB_{r,lr,t}^s / \sum_{lr} LAB_{r,lr,t}^s}{W_{r,l,t}^t} - 1 \quad (\text{F-9})$$

The final two sets of equations define the total labor supply by skill, equation (F-10), and the total labor supply, equation (F-11), respectively,  $L^s$  and  $TL^s$ . In comparative static simulations,  $LS^z$  is fixed by zone. In dynamic simulations,  $LS^z$  grows at some 'natural' rate of growth, but may also be adjusted by rural to urban migration. This is further described below.

$$L_{r,l}^s = \sum_z LS_{r,l,z}^z \quad (\text{F-10})$$

$$TL_r^s = \sum_l L_{r,l}^s \quad (\text{F-11})$$

### 3.8.2 Capital market

The model can operate in recursive dynamic mode or in comparative static mode. Capital market closure differs in each. In the comparative static mode, there is a single vintage and aggregate capital is allocated across activities using a CET transformation function. The latter allows two extreme assumptions—fixed capital per activity, i.e. no mobility, and perfect mobility and a uniform rates of return across activities. In dynamic mode, with vintage capital, *New* capital is fully mobile across sectors. *Old* capital is released on the secondary capital market if installed capital exceeds demand, for example in declining sectors.

#### Comparative static closure

A CET transformation function is used to allocate aggregate capital in the comparative static version of the model. Equation (F-12) describes the supply of capital to activity  $a$ ,  $K^s$ , where  $\omega^k$  determines the degree of mobility across activities.<sup>49</sup> With perfect mobility, the rate of return across activities is uniform. Equation (F-13) determines the aggregate rate of return to capital,  $TR$ . Equation (F-14) is the equilibrium condition for capital—supply of capital to each activity is equal to demand. In the model implementation the equilibrium condition is substituted out.

$$\begin{cases} K_{r,a,v}^s = \gamma_{r,a,v}^k \left( \frac{PK_{r,a,v}}{TR_r} \right)^{\omega_r^k} TK_r^s & \text{if } \omega_r^k \neq \infty \\ PK_{r,a,v} = TR_r & \text{if } \omega_r^k = \infty \end{cases} \quad (\text{F-12})$$

$$TR_r TK_r^s = \sum_a \sum_v PK_{r,a,v} K_{r,a,v}^s \quad (\text{F-13})$$

$$K_{r,a,v}^s = K_{r,a,v}^v \quad (\text{F-14})$$

#### Vintage capital closure

There are two types of capital in the vintage specification, *Old*, or installed capital, and *New* capital. Initial installed capital is set at the beginning of the period and is equal to the depreciated level of the previous period's aggregation of both *Old* and *New* capital. The demand for *New* capital will depend on market conditions, substitution and other factors. A sector is in expansion if demand

<sup>49</sup> The vintage index is for convenience only. In the comparative static version there is only a single vintage.

for *New* capital exceeds the initial installed capital. A sector is in contraction if demand for capital in the sector is less than its initial installed capital. This can come about, for example, if public policies affect demand for a good—such as a carbon tax leading to reduced demand for coal or oil. If a sector is in contraction, it will release its *Old* capital using an upward sloping supply schedule, potentially with an infinite elasticity.

Figure 10 represents the two possible regimes each under two possible capital demand schedules. The left panel shows the case for a flat supply curve, i.e. an infinite supply curve for *Old* capital. In that case, even if capital demand is contracting, the rate of return to *Old* capital is equated to the economy-wide rental rate, i.e. the ratio of returns, *Old* over *New* is 1. The right panel shows the case where the supply of *Old* capital follows an upward sloping schedule. If capital demand is in decline, the rate of return to *Old* capital will be less than the economy-wide return. In the case of an expanding demand, the rate of return to *Old* capital is equal to the economy-wide return and the supply of *Old* capital is capped at the initially installed level, i.e.  $K_0$ .

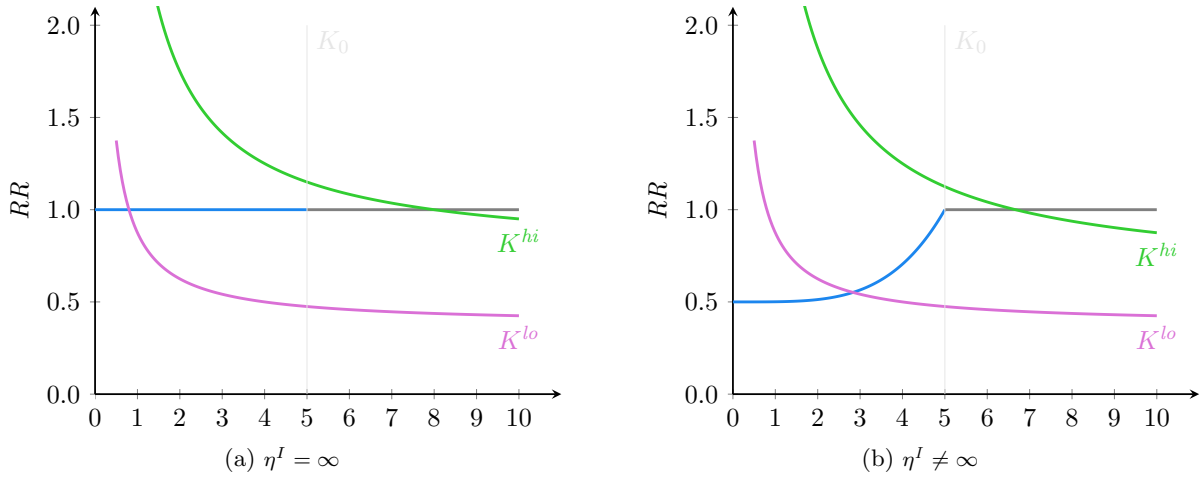


Figure 10: Capital supply regimes

Equation (F-15) represents the mixed complementarity condition for the supply of old capital. If demand is contracting, the upward sloping portion of the supply schedule is determined by the iso-elastic relation with the supply elasticity given by  $\eta^I$ . At the inflection point, the supply of *Old* capital cannot exceed the amount of the initially installed *Old* given by  $K^0$ . The complementarity condition then translates into: (a) all of the installed capital is needed,  $K^{lo} = K^0$  and the ratio of returns is equal to 1; or (b) the supply of old capital is less than the installed capital and the ratio of returns is less than 1. The second complementarity condition, equation (F-16), states that either the ratio of returns is 1 and thus supply of new capital is positive, or the ratio is less than 1 and then the supply of new capital is 0. Equation (F-17) states that total supply should meet total demand. The model allows for the possibility of a horizontal supply schedule, in which case the ratio of returns is equal to 1 irrespective of market conditions and supply will be set to demand.

$$K_{r,a}^{lo} \geq K_{r,a}^0 RR_{r,a}^{\eta^I} \perp RR_{r,a} \leq 1 \quad (\text{F-15})$$

$$RR_{r,a} \leq 1 \perp K_{r,a}^{hi} \geq 0 \quad (\text{F-16})$$

$$K_{r,a}^{lo} + K_{r,a}^{hi} \leq \sum_v K_{r,a,v}^v \quad (\text{F-17})$$

Equation (F-18) determines the economy-wide rate of return,  $TR$ . Equation (F-19) determines the sector- and vintage-specific rate of return. If the activity uses *New* capital,  $RR$  is equal to 1 and both *Old* and *New* capital obtain the rate of return  $TR$ . If the sector is in decline, it receives a rate of return lower than the economy-wide rate of return.

$$TK_r^s = \sum_a \sum_v K_{r,a,v}^v \quad (\text{F-18})$$

$$PK_{r,a,v} = RR_{r,a} TR_r \quad (\text{F-19})$$

### Determination of output by vintage

Aggregate output,  $XP$ , is determined by equilibrium conditions. Equation (F-20) represents the contemporaneous capital-output ratio for the *Old* vintage,  $kxRat$ . Equation (F-21) determines output with *Old* capital. *Old* output, will be equal to the capital output ratio applied to the supply of old capital,  $K^{lo}$ . Equation (F-22) ensures that the sum of output across vintages equals aggregate output. In the vintage capital version of the model, it essentially determines production with *New* capital. In the comparative static version, it trivially equates total production to production produced by the only vintage that is labeled *Old*.

$$kxRat_{r,a,Old} = \frac{K_{r,a,Old}^v}{XP_{r,a,Old}^v} \quad (\text{F-20})$$

$$kxRat_{r,a,Old} XP_{r,a,Old}^v = K_{r,a}^{lo} \quad (\text{F-21})$$

$$XP_{r,a} = \sum_v XP_{r,a,v} \quad (\text{F-22})$$

Equations (F-23) and (F-24) are added for accounting purposes as many of the various income identities are simplified with these substitutions. Equation (F-24) holds in all circumstances. If an activity is expanding the return on installed capital is equal to the return on *New* capital. If an activity is contracting, the only relevant return is on *Old* capital.

$$XF_{r,cap,a} = \sum_v K_{r,a,v}^v \quad (\text{F-23})$$

$$PF_{r,cap,a} = PK_{r,a,Old} \quad (\text{F-24})$$

### 3.8.3 Land market

The supply side of the land market has two components. The first component provides the aggregate supply of land. The second step allocates aggregate land across different activities allowing for a nested CET structure, the possibility of perfect mobility, and the use of the adjusted CET that preserves land additivity.

The aggregate land supply curve is allowed to have four shapes—constant elasticity, a logistic curve with an upward asymptote, a generalized hyperbola also with an upward asymptote, and

perfectly horizontal. The aggregate land supply curve is given by equation (F-25). The initial elasticity of aggregate land is given by  $\eta^t$ .

$$\left\{ \begin{array}{ll} TLand_r = \chi_r^t \left( \frac{PTLand_r}{PGDP_r} \right)^{\eta_r^t} & \text{iso-elastic} \\ TLand_r = \frac{TLand_r^{Max}}{1 + \chi_r^t e^{-\gamma_r^t (PTLand_r / PGDP_r)}} & \text{logistic} \\ TLand_r = TLand_r^{Max} - \chi_r^t \left( \frac{PTLand_r}{PGDP_r} \right)^{-\gamma_r^t} & \text{hyperbola} \\ PTLand_r = PGDP_r & \text{if } \eta_r^t = \infty \end{array} \right. \quad (F-25)$$

The iso-elastic curve by definition has a constant elasticity given by  $\eta^t$ . The logistic function, in a simplified form is given by:

$$S = \frac{\bar{S}}{1 + \chi e^{-\gamma P}}$$

where  $S$  is supply, the land maximum is given by  $\bar{S}$  and  $P$  is the real price of land. The supply elasticity is given by the following formula:

$$\eta = \gamma P \left( \frac{\bar{S} - S}{\bar{S}} \right)$$

The  $\gamma$  coefficient can be calibrated using the expression above given initial values for  $\eta$ ,  $S$ ,  $\bar{S}$  and  $P$ . The calibration formula is given by:

$$\gamma = \frac{\eta}{P} \left( \frac{\bar{S}}{\bar{S} - S} \right)$$

Calibration of the  $\chi$  parameter is given by the following expression and requires first calibrating the  $\gamma$  parameter:

$$\chi = e^{\gamma P} \left( \frac{\bar{S} - S}{\bar{S}} \right)$$

The generalized hyperbola, using the same notation is given by:

$$S = \bar{S} - \chi P^{-\gamma}$$

The supply elasticity is given by the following formula:

$$\eta = \gamma \left( \frac{\bar{S} - S}{S} \right)$$

The  $\gamma$  coefficient can be calibrated using the expression above given initial values for  $\eta$ ,  $S$ ,  $\bar{S}$  and  $P$ . The calibration formula is given by:

$$\gamma = \eta \left( \frac{S}{\bar{S} - S} \right)$$

Calibration of the  $\chi$  parameter is given by the following expression and requires first calibrating the  $\gamma$  parameter:



$$\chi = P^\gamma (\bar{S} - S)$$

The user chooses the functional form by setting the global parameter named **TASS** towards the top of a simulation file. Valid options are **KELAS** for the iso-elastic form, **LOGIST** for the logistic specification, **HYPERB** for the generalized hyperbola and **INFTY** for the horizontal supply function.<sup>50</sup>

Total land is allocated across different uses according to the relative rates of return across uses. A nested CET structure is implemented that allows for a degree of user control.<sup>51</sup> Figure 11 provides a graphic depiction of the land allocation specification. The number of land bundles (*XLB*) is under user control, but typically there should be at least one. The first land bundle is in the top CET nest with an intermediate land bundle (*XLB*). All other land bundles are nested within the intermediate land bundle.<sup>52</sup> It is possible to mimic the GTAP standard nesting, a single level nest, by defining only one land bundle and use the GTAP transformation elasticity for that bundle. The top level elasticity is essentially irrelevant since in this case the intermediate land bundle would not be used.

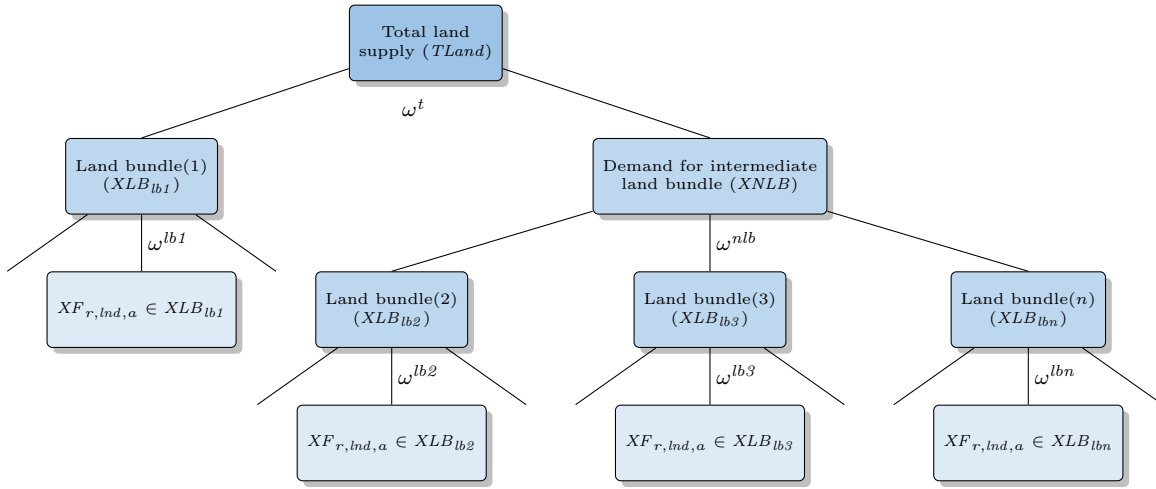


Figure 11: **CET nest for land allocation**

The top level nest decomposes total land supply into land bundle 1 and the intermediate land bundle. Equations (F-26) and (F-27) provide the first order conditions and hold irrespective of the CET specification used—either the standard, with **ifLandCET** set to 1, or the adjusted, volume preserving, CET specification. Equation (F-28) defines the aggregate price index of land,  $PTLandN$ . The index expression depends on the choice of CET specification. For the standard CET, the price index will be identical to the average price (defined next). For the adjusted CET, the price index will not generally be equal to the average price. Note that the price index expressions

<sup>50</sup> These options are defined in the **model.gms** file with a set named **landAssumptions**. The subset named **tass** contains the user-chosen specification.

<sup>51</sup> The implementation was first developed to mimic the CET structure of the land allocation specification in the MAGNET model.

<sup>52</sup> To mimic MAGNET's specification, **v-f**, **ocr-a**, **oap-a** are mapped to land bundle 1, **sug**, **ctl** and **rmk** are mapped to land bundle 2 and **ric**, **wht**, **gro** and **osd** are mapped to land bundle 2. The MAGNET model has two intermediate bundles and only three elasticities. In this specification, there are four intermediate bundles. Thus the top elasticity and the elasticity for land bundle 1 (horticulture, etc.), in MAGNET, are identical (and set to  $\sigma^1$  in the MAGNET documentation, see [edepot.wur.nl/310764](http://edepot.wur.nl/310764), page 74). The elasticity for the intermediate land bundle and the second land bundle (pasture, etc.) would also be identical (and set to  $\sigma^2$ ). And the third land bundle (cereals) would have the elasticity given by  $\sigma^3$ .

are not incorporated in the model under the assumption of perfect transformation, i.e. when the transformation elasticity is infinite. The average price of aggregate land,  $PTLand$  is given in equation (F-29) and holds for both versions of the CET.

$$\begin{cases} XLB_{r,lb1} = \gamma_{r,lb1}^{lb} \left( \frac{PLB_{r,lb1}}{PTLandN_r} \right)^{\omega_r^t} TLand_r & \text{if } \omega_r^t \neq \infty \\ PLB_{r,lb1} = PTLand_r & \text{if } \omega_r^t = \infty \end{cases} \quad (\text{F-26})$$

$$\begin{cases} XNLB_r = \gamma_r^{nlb} \left( \frac{PNLB_r}{PTLandN_r} \right)^{\omega_r^t} TLand_r & \text{if } \omega_r^t \neq \infty \\ PNLB_r = PTLand_r & \text{if } \omega_r^t = \infty \end{cases} \quad (\text{F-27})$$

$$\begin{cases} PTLandN_r = \left[ \gamma_{r,lb1}^{lb} PLB_{r,lb1}^{1+\omega_r^t} + \gamma_r^{nlb} PNLB_r^{1+\omega_r^t} \right]^{\frac{1}{1+\omega_r^t}} & \text{if } ifLandCET \\ PTLandN_r = \left[ \gamma_{r,lb1}^{lb} PLB_{r,lb1}^{\omega_r^t} + \gamma_r^{nlb} PNLB_r^{\omega_r^t} \right]^{\frac{1}{\omega_r^t}} & \text{if } \neg ifLandCET \end{cases} \quad (\text{F-28})$$

$$PTLand_r TLand_r = PLB_{r,lb1} XLB_{r,lb1} + PNLB_r XNLB_r \quad (\text{F-29})$$

The second level nest decomposes the intermediate land bundle (if it is defined) into a number of user determined land bundles, excluding the first land bundle. Equation (F-30) defines the demand for the land bundles as a share of the intermediate land bundle.<sup>53</sup> The price index of the intermediate land bundle is defined in equation (F-31), allowing for both CET specifications and is only included if the transformation elasticity at this nest is finite. The average price of the land bundle is defined in equation (F-32).

$$\begin{cases} XLB_{r,lb} = \gamma_{r,lb}^{lb} \left( \frac{PLB_{r,lb}}{PNLBN_r} \right)^{\omega_r^{nlb}} XNLB_r & \text{if } \omega_r^{nlb} \neq \infty \\ PLB_{r,lb} = PNLB_r & \text{if } \omega_r^{nlb} = \infty \end{cases} \quad (\text{F-30})$$

$$\begin{cases} PNLBN_r = \left[ \sum_{\{lb: lb \neq lb1\}} \gamma_{r,lb}^{lb} PLB_{r,lb}^{1+\omega_r^{nlb}} \right]^{\frac{1}{1+\omega_r^{nlb}}} & \text{if } ifLandCET \\ PNLBN_r = \left[ \sum_{\{lb: lb \neq lb1\}} \gamma_{r,lb}^{lb} PLB_{r,lb}^{\omega_r^{nlb}} \right]^{\frac{1}{\omega_r^{nlb}}} & \text{if } \neg ifLandCET \end{cases} \quad (\text{F-31})$$

$$PNLB_r XNLB_r = \sum_{\{lb: lb \neq lb1\}} PLB_{r,lb} XLB_{r,lb} \quad (\text{F-32})$$

The bottom land nest(s) decompose each of the land bundles into the primary land allocation to each activity  $a$ —each activity  $a$  (that uses land) must be mapped to one, and only one, land bundle. Equation (F-33) defines the supply of land to activity  $a$  that has been mapped to land bundle  $lb$ . Equation (F-34) defines the price index for the bundles that depends on the CET specification. And equation (F-35) defines the average price of the land bundles.

$$\begin{cases} Land_{r,a}^s = \gamma_{r,a}^t \left( \frac{PF_{r,land,a}}{PLBN_{r,lb}} \right)^{\omega_{r,lb}^{lb}} XLB_{r,lb} & \text{if } \omega_{r,lb}^{lb} \neq \infty \quad a \in \{lb\} \\ PF_{r,land,a} = PLB_{r,lb} & \text{if } \omega_{r,lb}^{lb} = \infty \end{cases} \quad (\text{F-33})$$

<sup>53</sup> The index  $lb$  excludes the first land bundle.

$$\begin{cases} PLBN_{r,lb} = \left[ \sum_{a \in \{lb\}} \gamma_{r,a}^t PF_{r,ln,d,a}^{1+\omega_{r,lb}^{lb}} \right]^{\frac{1}{1+\omega_{r,lb}^{lb}}} & \text{if } ifLandCET \\ PLBN_{r,lb} = \left[ \sum_{a \in \{lb\}} \gamma_{r,a}^t PF_{r,ln,d,a}^{\omega_{r,lb}^{lb}} \right]^{\frac{1}{\omega_{r,lb}^{lb}}} & \text{if } \neg ifLandCET \end{cases} \quad (F-34)$$

$$PLB_{r,lb} XLB_{r,lb} = \sum_{a \in \{lb\}} PF_{r,ln,d,a} Land_{r,a}^s \quad (F-35)$$

The final equation in the land market module, equation (F-36) equates land supply to activity  $a$  to its respective demand thus determining the equilibrium and sector-specific price for land,  $PF_{ln,d}$ . The equation is substituted out of the model specification.

$$Land_{r,a}^s = XF_{r,ln,d,a} \quad (F-36)$$

### 3.8.4 Market for natural resources

Natural resources are assumed to be activity-specific, for example crude oil reserves. The supply of natural resources,  $XNRS^s$ , is governed by a kinked iso-elastic supply curve. The model allows the supply response to differ according to whether market conditions are leading to increased supply, the 'hi' elasticity, versus a contracting market, the 'lo' elasticity. Ideally, we could implement this with a regime shift using continuous functions, but discontinuous derivatives. Instead, we add an equation to the model that determines the supply elasticity using the sigmoid function, also known as a logistic function.<sup>54</sup> The sigmoid function is defined by:

$$f(x) = \frac{1}{1 + e^{-x}}$$

Equation (F-37) defines the applied elasticity as a function of market conditions. The user enters the 'lo' and 'hi' elasticities in the parameters  $\eta^{nrsx}$ . One would typically assume that the 'lo' elasticity is higher than the 'hi' elasticity, i.e. it is harder to expand output than to contract output. If the two are identical, the equation is trivial. The argument in the sigmoid function reflects market growth and will be positive for expansion and negative for contraction. The  $\kappa$  parameter captures the 'speed' of adjustment between the 'lo' and 'hi' elasticities. The greater the value, the more rapid the transition. We have used a value of 30 with few numerical problems.<sup>55</sup> Figure 12 shows the kinked elasticity and the resulting supply curve as a function of price. It is centered at a price of 1. The 'lo' elasticity is assumed to be 2 and the 'hi' elasticity is 0.5. The 'kink' parameter is been set at 10 in this example.

Equation (F-38) reflects the supply function for the natural resource. For finite elasticities, it is an iso-elastic function, however, where the supply elasticity is determined by market conditions as described above. It is a function of the growth of the real price of the natural resource and the supply curve shifts according to lagged supply.<sup>56</sup> The (finite) supply curve has two shifters that can be used to target either the regional output price of the commodity being produced or the global price. To target the regional price, exogenize the output price,  $PX$ , and endogenize

<sup>54</sup> This is used for example in the CAPRI model of global agriculture with a focus on the EU, Britz et al. (2007).

<sup>55</sup> The sigmoid function is an intrinsic function in GAMS.

<sup>56</sup> This is somewhat different from the previous specification where the supply shifter was constant unless endogenous for calibration purposes.

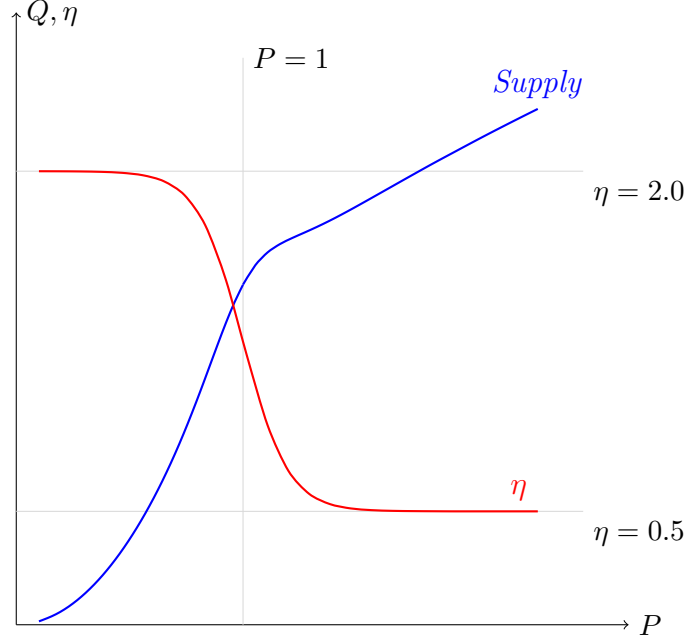


Figure 12: Example of a kinked supply curve using the sigmoid function

the regional shifter,  $\chi^{nrs}$ . To target the global price, exogenize the average producer price,  $PW$ ,<sup>57</sup> and endogenize the global shifter,  $\chi^{w,nrs}$ . The supply curve can also be horizontal, in which case the supply price is fixed relative to an anchor price, herein the GDP deflator. Equation (F-39) determines the equilibrium price. The latter equation is substituted out of the model.

$$\eta_{r,a,t}^{nrs} = \eta_{r,a,lo}^{nrsx} + \frac{\eta_{r,a,hi}^{nrsx} - \eta_{r,a,lo}^{nrsx}}{1 + \exp\left(-\kappa \left[\frac{XF_{r,nrs,a,t}}{XF_{r,nrs,a,t-1}} - 1\right]\right)} \quad (\text{F-37})$$

$$\begin{cases} XNRS_{r,a,t}^s = \chi_a^{w,nrs} \chi_{r,a}^{nrs} XNRS_{r,a,t-1}^s \left( \frac{\chi_{r,a}^{nrsp} PF_{r,nrs,a,t} / PGDPMP_{r,t}}{PF_{r,nrs,a,t-1} / PGDPMP_{r,t-1}} \right)^{\eta_{r,a}^{nrs}} & \text{if } \eta_{r,a}^{nrs} \neq \infty \\ \chi_{r,a}^{nrsp} PF_{r,nrs,a} = PGDPMP_r & \text{if } \eta_{r,a}^{nrs} = \infty \end{cases} \quad (\text{F-38})$$

$$XNRF_{r,a}^s = XF_{r,nrs,a} \quad (\text{F-39})$$

### 3.8.5 Water market

The market for water is somewhat similar to the land market. An aggregate supply curve is provided—and aggregate supply is then allocated to different uses using a nested CET structure. However, unlike land, water is in direct use in only the irrigated crop sectors. Demand for water in other sectors is specified using aggregate demand functions.

The aggregate water supply curve is allowed to have four shapes—constant elasticity, a logistic curve with an upward asymptote, a generalized hyperbola also with an upward asymptote, and

<sup>57</sup> Defined in equation (M-31).

perfectly horizontal. The aggregate water supply curve is given by equation (F-40). The initial elasticity of aggregate water is given by  $\eta^w$ .

$$\left\{ \begin{array}{ll} TH2O_r = \chi_r^{h2o} \left( \frac{PTH2O_r}{PGDP_r} \right)^{\eta_r^w} & \text{iso-elastic} \\ TH2O_r = \frac{TH2O_r^{Max}}{1 + \chi_r^{h2o} e^{-\gamma_r^{tw} (PTH2O_r / PGDP_r)}} & \text{logistic} \\ TH2O_r = TH2O_r^{Max} - \chi_r^{h2o} \left( \frac{PTH2O_r}{PGDP_r} \right)^{-\gamma_r^{tw}} & \text{hyperbola} \\ PTH2O_r = PGDP_r & \text{if } \eta_r^w = \infty \end{array} \right. \quad (F-40)$$

The iso-elastic curve by definition has a constant elasticity given by  $\eta^w$ . The calibration and other formulas for the logistic and hyperbola functions are the same as the land market expressions and are not repeated here.

Water use at the sub-aggregate level is represented by a number of 'bundles',  $H2OBnd$ . Table 3.4 lists the bundles and their relationships. The marketed water bundles are allocated in two CET nests. The top nest allocates water between agriculture and non-agriculture. These bundles are indexed by  $wb1$ . The second level nests include crops and livestock linked to the agriculture bundle, and industrial and municipal water use linked to the non-agriculture bundle. The second level bundles are indexed by  $wb2$ . The second level bundles are also segmented into two groups— $wba$  and  $wbi$ . The former allocates water across different activities—for example irrigated crops—and the latter is linked to an aggregate water demand—not explicitly linked to specific activities. The set  $wbx$  refers to exogenous water demand.

Table 3.4: Aggregate water bundles

Name	Description	Set	Dependency	Type
ENV	Environmental services		$TH2O$	$wbx$
GRD	Groundwater recharge		$TH2O$	$wbx$
AGR	All agriculture	$wb1$	$TH2O^m$	
CRP	Crops	$wb2$	AGR	$wba$
LVS	Livestock	$wb2$	AGR	$wbi$
NAG	Not agriculture	$wb1$	$TH2O^m$	
IND	Industrial water use	$wb2$	NAG	$wbi$
MUN	Municipal water use	$wb2$	NAG	$wbi$

At the top level water supply is allocated between marketed water and non-marketed water. The latter is exogenous and split into two variables—water for environmental services and water required for ground water recharge. In equilibrium, one would assume the latter is close to zero, i.e. ground water withdrawal closely matches ground water extraction. In the current database, both of these exogenous water uses are initially set at zero. A policy shock may include making these positive in a future time period—perhaps permanently for environmental services and temporarily for ground water recharge. Marketed water is then the difference between total water supply and exogenous demand. Equation (F-41) is the market supply equilibrium equation equating total water supply to marketed water plus the sum of all exogenous water demand. The variable  $H2O^m$  represents marketed water.

$$TH2O_r = TH2O_r^m + \sum_{wb \in wbx} H2OBnd_{r,wb} \quad (F-41)$$

Marketed water is allocated across the various water bundles using a nested (adjusted) CET specification.<sup>58</sup> Figure 13 illustrates the nested structure.

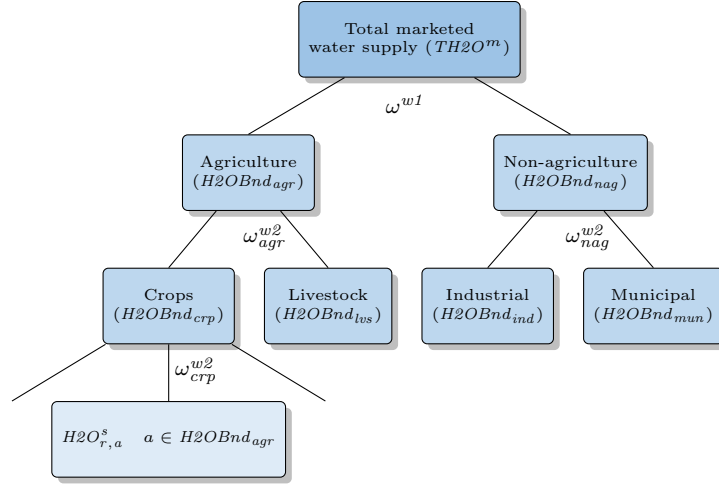


Figure 13: CET nest for water allocation

The first level nest allocates aggregate marketed water to the agricultural and non-agricultural bundles using an adjusted CET specification. The supply expressions are given in equation (F-42) that allows for perfect transformation. With reference to Table the index *wb1* covers the AGR and NAG water bundles.<sup>59</sup> If the transformation elasticity is finite, equation (F-43) provides the expression for the CET aggregate price index for  $TH2O^m$ . The aggregate water supply price is given by equation (F-44).

$$\begin{cases} H2OBnd_{r,wb1} = \gamma_{r,wb1}^{w1} \left( \frac{PH2OBnd_{r,wb1}}{PTH2O_r^n} \right)^{\omega_r^{w1}} TH2O_r^m & \text{if } \omega_r^{w1} \neq \infty \\ PH2OBnd_{r,wb1} = PTH2O_r & \text{if } \omega_r^{w1} = \infty \end{cases} \quad (\text{F-42})$$

$$PTH2O_r^n = \left[ \sum_{wb1} \gamma_{r,wb1}^{w1} PH2OBnd_{r,wb1}^{\omega_r^{w1}} \right]^{1/\omega_r^{w1}} \text{ if } \omega_r^{w1} \neq \infty \quad (\text{F-43})$$

$$PTH2O_r TH2O_r^m = \sum_{wb1} PH2OBnd_{r,wb1} H2OBnd_{r,wb1} \quad (\text{F-44})$$

The next nest decomposes the AGR and LVS bundles into crops (CRP) and livestock (LVS) on the one hand and industrial (IND) and municipal (MUN) water demand on the other hand. Equation (F-45) determines the supply for the second level bundles—indexed by *wb2*. Each of the *wb2* bundles is mapped to one, and only one, top level bundle indexed by *wb1*. Equation (F-46) defines the relevant aggregate price index that is used to determine the allocation. It is only used in the case of a finite transformation elasticity. Equation (F-47) defines the price of the aggregate *wb1* bundles, i.e. the price of the AGR and NAG bundles.

$$\begin{cases} H2OBnd_{r,wb2} = \gamma_{r,wb2}^{w2} \left( \frac{PH2OBnd_{r,wb2}}{PH2OBnd_{r,wb1}^n} \right)^{\omega_{r,wb1}^{w2}} H2OBnd_{r,wb1} & \text{if } \omega_{r,wb1}^{w2} \neq \infty \\ PH2OBnd_{r,wb2} = PH2OBnd_{r,wb1} & \text{if } \omega_{r,wb1}^{w2} = \infty \end{cases} \quad \text{wb2} \in \text{wb1} \quad (\text{F-45})$$

<sup>58</sup> The adjusted CET specification preserves quantity additivity—see Appendix A.

<sup>59</sup> The GAMS code is written in a relatively flexible way so that the user can largely change the set definitions for all water bundles and the relevant subsets and mappings.

$$PH2OBnd_{r,wb1}^n = \left[ \sum_{wb2 \in wb1} \gamma_{r,wb2}^{w2} PH2OBnd_{r,wb2}^{\omega_{r,wb1}^{w2}} \right]^{1/\omega_{r,wb1}^{w2}} \quad \text{if } \omega_{r,wb1}^{w2} \neq \infty \quad (\text{F-46})$$

$$PH2OBnd_{r,wb1} H2OBnd_{r,wb1} = \sum_{wb2 \in wb1} PH2OBnd_{r,wb2} H2OBnd_{r,wb2} \quad (\text{F-47})$$

The second level bundles, indexed by  $wb2$ , are segmented into two subsets. The first subset, indexed by  $wba$ , has an additional CET nest that maps the aggregate supply to individual production activities ( $a$ ). With the current version of the database, the AGR bundle is the only bundle that maps to water demand by production activities. The water bundles in the second subset, indexed by  $wbi$ , are set equal to an aggregate demand for water over a relevant subset of activities mapped to that bundle.<sup>60</sup>

Equation (F-48) determines the supply of water to each activity  $a$  that is mapped to the aggregate water supply indexed by  $wba$  (currently limited to irrigated crops mapped to the AGR bundle).<sup>61</sup> Equation (F-49) defines the relevant aggregate price index that is used to determine the allocation. It is only used in the case of a finite transformation elasticity. Equation (F-50) defines the price of the aggregate  $wba$  bundle(s).

$$\begin{cases} H2O_{r,a}^s = \gamma_{r,a}^{w3} \left( \frac{PF_{r,wat,a}}{PH2OBnd_{r,wba}^n} \right)^{\omega_{r,wba}^{w2}} H2OBnd_{r,wba} & \text{if } \omega_{r,wba}^{w2} \neq \infty \\ PF_{r,wat,a} = PH2OBnd_{r,wba} & \text{if } \omega_{r,wba}^{w2} = \infty \end{cases} \quad a \in wba \quad (\text{F-48})$$

$$PH2OBnd_{r,wba}^n = \left[ \sum_{a \in wba} \gamma_{r,a}^{w3} (PF_{r,wat,a})^{\omega_{r,wba}^{w2}} \right]^{1/\omega_{r,wba}^{w2}} \quad \text{if } \omega_{r,wba}^{w2} \neq \infty \quad (\text{F-49})$$

$$PH2OBnd_{r,wba} H2OBnd_{r,wba} = \sum_{a \in wba} PF_{r,wat,a} H2O_{r,a}^s \quad (\text{F-50})$$

Equation (F-51) determines the equilibrium market price of water,  $PF_{wat}$ , delivered to activity  $a$ . Water demand is provided by equation (P-36). The equilibrium condition is substituted out of the model.

$$H2O_{r,a}^s = XF_{r,wat,a} \quad (\text{F-51})$$

For the aggregate water bundles indexed by  $wbi$ , demand is constructed as a function of the aggregate production volume of the constituent activities and includes a price sensitive component. Equation (F-52) defines demand for aggregate water,  $H2OBnd^d$ . Demand is a function of a volume index over the activities mapped to the water bundles—subject to an elasticity given by  $\eta^{h2ob}$ . In addition, demand responds to the real price of the relevant bundle. Equation (F-53) determines the equilibrium price for the aggregate bundle. The equilibrium expression is substituted out of the model.

$$H2OBnd_{r,wbi}^d = \frac{\alpha_{r,wbi}^{h2ob}}{\lambda_{r,wbi}^{h2ob}} \left( \frac{PH2OBnd_{r,wbi}^d}{PGDPMP_r} \right)^{-\epsilon_{r,wbi}^{h2ob}} \left[ \frac{\sum_{a \in wbi} PX_{r,a,0} XP_{r,a}}{\sum_{a \in wbi} PX_{r,a,0} XP_{r,a,0}} \right]^{\eta_{r,wbi}^{h2ob}} \quad (\text{F-52})$$

<sup>60</sup> This is an ad hoc treatment of water demand that generates an equilibrium price. However, the demand for water in the relevant production activities do not form part of the cost structure of these activities—unlike the case for irrigated crops.

<sup>61</sup> The same elasticity parameter,  $\omega^{w2}$  is used for all of the aggregate water bundles, irrespective of the nest level.

$$H2OBnd_{r,wbi} = H2OBnd_{r,wbi}^d \quad (\text{F-53})$$

### 3.8.6 Factor taxes and prices

Producers pay the market price for factors adjusted by factor taxes. Equation (F-54) determines the producers purchase price of factors for all factors of production. Equation (F-55) is the relevant equation for the price of capital across vintages.

$$PF_{r,f,a}^p = (1 + \tau_{r,f,a}^v) PF_{r,f,a} \quad (\text{F-54})$$

$$PK_{r,a,v}^p = (1 + \tau_{r,cap,a}^v) PK_{r,a,v} \quad (\text{F-55})$$

## 3.9 National accounts and model closure

### 3.9.1 National accounts

Definition of the national accounts would normally be part of the post-processing portion of a simulation, i.e. it is not necessarily part of the core definition of a CGE model. However, when running the model in dynamics, it is typically useful to have some of the key national accounting identities available for targeting purposes, for example to target the growth of per capita GDP, or a macro ratio such as the share of government or investment expenditures relative to GDP.

Equation (M-1) defines nominal GDP at market price,  $GDPMP$ . The expression uses the indicator  $QGDP$  that represents the value of GDP using prices and quantities evaluated at different periods—similar to the absorption indicator above:

$$\begin{aligned} QGDP_{r,tp,tq} = & \sum_{fd} \sum_i PA_{r,i,fd,tp}^a XA_{r,i,fd,tq} + \sum_m PD_{r,m,tp} XTT_{r,m,tq} \\ & + \sum_i \left\{ \sum_d PWE_{r,i,d,tp} XW_{r,i,d,tq}^s - \sum_s PWM_{s,i,r,tp} XW_{s,i,r,tq}^d \right\} \end{aligned}$$

i.e. it is the sum of domestic absorption, including the exports of international trade and transport services and the net trade balance for all other goods and services evaluated at their respective border prices.

$$GDPMP_{r,t} = QGDP_{r,t,t} \quad (\text{M-1})$$

Equation (M-2) defines the GDP at market price deflator,  $PGDPMP$ , which is calculated as the Fisher price index of GDP.<sup>62</sup> Equation (M-3) then defines the volume of GDP at market price.

$$PGDPMP_{r,t} = PGDPMP_{r,t-1} \sqrt{\frac{QGDP_{r,t,t-1}}{QGDP_{r,t-1,t-1}} \cdot \frac{QGDP_{r,t,t}}{QGDP_{r,t-1,t}}} \quad (\text{M-2})$$

$$RGDPMP_r = GDPMP_r / RGDPMP_r \quad (\text{M-3})$$

Equation (M-4) defines real per capita GDP,  $RGDPpc$ . Equation (M-5) defines the growth in real per capita GDP,  $g^y$ . In comparative static exercises, it measures the difference in real per capita

<sup>62</sup> When prices and quantities are evaluated at the same year, nominal GDP can replace the relevant  $QGDP$  indicator.



GDP between the baseline and the shock simulation.<sup>63</sup> In dynamic simulations the equation allows for multi-year time steps as measured by  $n$ . In most dynamic baseline or reference simulations, the growth rate of real per capita GDP will be exogenous and an economy-wide variable, for example a uniform labor productivity shifter, is endogenous and will serve as the instrument to target GDP growth. In policy or alternative simulations, the economy-wide factor would normally be exogenous and the growth rate of GDP would be endogenous. Equation (M-6) defines the average capital labor ratio in efficiency units. Under a balanced growth scenario, this ratio is intended to be uniform across time. It could be used to target the growth of capital productivity independently of the growth of labor productivity.

$$RGDPpc_r = \frac{RGDPMP_r}{Pop_r} \quad (M-4)$$

$$RGDPpc_{r,t} = (1 + g_{r,t}^y)^n RGDPpc_{r,t-n} \quad (M-5)$$

$$KLRat_r = \frac{\sum_a \sum_v PK_{r,a,v,0} \lambda_{r,cap,a}^f K v_{r,a,v}}{\sum_a \sum_l PF_{r,l,a,v,0} \lambda_{r,l,a}^f XF_{r,l,a}} \quad (M-6)$$

### 3.9.2 Closure

#### Government accounts

Equation (M-7) defines nominal government savings,  $S^g$ , as the difference between government revenues and government's current expenditures on goods and services. In the standard model, there is only one stream of government expenditures. Alternative versions allow for additional streams, such as expenditures on research and development, public investment, etc. The subset *Gov* includes all final demand expenditures linked to government accounts.<sup>64</sup> When an emissions quota regime is put into place, revenues from the quotas<sup>65</sup> accrue to the government and eventually to households using the default closure with a lump sum tax on households. Note that the counterpart account of the quota income appears in the balance of payments expression, i.e. as a flow from the rest of the world. The equation also includes net government to government transfers and the government's income from NTM-related measures—both domestic and foreign.<sup>66</sup> Equation (M-7) defines real government savings,  $RS^g$ . In the standard closure, the public deficit is held fixed in real terms (in order to preserve price homogeneity). The direct tax rate,  $\kappa^h$ , adjusts to the target for the public deficit. Because there is a single household, this is equivalent to a change in lump sum taxes.

$$\begin{aligned} S_r^g &= \sum_{gy} YGOV_{r,gy} + \sum_{em} Quota Y_{r,em}^{Emi} + \sum_s \chi_{s,r}^{gNTM} ntm Y_s \\ &+ ODAIn_r - ODAOut_r - \sum_{fd \in Gov} YFD_{r,fd} \end{aligned} \quad (M-7)$$

$$RS_r^g = \frac{S_r^g}{PGDPMP_r} \quad (M-8)$$

<sup>63</sup> The value of  $n$  is always 1 in comparative static simulations.

<sup>64</sup> [NEW: 20-JAN-2019] In the current model code, the subset *Gov* does not exist. Instead the sum is over all *fd* that equal either *gov* or *r\_d*.

<sup>65</sup> Revenues are positive if a region is a net seller and negative if a region is a net buyer.

<sup>66</sup> [NEW: 4-OCT-2018]

### Macro closure

Equations (M-9) and (M-10) define respectively the real and nominal share of aggregate domestic final demand expenditures with respect to GDP, respectively real and nominal. These equations are often used to calibrate the model in the dynamic baseline. For example, real government and research and development expenditures may be fixed relative to real GDP in the baseline. In shock simulations, they may be fixed at baseline levels, and then these equations determine how the shares change with shocks. For investment, the baseline may target the ratio of investment relative to GDP, with an endogenous savings rate. The savings rate would then be exogenous in shock simulations.

$$rFDShr_{r,fd} = \frac{XFD_{r,fd}}{RGDPMP_r} \quad (M-9)$$

$$nFDShr_{r,fd} = \frac{YFD_{r,fd}}{GDPMP_r} \quad (M-10)$$

### Capital account

The model has four closures for the capital account. The standard closure fixes foreign savings in real terms. Equation (M-11) determines  $S^f$ . For all but one region, the residual region, foreign saving is equal to a fixed level. The second part of the equation, guarantees that the sum across regions is identically zero, and thus sets the level for the residual region.<sup>67</sup>

$$\begin{cases} S_r^f = \bar{S}_r^f & \text{if } r \neq rres \\ \sum_r S_r^f \equiv 0 & \text{if } r = rres \end{cases} \quad (M-11)$$

In addition, equation (M-12) is added to the model specification in the case of the fixed capital account closure and simply evaluates  $R^g$  as a weighted average of regional expected returns, with the weights given in equation (M-13).

$$R^g = \sum_r \varphi_r^r R_r^e \quad (M-12)$$

$$\varphi_r^r = \frac{PFD_{r,inv} (XFD_{r,inv} - \delta_r TK_r^s)}{\sum_s PFD_{s,inv} (XFD_{s,inv} - \delta_s TK_s^s)} \quad (M-13)$$

The second alternative fixes the capital account relative to GDP, equation (M-14). Note that in this case the capital flows are deflated by the world price of savings since the variable  $S^f$  is a volume, not a value. The residual region becomes a borrower or lender of last resort as there is nothing that guarantees that if the ratio is fixed for all regions that their sum will add to zero. As in the case of the first closure, equation (M-12) is part of the model specification for this capital account closure.

$$\begin{cases} S_r^f = \chi_r^{sf} GDPMP_r / PW^{sav} & \text{if } r \neq rres \\ \sum_r S_r^f \equiv 0 & \text{if } r = rres \end{cases} \quad (M-14)$$

<sup>67</sup> It is best if the user makes sure that the sum of exogenous foreign savings equals 0 to avoid unintended consequences, but at least with this formulation, a small deviation will not create any problems.

The third alternative, inspired from the GTAP model (Hertel (1997)), allows foreign savings to adjust to relative differentials in expected rates of return—with a specific definition of what is meant by the expected rate of return.

We start with the definition of the expected end-of-period capital stock, equation (M-15). This is simply the standard motion equation for the capital stock based on this period's beginning-of-period capital stock and investment level.

$$TK_{r,t}^e = (1 - \delta_{r,t}) TK_{r,t} + XFD_{r,inv,t} \quad (\text{M-15})$$

Equation (M-16) defines the aggregate after-tax rate of return. The factor  $\chi^k$  converts the normalized rate of return to the non-normalized rate of return.<sup>68</sup> Equation (M-17) defines the net current rate of return, the rental rate adjusted for the price of replacing capital, less the depreciation rate. Therefore, a tariff reduction, that would lower the price of investment goods, would reduce the replacement price of capital and increase the rate of return (all else equal). The expected rate of return is given by equation (M-18). Regions with relatively high investment rates, i.e., that would lead to relatively high growth in the capital stock and thus depress the capital rate of return, would tend to see a reduction in the expected rate of return. The degree of adjustment is controlled by the parameter  $\epsilon^{ror}$ .

$$R_{r,t} = \frac{\sum_a \sum_v \left(1 - \kappa_{r,cap,a}^f\right) PK_{r,a,v} K_{r,a,v}^v}{PFD_{r,inv,t} TK_{r,t}^s} \quad (\text{M-16})$$

$$R_{r,t}^c = \frac{R_{r,t}}{PFD_{r,inv,t}} - \delta_{r,t} \quad (\text{M-17})$$

$$R_{r,t}^e = R_{r,t}^c \left[ \frac{TK_{r,t}^e}{TK_{r,t}} \right]^{-\epsilon^{ror}} \quad (\text{M-18})$$

The level of foreign saving for each region will then be determined by equation (M-19), which equates the expected regional rates of return to the global average,  $R^g$ , adjusted for any risk premium that is embedded in the base year data. Equation (M-20) guarantees that global savings equals global investment and in essence determines  $R^g$ . High values for  $\epsilon^{ror}$  lead to relatively low adjustment in foreign savings. The default value for the GTAP model is 10.<sup>69</sup>

$$\pi_{r,t}^p R_{r,t}^c = R_t^g \quad (\text{M-19})$$

$$\sum_r S_r^f \equiv 0 \quad (\text{M-20})$$

The fourth specification is inspired by the capital allocation mechanism in the Monash and USAGE models (Dixon and Rimmer (2002) and Dixon and Rimmer (2005)).<sup>70</sup> The MONASH uses a capital supply schedule that links the growth in capital (for each region),  $grK$ , to deviations in the region's rate of return,  $\Delta RoR$ , where the deviation is relative to an exogenous normal rate of return,  $R^n$ . A positive deviation leads to growth in capital above trend,  $grTrend$ . Growth is bounded by  $grKMin$  at the lower end and  $grKMax$  at the upper end. A logistic function describes

<sup>68</sup> In the base year,  $TR$  is set to 1, and the non-normalized return will be a percent rate between 0 and perhaps up to 30%.

<sup>69</sup> Users are urged to read the discussion in Hertel (1997), page 57.

<sup>70</sup> See also the discussion in the freely available Monash notes—Chapter 5, pages 317–326. The notes can be downloaded from <https://www.copsmodels.com/ftp/monbook1/m1-chap5.pdf> (accessed 28-May-2017).

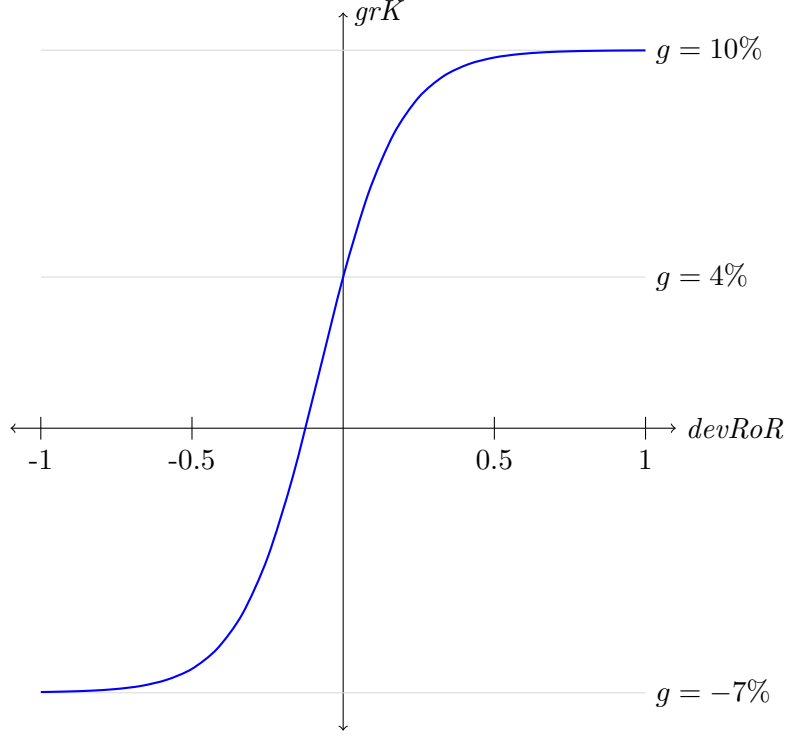


Figure 14: **Capital supply as a function of deviations in the rate of return**

the capital growth schedule where the curvature parameter,  $\chi^k$ , is calibrated at  $\Delta RoR = 0$  to an econometric estimate. Figure 14 shows the shape of the logistic curve with the growth of capital on the vertical axis and the deviation in the rate of return on the horizontal axis. The curve crosses at the normal or trend growth rate (4% in the graph) when the deviation in the rate of return is equal to its long-term average (i.e. 0%). The maximum growth in the example is 10% and the minimum is -7%.

We start by describing the expected rate of return. The expected rate of return is the sum of two components. The first is the expected nominal gross return in year  $t + 1$ ,  $R_{t+1}$ , relative to the unit cost of investing in year  $t$ ,  $PFD_{inv}$ . The second term represents the capital gains on the unit of investment, i.e. the difference in the unit cost of investing between  $t$  and  $t + 1$  (adjusted by depreciation). Both terms are discounted by the real rate of interest,  $r$ .

$$E(R_{r,t}^e) = \frac{1}{1+r} \left[ E\left(\frac{R_{r,t+1}}{PFD_{r,inv,t}}\right) + (1 - \delta_{r,t}) E\left(\frac{PFD_{r,inv,t+1}}{PFD_{r,inv,t}}\right) \right] - 1$$

With myopic expectations, the expected future values of  $R$  and  $PFD_{inv}$  are the current values and the expression above simplifies to equation (M-21), that defines the expected rate of return,  $R^e$ , at time  $t$  for region  $r$ . Equation (M-22) defines the deviation of the expected rate of return from its 'normal' or trend path,  $\Delta RoR$ , which is used to determine the desired growth rate of the stock of capital in region  $r$ , where the expected rate of return is given by  $R^n$  and is an exogenous input. If the deviation is 0, the desired growth rate of the stock of capital is the trend growth rate (see equation (M-23) below). The equation allows for region-specific,  $R^d$  and a global shifter,  $R^g$  that

<sup>71</sup> See equation (M-16).

allow for deviations from the normal trend.

$$R_{r,t}^e = \frac{1}{1+r} \left[ \frac{R_{r,t}}{PFD_{r,inv,t}} + (1 - \delta_{r,t}) \right] - 1 \quad (\text{M-21})$$

$$\Delta RoR_{r,t} = R_{r,t}^e - R_{r,t}^n - R_{r,t}^d - R_t^g \quad (\text{M-22})$$

Equation (M-23) defines the growth in the capital stock,  $grK$ , using the doubly bounded logistic curve as a function of the deviation in the real rate of return from its long-run average.<sup>72</sup> Equation (M-24) defines the demand for new capital, i.e. investment,  $XFD_{inv}$ , for each region  $r$ . With the standard motion equation that sets end-of-period capital stock equal to the depreciated level of the beginning-of-period capital stock plus investment and the notional growth in the capital stock as given by  $grK$ , 'desired' investment is equal to the beginning-of-period capital stock multiplied by the sum of the growth rate and depreciation. It is bounded below by zero since  $grKMin$  is set at minus the depreciation rate. Given that we also have equation (Y-20), this equation can be thought of as determining foreign savings. Equation (M-25) ensures the global savings/investment balance and in essence determines the  $R_t^g$ .<sup>73</sup>

$$grK_{r,t} = \frac{grKMax_{r,t} e^{\chi_{r,t}^k \Delta RoR_{r,t}} + grKMin_{r,t} \left[ \frac{grKMax_{r,t} - grKTrend_{r,t}}{grKTrend_{r,t} - grKMin_{r,t}} \right]}{e^{\chi_{r,t}^k \Delta RoR_{r,t}} + \left[ \frac{grKMax_{r,t} - grKTrend_{r,t}}{grKTrend_{r,t} - grKMin_{r,t}} \right]} \quad (\text{M-23})$$

$$\begin{aligned} K_{r,t+1} &= (1 - \delta_{r,t}) K_{r,t} + XFD_{r,inv,t} = (1 + grK_{r,t}) K_{r,t} \\ &\quad \Updownarrow \\ XFD_{r,inv,t} &= K_{r,t} (grK_{r,t} + \delta_{r,t}) \end{aligned} \quad (\text{M-24})$$

$$\sum_r S_r^f \equiv 0 \quad (\text{M-25})$$

The user provides values for  $grKMin$ ,  $grKMax$ ,  $grKTrend$ ,  $\chi^k$  and  $R^n$ . The variables  $grK$  and  $R^e$  are initialized to base year data. Equation (M-23) can be inverted to yield the initial level for  $\Delta RoR$ . And finally, equation (M-22) is used to initialize  $R^d$ .

The capital closure specification is governed by setting the `savfFlag`. The valid options are:

1. **capFix**. With this setting the capital account is fixed.
2. **capRFix**. With this setting the capital account is fixed relative to GDP for all but a residual region, which becomes the lender or borrower of last resort.
3. **capFlexGTAP**. With this setting the capital account is flexible and uses the GTAP specification.
4. **capFlexUSAGE**. With this setting the capital account is flexible and uses an allocation mechanism similar in spirit to the domestic allocation mechanism used in the USAGE model.

<sup>72</sup> Further discussion is available in [Dixon and Rimmer \(2002\)](#), pp. 190-193. The version described herein inverts the functional form, but is otherwise identical. The inverted function yields the following, which can be used to initialize  $\Delta RoR$ :

$$\Delta RoR_{r,t} = \frac{1}{\chi_{r,t}^k} \ln \left[ \left( \frac{grKMax_{r,t} - grKTrend_{r,t}}{grKTrend_{r,t} - grKMin_{r,t}} \right) \left( \frac{grK_{r,t} - grKMin_{r,t}}{grKMax_{r,t} - grK_{r,t}} \right) \right]$$

<sup>73</sup> Though  $R^g$  is used for all closures it has a different significance across closures. In the USAGE closure, one would anticipate that in the long-run it is 0.

### Price anchors

Equation (M-26) defines the so-called manufactures unit value index,  $PMUV$  and it is based on the Fisher index of export prices of selected goods (often manufactures) over selected exporters (often high-income regions). The  $QMUV$  indicator is based on the following expression:

$$QMUV_{tp,tq} = \sum_{r \in RMUV} \sum_{i \in IMUV} \sum_d PWE_{r,i,d,tp} XW_{r,i,d,tq}^s$$

and represents the an export indicator using time dependent prices and quantities.

$$PMUV_t = PMUV_{t-1} \sqrt{\frac{QMUV_{t,t-1}}{QMUV_{t-1,t-1}} \cdot \frac{QMUV_{t,t}}{QMUV_{t-1,t}}} \quad (M-26)$$

In a similar vein, the model has a factor price index—one per region—that can be used to target the real exchange rate. The regional factor price index is based on the following formula:

$$QFACT_{tp,tq} = \sum_f \sum_a PF_{r,f,a,tp} XF_{r,f,a,tq}$$

Equation (M-27) defines the regional factor price index. It is the Fisher price index of regional factor prices aggregated over all factors and activities.<sup>74</sup>

$$PFACT_{r,t} = PFACT_{r,t-1} \sqrt{\frac{QFACT_{t,t-1}}{QFACT_{t-1,t-1}} \cdot \frac{QFACT_{t,t}}{QFACT_{t-1,t}}} \quad (M-27)$$

Foreign savings flows,  $S^f$ , are determined in real terms and are evaluated at the price  $PW^{sav}$ , that is set equal to  $PMUV$ , equation (M-29). The model requires at least one price to be fixed and treated as the numéraire. Equation (M-30) defines the numéraire,  $PNUM$ . Herein it is set equal to a world GDP price deflator, but it could easily be set to  $PMUV$ . The left-out equation, as discussed above, is one of the investment/savings balance equations. A variable called *Walras* has been added to the investment/savings balance equation for a single region. It should evaluate to zero.

$$PWGDP_t = \frac{\sum_r GDPMP_t}{\sum_r RGDPMPT_t} \quad (M-28)$$

$$PW^{sav} = PMUV \quad (M-29)$$

$$PNUM = PWGDP \quad (M-30)$$

In summary, the standard macro closure has household savings determined by the ELES demand system, government savings are fixed (in real terms) with the direct household tax schedule endogenous, and foreign savings are fixed (with the real exchange rate adjusting to *ex ante* changes in the trade balance. Investment is savings determined. The model numéraire is the export price index of high-income manufactured exports.

<sup>74</sup> We have done some limited testing with fixing the real exchange rate—using the factor price index as a measure of the real exchange rate. If the real exchange rate is fixed, then the capital account must adjust. In this case, the set `fixER` must be set to 'YES' for the region with the fixed exchange rate (under the 'capFix' capital account closure). In this case the region will be excluded from the capital account closure equation and the capital account adjustments will be assumed by the residual region. This closure should be used with the appropriate caution.

### 3.9.3 Targeting prices

It is often the case, particularly in developing a baseline, that it is desirable (or required) to target price trends for key commodities, for example crude oil. The first step is to define what is meant by a 'world' price. The current version of the model defines the 'world' price to be the average producer price—defined using a Fisher price index. The key reason is that typically the commodities to be targeted have a sector-specific resource base that can be adjusted to target an exogenous price.<sup>75</sup> Equation (M-31) defines the world price for activity  $a$ . It is based on the following indicator:

$$QX_{a,tp,tq} = \sum_r PX_{r,a,tp} XP_{r,a,tq}$$

$$PW_{a,t} = PX_{a,t-1} \sqrt{\frac{QX_{a,t,t-1}}{QX_{a,t-1,t-1}} \cdot \frac{QX_{a,t,t}}{QX_{a,t-1,t}}} \quad (\text{M-31})$$

To target a price, the variable  $PW$  needs to be exogenized and a counterpart needs to be endogenized. The most logical choice is to endogenize the shifter on the natural resource supply curve,  $\chi_a^{w,nrs}$ .

### 3.9.4 Regional and global welfare

[NEW 20-Jan-2019] This section will be revised in the future.

#### Equivalent variation

Equivalent variation is defined to be the level of income needed to attain utility level  $u$  at base year prices. The formulation depends on the specific utility function—and in many cases will need to be derived using a solver given implicit functional forms. For the CDE utility function, equivalent variation,  $EV$ , is defined implicitly by the following expression:

$$\sum_k \alpha_{r,k,h}^h b_{r,k,h}^h u_{r,h}^{e_{r,k,h}^h b_{r,k,h}^h} \left( \frac{PC_{r,k,h,0}}{EV_{r,h}/Pop_r} \right)^{b_{r,k,h}^h} = 1 \quad (\text{M-32})$$

For the AIDADS class of utility functions, which includes the LES and Cobb-Douglas, the expression for equivalent variation is provided by the following:

$$\ln \left( \frac{EV_{r,h}}{Pop_r} - \sum_k PC_{r,k,h,0} \gamma_{r,k,h} \right) = 1 + \ln(A_{r,h}^{ad}) + u_{r,h} + \sum_k \mu_{r,k,h}^c \ln \left( \frac{PC_{r,k,h,0}}{\mu_{r,k,h}^c} \right) \quad (\text{M-33})$$

Evaluation of  $EV$  is incorporated in the model code and is a model output.

Utility can also be derived from other final demand expenditures. These all rely on some form of a CES expenditure function, for which the expenditure function is relatively straightforward. Equation (M-34) defines the equivalent variation for each of the other final demand expenditure components—the real value of current expenditures, times the initial aggregate price. Equation (M-35) defines the expenditure function linked to domestic savings. In the absence of capital flows, it will be equal to the expenditure function of investment. The GTAP model (McDougall (2003) and Corong et al. (2017)) defines total equivalent variation as the sum of equivalent variation of households, public expenditures and savings. In the GREEN Model (Burniaux et al. (1992)),

<sup>75</sup> With a diagonal make matrix, targeting the producer price and the commodity price would be identical.

total equivalent variation is equated to private and public expenditures in each year. These are discounted to the base year to evaluate the net discounted change in equivalent variation and used as one metric to summarize the net impacts of a policy change. In GREEN, savings (or investment) are not included in the definition of welfare, as the welfare from savings are derived in future periods (as increments to private or public consumption). This does not take into consideration the utility derived from savings in periods beyond the model's time horizon.

$$EVF_{r,fdc} = PFD_{r,fdc,0} \frac{YFD_{r,fdc,t}}{PFD_{r,fdc,t}} \quad (\text{M-34})$$

$$EVS_r = PSAVE_{r,0} \frac{\sum_h S_{r,h,t}^h + S_{r,t}^g}{PSAVE_{r,t}} \quad (\text{M-35})$$

The price of savings formula is derived from an adaptation of the price of savings from the GTAP model ([van der Mensbrugghe \(2018\)](#)). In the absence of capital flows, the price of savings is equal to the price of investment. Equation (M-36) adjusts the definition to take into account foreign savings.

$$PSAVE_{r,t} = \chi_t^s \frac{PSAVE_{r,t-1}}{PFD_{r,inv,t-1}} PFD_{r,inv,t} \quad (\text{M-36})$$

The adjustment factor  $\chi^s$  is updated each period and is given by the following expression:

$$\chi_t^s = \frac{\sum_r \varphi_{r,t}^i \frac{PFD_{r,inv,t}}{PFD_{r,inv,t-1}}}{\sum_r \varphi_{r,t}^s \frac{PSAVE_{r,t}}{PSAVE_{r,t-1}}}$$

where  $\varphi^i$  represents the regional share in global net investment:

$$\varphi_{r,t}^i = \frac{PFD_{r,inv,t-1} (XFD_{r,inv,t-1} - \delta_{r,t-1} KStock_{r,t-1})}{\sum_s PFD_{s,inv,t-1} (XFD_{s,inv,t-1} - \delta_{s,t-1} KStock_{s,t-1})}$$

and  $\varphi^s$  represents the regional share in global savings:

$$\varphi_{r,t}^s = \frac{\sum_h S_{r,h,t-1}^h - S_{r,t}^g}{\sum_s \left( \sum_h S_{s,h,t-1}^h - S_{s,t}^g \right)}$$

### Global welfare function

The model includes an objective function, which for most simulations is redundant as the model is intended to be square with a unique solution. In principle, it would be possible to run the model with more variables than equations and to use an optimizing package to maximize an objective, with all model equations acting as constraints. In the current version, the social welfare function is a CES aggregate of per capita real income, as measured by equivalent variation, that is population weighted and with welfare weights—currently initialized at the value of 1. For a value of  $\epsilon^w = 0$  the social welfare function is simply the average global level of real income. The objective function in any given time period is set to the level of the social welfare function. The model also calculates a two additional global social welfare functions: (1) one that takes into account the utility derived



from public expenditures, equation (M-38); and (2) one consistent with the GTAP definition of welfare that includes private and public expenditures and domestic savings, equation (M-39).

$$SW = \sum_r \frac{\varphi_r Pop_r}{1 - \epsilon^w} \left( \frac{\sum_h EV_{r,h}}{Pop_r} \right)^{1-\epsilon^w} \bigg/ \sum_r Pop_r \quad (M-37)$$

$$SWT = \sum_r \frac{\varphi_r Pop_r}{1 - \epsilon^w} \left( \frac{\sum_h EV_{r,h} + EVF_{r,gov}}{Pop_r} \right)^{1-\epsilon^w} \bigg/ \sum_r Pop_r \quad (M-38)$$

$$SWT2 = \sum_r \frac{\varphi_r Pop_r}{1 - \epsilon^w} \left( \frac{\sum_h EV_{r,h} + EVF_{r,gov} + EVS_r}{Pop_r} \right)^{1-\epsilon^w} \bigg/ \sum_r Pop_r \quad (M-39)$$

## 3.10 Emissions

### 3.10.1 Emissions

The model incorporates a number of greenhouse gas (GHG) emissions and non-greenhouse gas emissions, indexed by *em*.

The standard database includes the following greenhouse gas emissions: carbon dioxide (CO<sub>2</sub>), methane (CH<sub>4</sub>), nitrous oxide (N<sub>2</sub>O) and an aggregate emissions bundle of fluoridated gases (F-gases). The input data for carbon is in millions of metric tons of CO<sub>2</sub> and the data for the other GHGs are both in physical units (metric tons) as well as in CO<sub>2</sub>-equivalent.<sup>76</sup> The model allows for conversion between CO<sub>2</sub>*e* and *Ce*, i.e. carbon equivalent, depending on the needs of the user.<sup>77</sup>

The non-greenhouse gases are described in Table 3.5.<sup>78</sup> These emissions have no corresponding 'carbon equivalent' component, though they can impact on climate and are an input to many climate models, for example MAGICC.<sup>79</sup> They can also have an impact on morbidity and mortality—see for example Chepeliev and van der Mensbrugghe (2017) and OECD (2016).

<sup>76</sup> The conversion from tons to CO<sub>2</sub>-equivalent uses the standard global warming potential (gwp) coefficients defined by the Intergovernmental Panel on Climate Change (IPCC).

<sup>77</sup> The *ifCEQ* flag determines the units. If set to 1, the model will use *Ce*, else CO<sub>2</sub>*e*. The main impact will be on the level of any endogenously determined carbon tax. The factor (44/12) is used to convert from *Ce* to CO<sub>2</sub>*e* in volume terms. The inverse ratio is used to convert prices and taxes.

<sup>78</sup> The non-greenhouse gas database has been compiled by Maksym Chepeliev at the Center and incorporated in ENVISAGE in May 2018. See Chepeliev (2018).

<sup>79</sup> Meinshausen et al. (2011)

Table 3.5: **Non-greenhouse gas emissions**

<i>Gas</i>	<i>Description</i>
BC	Black carbon
CO	Carbon monoxide
NH <sub>3</sub>	Ammonia
NMVB	Non-methane volatile organic compounds (short carbon cycle)
NMVF	Non-methane volatile organic compounds (long carbon cycle)
NO <sub>x</sub>	Nitrogen oxides
OC	Organic carbon
PM10	Particulate matter 10
PM2.5	Particulate matter 2.5
SO <sub>2</sub>	Sulfur dioxide

Emissions are generated by three sources: 1) direct consumption of a commodity; 2) factor-based emissions (e.g. capital, i.e. herds, in the livestock sectors); and 3) output or processed based emissions (e.g. methane from landfills). Carbon emissions in the current model version are only generated by the combustion of fossil fuels—coal, oil (crude and refined) and natural gas. Other emissions can be a combination of all three sources of emissions. Equation (E-1) reflects the consumption based emissions over all Armington agents ( $aa$ ) and for all commodities ( $i$ ). Emissions are based on the consumption level multiplied by the rate of emission per unit of consumption ( $\rho^{Emi}$ ). The specification allows for both national- and agent-specific sourcing of imports. In the case of the latter, the relevant volume of consumption is  $XA$ . In the case of the latter, the relevant consumption volumes are  $XD$  and  $XM$  and the emission rates are allowed to vary across bundles.<sup>80</sup> Equation (E-2) reflects the factor based emissions for labor, capital, land and natural resources. Finally, equation (E-3) represents the processed based emissions.<sup>81</sup>

$$Emi_{r,em,i,aa} = \begin{cases} \chi_{em}^{Emi} \rho_{r,em,i,aa}^{Emi} XA_{r,i,aa} & \text{if } ArmFlag = 0 \\ \chi_{em}^{Emi} \rho_{r,em,i,aa}^{Emi,d} XD_{r,i,aa} + \chi_{em}^{Emi} \rho_{r,em,i,aa}^{Emi,m} XM_{r,i,aa} & \text{if } ArmFlag \neq 0 \end{cases} \quad (E-1)$$

$$Emi_{r,em,f,a} = \chi_{em}^{Emi} \rho_{r,em,f,a}^{Emi} XF_{r,f,a} \quad (E-2)$$

$$Emi_{r,em,tot,a} = \chi_{em}^{Emi} \rho_{r,em,tot,a}^{Emi} XP_{r,a} \quad (E-3)$$

Equation (E-4) is the aggregate level of emissions for region  $r$ . It is the sum across all potential sources of emissions represented by the indices  $is$  and  $aa$ . The equation allows for a region specific level of emissions from exogenous sources—that could be used for example to introduce forest-based emissions. Equation (E-5) defines the global level of emissions, the sum of total regional emissions, and also allows for an exogenous level of emissions that is not region-specific.

$$EmiTot_{r,em} = \sum_{is} \sum_{aa} Emi_{r,em,is,aa} + EmiOth_{r,em} \quad (E-4)$$

<sup>80</sup> The CO<sub>2</sub> emissions database distinguishes between domestic and import emissions, however the non-CO<sub>2</sub> database does not and the emission rates are assumed to be identical.

<sup>81</sup> There is a master set,  $is$ , that represents all of the labels of the SAM—including one for the total that is represented by the subset  $tot$ .

$$EmiGbl_{em} = \sum_r EmiTot_{r,em} + EmiOthGbl_{em} \quad (E-5)$$

### 3.10.2 Emission price regimes

The model exhibits a degree of flexibility in implementing carbon price regimes. The simplest option is to set an emissions tax exogenously and all the required equations have already been described.<sup>82</sup> If the model is being solved in terms of CO<sub>2</sub>, a tax of \$x/TCO<sub>2</sub> is implemented as:

```
emiTax.fx(r,"co2",t) = cscale*x ;
```

If the model is being solved in terms of carbon, the same tax would need to be specified as:<sup>83</sup>

```
emiTax.fx(r,"co2",t) = cscale*x*(44/12) ;
```

The formulas would need adjustment if the exogenous carbon tax is implemented in terms of tons of carbon instead of tons of CO<sub>2</sub>. The rest of this section describes an alternative carbon price regime that relies on an emissions cap, with and without trading.

An emissions cap requires two additional equations—one that calculates the carbon tax for all regions within a coalition of regions (that may be as small as one region). The second equation maps the uniform coalition-wide carbon tax the regional carbon tax. Equation (E-6) defines the emissions cap constraint. Coalitions are represented by the index  $rq$ . This index is a subset of  $ra$  that represents a user-defined set of aggregate regions that can be used by the model (and the output routines). It normally includes all of the separate regions of the model—so that each modeled region can be its own coalition—plus a number of additional aggregate regions to which the individual regions are mapped. Coalitions are then setup by defining the subset  $ra$  to which the regional mappings will already have been specified. The emissions cap equation then implicitly determines the coalition-wide emissions tax,  $\tau_{rq,em}^{EmiQ}$  that is compatible with the emissions constraint. The tax will be calculated for each coalition. Equation (E-7) maps the coalition-wide tax to the emissions tax of each individual region that is part of the  $rq$  coalition. In the absence of a permit trading regime, the emission tax revenues collected in each region are fully allocated to the regional government.

$$\sum_{r \in \{rq\}} EmiTot_{r,em} = \chi_{em}^{Cap} Emi_{rq,em}^{Cap} \quad (E-6)$$

$$\tau_{r,em}^{Emi} = \tau_{rq,em}^{EmiQ} \quad \text{if } r \in \{rq\} \quad (E-7)$$

The following code snippet shows how to implement a cap on emissions for two coalitions—one composed of developing countries indexed by "lmy" and one composed of developed countries indexed by "hic". The first line sets the `ifEmiCap` flag. The second line sets the `emFlag` flag to indicate a tax on carbon only. Lines 3 and 4 define the two coalitions. The fifth line sets the cap for each of the two coalitions. The caps represent a 20 percent reduction in aggregate carbon emissions relative to the level of emissions of the previous period. Assuming all regions are mapped to the two coalition regions, this will also reduce global emissions by 20 percent. The final four lines are

<sup>82</sup> This version of the model is really only set up to deal with carbon taxes. The effect of an emission tax on other GHGs is only captured through limited substitution and income effects as marginal abatement curves (MAC) have yet to be implemented.

<sup>83</sup> Given the `ifCEQ` flag, the two expressions can be combined into the following expression:

```
emiTax.fx(r,"co2",t) = cscale*x*(1$(not ifCEQ) + (44/12)$ifCEQ) ;
```

required to endogenize the coalition-wide emission tax and the region-specific emission tax.<sup>84</sup> The latter will be uniform across all regions within the same coalition.

Listing 3.1: Example of a cap regime with two coalitions

---

```

1  ifEmiCap      = 1 ;
2  emFlag("co2") = 1 ;
3  rq("lmy")     = yes ;
4  rq("hic")     = yes ;
5  emiCap.fx(rq, "co2", tsim) = 0.8*sum(r$mapr(rq,r), emiTot.l(r, "co2", tsim-1)*emiTot0(r, "co2") ;
6  emiRegTax.lo(rq, "co2", tsim) = -inf ;
7  emiRegTax.up(rq, "co2", tsim) = +inf ;
8  emiTax.lo(r, "co2", tsim) = -inf ;
9  emiTax.up(r, "co2", tsim) = +inf ;

```

---

Emissions trading is setup by specifying regional emission quotas. The sum of the quotas within a coalition must add up to the emissions cap for the coalition. Equation (E-8) determines the value of emissions trade for each region within a coalition. It is equal to the level of the coalition-wide emission tax times the difference between the quota allocation and actual emissions. If the difference is positive, the value of trade will be positive. It is assumed that the value of the trading rights goes directly to the government (with the counterbalancing account in the omitted balance of payments constraint), see equation (M-7). The code snippet below shows how to set up a cap and trade regime at the global level. It is assumed that developing countries are allocated quotas equivalent to their emission level from the previous period. The residual quotas, for developed countries, are allocated based on their share of emissions of the previous period. Note that in lines 10-12 where the quotas are set for the regions in the high-income coalition the quotas for the developing regions is subtracted from the overall cap using the lagged value of emissions and not the variable `emiQuota`. This is because the order of the loop over `ra` is undetermined and it is possible the quotas for the developed regions will be set before the quotas for the developing regions.

$$EmiQY_{r,em} = \tau_{r,em}^{Emi} (EmiQ_{r,em} - EmiTot_{r,em}) \quad (E-8)$$

Listing 3.2: Example of a cap and trade regime

---

```

1  ifEmiCap      = 1 ;
2  emFlag("co2") = 1 ;
3  rq("wld")     = yes ;
4  emiCap.fx("wld", "co2", tsim) = 0.8*sum(r$mapr("wld",r), emiTot.l(r, "co2", tsim-1)*emiTot0(r, "co2"
5  )) ;
6  ifEmiQuota(r) = 1 ;
7  loop(ra,
8    if(sameas(ra, "lmy"),
9      emiQuota.fx(r, "co2", tsim)$mapr(ra,r) = emiTot.l(r, "co2", tsim-1)*emiTot0(r, "co2") ;
10     elseif(sameas(ra, "hic"),
11       emiQuota.fx(r, "co2", tsim)$mapr(ra,r)
12       = (emiCap.l("wld", "co2", tsim)
13         - sum(rp$mapr("lmy",rp), emiTot.l(rp, "co2", tsim-1)*emiTot0(rp, "co2")))
14         * emiTot.l(r, "co2", tsim-1)*emiTot0(r, "co2")
15         / sum(rp$mapr(ra,rp), emiTot.l(rp, "co2", tsim-1)*emiTot0(rp, "co2")) ;
16     ) ;
17  emiRegTax.lo("wld", "co2", tsim) = -inf ;
18  emiRegTax.up("wld", "co2", tsim) = +inf ;
19  emiTax.lo(r, "co2", tsim) = -inf ;
20  emiTax.up(r, "co2", tsim) = +inf ;
21  emiQuotaY.lo(r, "co2", tsim) = -inf ;
22  emiQuotaY.up(r, "co2", tsim) = +inf ;

```

---

<sup>84</sup> The last two lines can potentially lead to errors for more complex coalitions as it endogenizes the emissions tax for all model regions. We are assuming here that all model regions are in one of the coalitions

## Chapter 4

# Model dynamics

The model as defined above represents the fully static version.<sup>1</sup> The factor stocks and technology are largely exogenous. This section describes the additional equations needed for the standard dynamics.

### 4.1 Labor supply

Labor is assumed to grow at some 'natural' rate for each skill level and by zone (in the case of segmented labor markets). At the moment, the natural rate of growth is equated to some measure of the growth in the relevant population and is assumed to be uniform for all skills and zones. The relevant population growth is defined by the following expression:

$$TLAB_{r,t} = \frac{\sum_p \phi_{r,p,t}^l Pop_{r,p,t}^t}{\sum_p \phi_{r,p,t-n}^l Pop_{r,p,t-n}^t} TLAB_{r,t-n}$$

where  $\phi^l$  is the relevant labor force participation rate. Population growth by cohort is assumed to be exogenous and is typically taken from a population scenario such as the UN or from those developed by the IAMC community known as the shared socio-economic pathways, or SSPs.

In the case of labor market segmentation with rural to urban migration, the level of migration is defined by equation (G-1). This simple migration equation is similar in spirit to the so-called Harris-Todaro model of migration.<sup>2</sup> Migration is a function of the expected urban wage relative to the expected rural wage, where  $\omega^m$  represents the migration elasticity.<sup>3</sup> Labor supply in each zone is equal to some exogenous component based on the 'natural' growth rate adjusted by migration, equation (G-2). The growth rate  $g^{lz}$  is typically exogenous and for the moment uniform across skills and zones. The parameter  $\delta^m$  is equal to -1 in the case of the rural zone and is equal to 1 for the urban zone. The parameter  $\mu^m$  is a multiplier for migration that accounts for migration occurring between periods in a multi-year time step. The level of the multiplier is given in equation (G-3). It is equal to 1 for single-year time steps. It is approximately equal to  $n$  for time steps of  $n$  years.

<sup>1</sup> In the model implementation, all of the equations listed till now are part of the **core** model—save for those that are substituted out.

<sup>2</sup> See [Harris and Todaro \(1970\)](#).

<sup>3</sup> There are more complex specifications of the migration function. See for example [Burniaux \(1987\)](#) and [Burniaux and van der Mensbrugghe \(1991\)](#).

Equation (G-4) calculates implicitly the growth rate of labor by skill at the economy-wide level.

$$MIGR_{r,l} = \chi_{r,l}^m \left( \pi_{r,l}^{Urb} \right)^{\omega_{r,l}^m} \quad (G-1)$$

$$L_{r,l,z,t}^s = \left( 1 + g_{r,l,z,t}^{lz} \right)^n L_{r,l,z,t-n}^s + \delta^m \mu_{r,l,z,t}^m MIGR_{r,l,t} \quad (G-2)$$

$$\mu_{r,l,z,t}^m = \frac{\left( 1 + g_{r,l,z,t}^{lz} \right)^n \left( \frac{\pi_{r,l,t-n}^{Urb}}{\pi_{r,l,t}^{Urb}} \right)^{\omega_{r,l}^m} - 1}{\left( 1 + g_{r,l,z,t}^{lz} \right)^n \left( \frac{\pi_{r,l,t-n}^{Urb}}{\pi_{r,l,t}^{Urb}} \right)^{\omega_{r,l}^m/n} - 1} \quad (G-3)$$

$$LAB_{r,l,t}^s = \left( 1 + g_{r,l,t}^l \right)^n LAB_{r,l,t-n}^s \quad (G-4)$$

## 4.2 Capital supply

The total capital stock,  $K^s$  evolves according to the standard capital accumulation formula, equation (G-5). Since both terms on the right-hand side are lagged, with a single year time step the supply of capital is in effect exogenous. Annex C.1 explains how the equation is implemented for multi-year time steps. In this case, the equation depends on the contemporaneous level of investment and the capital stock is therefore no longer pre-determined.

$$K_{r,t}^s = (1 - \delta_t) K_{r,t-1}^s + XFD_{r,Inv,t-1} \quad (G-5)$$

Equation (G-6) defines the normalized capital stock. It is a fixed scalar of the non-normalized level where the scaling factor is initialized using base year data, i.e.  $\chi_r^k = TK_{r,0}^s / K_{r,0}^s$ . The normalized capital stock is equal to aggregate capital remuneration in the base year, i.e. it is the level of the capital stock with the price of aggregate capital set to 1.

$$TK_{r,t}^s = \chi_r^k K_{r,t}^s \quad (G-6)$$

## 4.3 Depletion module

A depletion module has been developed for the coal, oil and gas reserves. Extraction of the relevant resource comes from proven reserves at a potentially sensitive extraction rate. The model assumes that for coal the extraction rate is price sensitive and that the elasticity of extraction depends on market conditions. The price sensitivity is lower on the up-side than on the down-side, i.e. it is easier to contract than to expand. The oil and gas depletion specification is somewhat more complicated. Two reserve notions are implemented for these two fuels—proven and unproven. Extraction from the proven reserve occurs at a fixed extraction rate. Unproven, or yet-to-discover reserves, are converted to proven reserves using a price-sensitive conversion rate.

## 4.4 Knowledge module

[NEW 20-JAN-2019] The model includes a knowledge module that operates somewhat similarly to investment and the capital stock. Each year there is aggregate expenditure on research and development (R&D), that is captured by the variable  $XFD_{r,d}$ . This expenditure flow increases

knowledge, though knowledge also depreciates each year. The main difference with the usual capital/investment dynamic, is that the impact of R&D expenditures happens with a distributed lag.<sup>4</sup> The latter is governed by a *Gamma* distribution function, the parameters and length of which are region specific.<sup>5</sup>

Equation (G-7) describes the motion equation for knowledge. The variable  $KN$  represents the stock of knowledge. It is equal to the previous period's knowledge stock—adjusted by depreciation,  $\delta^k$ —, plus the distributed lag of current and previous expenditures on R&D, where  $\gamma^k$  represents the lag coefficients.<sup>6</sup> Annex C.2 describes how the model implements the knowledge module. Implementation requires taking into account two features. (1) It is implemented using the variable time step that is a central part of the model. Because of the distributed lags, this also requires that the R&D expenditures be evaluated each *year*, independently of the time step. (2) The first years of the simulation require a backward projection of the R&D expenditures, i.e. prior to the initial year. If  $N$  is 50 for example, it requires projecting R&D expenditures back some 50 years.

$$KN_{r,t} = (1 - \delta_r^k) KN_{r,t-1} + \sum_{i=0}^N \gamma_{r,i}^k XFD_{r,r-d,t-i} \quad (\text{G-7})$$

where

$$\gamma_i^k = \chi (i + 1)^{\delta/(1-\delta)} \lambda^i$$

The knowledge module is designed to impact on labor productivity growth. Equation (G-8) describes the relation between the endogenous component of labor productivity in sector  $a$ ,  $\pi^k$ , relative to the growth rate of the stock of knowledge. The elasticity is given by  $\epsilon^r$  and is allowed to be sector specific. The shifter,  $\gamma^r$ , is initialized at 1.

$$\pi_{r,a,t}^k = \gamma_{r,a,t}^r \epsilon_{r,a,t}^r \left( \frac{KN_{r,t}}{KN_{r,t-1}} - 1 \right) \quad (\text{G-8})$$

## 4.5 Technology shifts

By default, changes in technology are largely exogenous, for example autonomous energy efficiency improvement (AEEI) and improvements in international trade and transport services. The one exception is labor productivity. Equation (G-9) describes the growth in labor productivity. In the base year,  $\lambda^f$  is initialized to 1, and is set to grow at a rate of  $\gamma^l$  each year. If  $\alpha^{gl}$  is set to 0 and  $\beta^{gl}$  is set to 1, then labor productivity grows at a uniform rate across all activities. There are multiple strategies to put values on these parameters. These may include setting  $\beta^{gl}$  to 0 for a subset of activities and initializing  $\alpha^{gl}$  to some exogenous level based on external studies—for example setting  $\alpha^{gl}$  to 2 percent in the agricultural sectors. The remaining sectors could be segmented into two subsets—one with high productivity, e.g. set  $\alpha^{gl}$  to 2 percent, and one with low productivity, e.g. set  $\alpha^{gl}$  to 0, with  $\beta^{gl}$  set to 1 in both cases.

Two other factors can influence sector-specific labor productivity. The first is a skill-specific shifter,  $\chi^{gl}$ , that introduces biased technical change across skill types. In typical simulations, this factor is initialized at 0 and there is no difference in labor productivity growth across skills. One alternative strategy is to have some (exogenous) assumption on the skill premium, for example no change in the skill premium. It is then possible to endogenize skill-biased technical change to see

<sup>4</sup> This formulation is inspired by Smeets Křístková et al. (2016).

<sup>5</sup> Additional details and example distribution functions are provided in Annex C.2.

<sup>6</sup> If  $\gamma_0^k = 0$  and  $\gamma_1^k = 1$ , the equation would have the same assumption as the standard capital accumulation equation.

how much this parameter must change to maintain the exogenous assumption on the skill premium. The second factor is the knowledge-induced endogenous technical change,  $\pi^k$ . This formulation is described in the knowledge module, see for example equation (G-8).

In the baseline or reference scenario,  $\gamma^l$  will be calibrated to target a given level of per capita GDP growth and subject to the assumptions made on  $\alpha^{gl}$  and  $\beta^{gl}$ . In subsequent scenarios,  $\lambda^f$  will be exogenous having been calibrated in the reference scenario, subject to any changes in the two other factors:  $\chi^{gl}$  and  $\pi^k$ .

$$\lambda_{r,l,a,t}^f = \left(1 + \chi_{r,l,t}^{gl} + \pi_{r,a,t}^k + \alpha_{r,l,a,t}^{gl} + \beta_{r,l,a,t}^{gl} \gamma_{r,t}^l\right)^n \lambda_{r,l,a,t-n}^f \quad (\text{G-9})$$

## 4.6 Preference shifters

### 4.6.1 Armington twists

In the standard version of the model, the Armington preference parameters—both top and second levels—are held fixed at their base year calibrated levels. Thus changes in import shares are linked only to changes in relative prices. Historically, we have observed large changes in import shares without a concomitant change in relative prices. [Dixon and Rimmer \(2002\)](#) have developed a specification for the MONASH (and USAGE) model that ‘twists’ the relevant share parameters. The ‘twist’ targets a percentage change in the ratio of imports to domestic demand, which at unchanged component prices, is cost-neutral. The mechanics of the twist methodology is described in [Appendix A](#).

For the ENVISAGE model, we have incorporated three versions of the twist—for the dual nested sourcing decision. The first two deal with the top-level Armington decision—purchases of domestic goods and the aggregate import bundle. Since we have two possible specifications—agent-based sourcing or national sourcing, we provide two twist specifications. The third is for the second level nest that sources aggregate imports by region of origin.

### National sourcing

The equations for national sourcing of domestic purchases and aggregate imports are provided in equations (T-2) and (T-3), with the domestic and import share parameters given respectively by  $\alpha^{dt}$  and  $\alpha^{mt}$ . Equation (G-10) provides the ‘twisted’ domestic and imported share parameters where  $tw^{tl}$  is the level of the desired twist, i.e. the percent change in the ratio of imports to domestic purchases at unchanged prices. The twist has no impact on the share parameters if it is set to its default value of 0. The twist updating uses the (lagged) value share of imports in the top-level Armington bundle.<sup>7</sup> Note that the formulas apply the twist factor directly to the share coefficients themselves and we drop the  $\lambda$  factor, i.e. preference change variable, from the formulas as these can be substituted out (refer to [Appendix A](#)).

$$\begin{aligned} \alpha_{r,i,t}^{dt} &= \alpha_{r,i,t-1}^{dt} \frac{1}{1 + s_{r,i,t-1}^{tl} tw_{r,i,t}^{tl}} \\ \alpha_{r,i,t}^{mt} &= \alpha_{r,i,t-1}^{mt} \frac{1 + tw_{r,i,t}^{tl}}{1 + s_{r,i,t-1}^{tl} tw_{r,i,t}^{tl}} \end{aligned} \quad (\text{G-10})$$

where

<sup>7</sup> In the MONASH and USAGE models, the twist parameters are calculated contemporaneously, i.e. they are part of the model equations. In ENVISAGE, the twists are applied between periods. This should have only minimal impact on model results.



$$s_{r,i,t-1}^{t1} = \frac{PMT_{r,i,t-1} XMT_{r,i,t-1}}{PAT_{r,i,t-1} XAT_{r,i,t-1}}$$

### Agent sourcing

The expressions for agent-based sourcing of domestic goods and aggregate imports are nearly identical with the inclusion of the index for the Armington agent,  $aa$ . The share parameters are part of equations (T-6) and (T-7). Equation (G-11) provides the relevant formulas for the 'twisted' shares using the appropriate import share level. Note that the twists are allowed to be agent-specific, though in practice, there is probably little evidence to differentiate across agents.

$$\begin{aligned} \alpha_{r,i,aa,t}^d &= \alpha_{r,i,aa,t-1}^d \frac{1}{1 + s_{r,i,aa,t-1}^I tw_{r,i,aa,t}^I} \\ \alpha_{r,i,aa,t}^m &= \alpha_{r,i,aa,t-1}^m \frac{1 + tw_{r,i,aa,t}^I}{1 + s_{r,i,aa,t-1}^I tw_{r,i,aa,t}^I} \end{aligned} \quad (G-11)$$

where

$$s_{r,i,aa,t-1}^I = \frac{PM_{r,i,aa,t-1} XM_{r,i,aa,t-1}}{PA_{r,i,aa,t-1} XA_{r,i,aa,t-1}}$$

### Second level Armington twists

The second level of the Armington nest allocates aggregate imports across source regions (see equation (T-13)). With the sole exception of a 2-region model, the standard twist does not apply and we thus apply the twist to a target set of source countries, i.e. we segment the source regions into 2.<sup>8</sup> The targeted regions will see their share increase (by the same twist), and the non-targeted regions will see their share decline.

$$\begin{aligned} \alpha_{s,i,r,t}^w &= \alpha_{s,i,r,t-1}^w \frac{1}{1 + s_{i,r,t-1}^2 tw_{r,i,t}^2} \quad \text{if } s \notin rtwtgt(s,r) \\ \alpha_{s,i,r,t}^w &= \alpha_{s,i,r,t-1}^w \frac{1 + tw_{r,i,t}^2}{1 + s_{i,r,t-1}^2 tw_{r,i,t}^2} \quad \text{if } s \in rtwtgt(s,r) \end{aligned} \quad (G-12)$$

where

$$s_{i,r,t-1}^2 = \frac{\sum_{s \in rtwtgt(s,r)} PDM_{s,i,r,t-1} XW_{s,i,r,t-1}}{\sum_s PDM_{s,i,r,t-1} XW_{s,i,r,t-1}}$$

### 4.6.2 Introducing cost curves

Changes in costs in the normal functioning of the model depend on changes in input prices and the standard assumptions regarding technology change. This section explains how to introduce an acceleration in cost reduction coming from some exogenous phenomenon such as learning by doing. The basic idea is to start from an initial price, say  $P_0$ , and to reduce costs over time, albeit with a lower limit given by  $PMIN$ . There also exists a target price for year  $T$ , given by  $P_T$ , that must be greater than  $PMIN$ . Let  $\alpha$  represent the ratio of the final price,  $P_T$ , relative to the base year price,  $P_0$ , i.e.:

$$P_T = \alpha P_0$$

<sup>8</sup> The targeted regions can include only a single region.

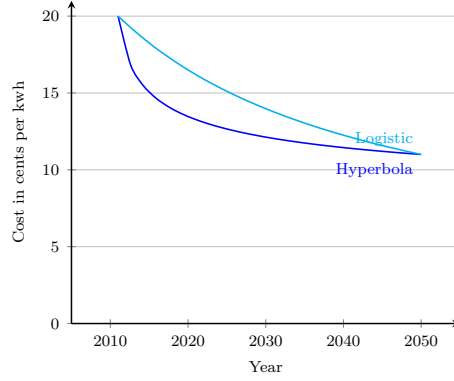


Figure 15: Examples of cost curves

and let  $\gamma$  represent the ratio of the minimum price,  $PMIN$ , relative to the initial price, i.e.:

$$PMIN = \gamma P_0$$

where  $\alpha > \gamma$ . Two functional forms are introduced to represent the cost curve—the hyperbola and the logistic functions. The hyperbola takes the following form:

$$P_t = PMIN + \chi t^{-\beta}$$

The logistic function takes the form:

$$P_t = \frac{PMIN}{1 + \chi e^{-\beta t}}$$

Calibration of the parameters for the hyperbola function involve the following expressions:

$$\beta = \frac{\ln\left(\frac{\alpha - \gamma}{1 - \gamma}\right)}{\ln\left(\frac{t_0}{T}\right)} \quad \chi = P_0 (1 - \gamma)$$

where we define  $t_0$  to start with 1.<sup>9</sup> In the case of the logistic function, the calibration formulas are:

$$\beta = -\frac{1}{T} \ln\left(\frac{\alpha - \gamma}{\alpha(1 - \gamma)}\right) \quad \chi = \gamma - 1$$

where we define  $t_0$  to start with 0.<sup>10</sup> Figure 15 depicts the shapes of the two cost curves for the same initial and end points and price limit. From the shape of the curves it is clear that costs drop sharply in the initial years using the hyperbola specification. The drop is more gradual (and constant in percentage terms) for the logistic function.

The adjustment to the cost curve is made on total factor productivity (TFP) for the relevant activity. The aggregate cost is given by the following formula:

<sup>9</sup> If the first year is 2011,  $t$  is then defined as  $t - 2010$ . If the final year is 2050, then  $T$  is equal to  $2050 - (2011 - 1) = 40$ .

<sup>10</sup> If the first year is 2011,  $t$  is then defined as  $t - 2011$ . If the final year is 2050, then  $T$  is equal to  $2050 - 2011 = 39$ .

$$P_t = \frac{1}{\lambda_t} F(P_{i,t})$$

where  $F$  is the cost function and depends on the prices of the various inputs given by  $P_i$ . In most cases, the cost function will be the dual price expression of the CES function.<sup>11</sup> TFP is assumed to behave according to the following expression:

$$\lambda_t = (1 + \pi_t)^n \lambda_{t-n}$$

Assuming that input prices are invariant, the growth in TFP must be equal to the following expression:

$$\pi_t = \left( \frac{P_{t-n}}{P_t} \right)^{1/n} - 1$$

We can introduce the formulas for the two cost curves to derive expressions for  $\pi$ . For the hyperbola, we have:

$$\pi_t = \left( \frac{PMIN + \chi(t-n)^{-\beta}}{PMIN + \chi t^{-\beta}} \right)^{1/n} - 1 \iff \left( \frac{\gamma + (1-\gamma)(t-n)^{-\beta}}{\gamma + (1-\gamma)t^{-\beta}} \right)^{1/n} - 1$$

For the logistic function, the expression for  $\pi$  is given by:

$$\pi_t = \left( \frac{e^{\beta t} + \chi}{e^{\beta t} + \chi e^{\beta n}} \right)^{1/n} - 1$$

Both changes in preferences and introduction of the cost curves can be implemented to change the share of any technology in a demand bundle. The latter operates directly on the preference shares of the buyer. The former operates on the perceived cost of the technology. These are formulated independently of the actual cost changes that arise from the endogenous changes in input prices.

## 4.7 Implementation

The **core** model includes all of the equations from Chapter 3.<sup>12</sup> There are two versions of the dynamic model. The reference simulation, that targets GDP growth, is defined as the **core** model plus the GDP growth equation (M-5), migration (G-1), labor supply by skill and zone (G-2), total labor supply by skill (G-4), investment growth factor (C-1), non-normalized capital stock (C-2), normalized capital stock (G-6), and labor productivity (G-9). This model is called **coreBaU** and it essentially calibrates the economy-wide labor productivity factor. All other dynamic simulations assume that labor productivity is exogenous i.e. per capita GDP growth is endogenous. This model is called **coreDyn** and includes the **core** model plus the GDP growth equation (M-5), migration (G-1), labor supply by skill and zone (G-2), total labor supply by skill (G-4), investment growth factor (C-1), non-normalized capital stock (C-2), and normalized capital stock (G-6).

<sup>11</sup> In most cases  $\lambda$  is initialized at 1 and is held constant. Technology is introduced as labor-augmenting technical change and with an exogenous improvement in energy efficiency.

<sup>12</sup> With the exception of the GDP per capita growth equation as this requires a different MCP formulation.

## 4.8 Diagnostics

There are a number of post-simulation diagnostics available with the model. Most are relatively straightforward and need no additional explanation, for example the GDP deflator at market price—these are taken directly from the model solution file and aggregated where appropriate.

### 4.8.1 Sectoral investment

Sectoral investment can be backed out by the capital accumulation motion equation:

$$I_{a,t} = K_{a,t+1} - (1 - \delta)K_{a,t}$$

## 4.9 Exogenous variables

In this section, we summarize the standard list of exogenous variables in the model. These are partitioned into sub-groups depending on their role in the model specification.

### 4.9.1 Factors of production

Population and labor force growth are typically exogenous. The latter may be defined over rural and urban if using the migration module. In this case, the growth of labor is exogenous, but the growth per zone may be influenced by endogenous rural to urban migration. The capital stock is exogenous in the comparative static version of the model and is in principle pre-determined in the dynamic model with annual step sizes. In the case of multi-year step size, current investment will influence the current stock level (see equation (C-2)). Natural resources are supplied with an upward sloping supply curve with two, normally fixed, shifters. The first shifter is a regional shifter. The second is a global shifter. The latter has been used to target a trend for world prices (e.g. oil or gas). In this case the world price for the relevant sector is fixed to the desired trend level and the global resource supply shifter is endogenous. This would typically be implemented only in a baseline. The land and water asymptotes are used-entered, as is the demand for exogenous water bundles.

Table 4.1: **Exogenous factor variables**

Variable	GAMS	Description
$Pop^t$	pop	Population
$g^{lz}$	glabz	Growth of labor by zone
$\chi^{rw}$	chirw	Reservation wage shifter
$K^s$	kstock	Capital stock (endogenous in dynamic version)
$TK^s$	tkaps	Normalized capital stock (endogenous in dynamic version)
$\chi^{w,nrs}$	wchinrs	Global natural resource supply shifter
$\chi^{nrs}$	chinrs	Region specific natural resource supply shifter
$PW$	pw	World price by activity
$TLand^{Max}$	landmax	Land maximum
$TH2O^{Max}$	h2oMax	Water maximum
$H2OBnd_{wbx}$	h2obnd	Exogenous water bundle demand

### 4.9.2 Macro closure

There are three different closure rules—all with default assumptions. The savings-investment closure fixes the household savings rate,  $aps$ , adjusted by the shift parameter  $\chi^s$ . The former is always fixed. The latter can be endogenous to achieve a given target—for example the investment to GDP ratio. The latter is endogenous under the default closure rule but may be over-ridden. Typically, in the baseline, one will target some level of the investment to GDP ratio and allow the private savings rate adjust.<sup>13</sup> In a similar vein, the wage premium, normally the skilled wage relative to unskilled wages, is endogenous. One could target the wage premium, i.e. make it exogenous. The endogenous variable could be the relative growth of skilled and unskilled labor in levels or the relative growth in labor productivity. The default public closure is that government expenditures are fixed relative to GDP in real terms (in the baseline) and real government savings are fixed. The direct tax rate,  $\kappa^h$ , on household income, is the swing variable that ensures the fiscal target. The default balance of payment closure fixes net foreign saving. The real exchange rate is endogenous. Two additional closures are available. One fixes the share of foreign savings relative to GDP—similar in many ways to the default closure. The third allows for some endogeneity in foreign savings, in which case the deviation of expected capital returns is equalized across all regions.

Table 4.2: **Exogenous macro closure variables**

Variable	GAMS	Description
$aps$	<b>aps</b>	Average savings propensity (not for ELES)
$\chi^s$	<b>chiaps</b>	Private savings shifter (depends on investment/savings closure)
$\pi^s$	<b>wprem</b>	Wage premium
$RS^g$	<b>rsg</b>	Real public savings (depends on fiscal closure)
$YFD_{gov}$	<b>xfd(gov)</b>	Government expenditures
$S^f$	<b>savf</b>	Foreign saving (depends on BoP closure)
$\chi^{sf}$	<b>savRat</b>	Foreign saving as share of GDP (depends on BoP closure)
$PNUM$	<b>pnum</b>	Model numeraire

### 4.9.3 Tax rates

In the default specification all tax rates are exogenous. The actual tax rate used could depend on model specification choices, for example, a national versus agent-based representation of the second level Armington specification.

<sup>13</sup> The targeted investment variable is **rinvshr** in the GAMS code.

Table 4.3: **Exogenous tax rate**

Variable	GAMS	Description
$\tau^p$	ptax	Production tax
$\tau^{uc}$	uctax	Unit cost tax
$\tau^v$	pftax	Tax on factor use
$\kappa^f$	kappaf	Factor income tax by source activity and factor
$\tau^a$	patax	Tax on Armington consumption (depends on ArmFlag)
$\tau^{ad}$	pdtax	Tax on domestic sales of domestic goods (depends on ArmFlag)
$\tau^{am}$	pmtax	Tax on domestic sales of imported goods (depends on ArmFlag)
$\tau^e$	etax	Bilateral export tax
$\tau_{s,i,r}^m$	mtax	Bilateral import tariff
$\tau^{ma}$	mtaxa	Import tariff for MRIO version of model

#### 4.9.4 Technology variables

In most cases all of the technology variables are exogenous, though not necessarily fixed at base year levels—for example, the international trade margins or energy efficiency. The model calibrates labor-biased technological change (the  $\lambda^f$  parameter for labor) in the baseline to target GDP growth. In this case,  $g^y$ , i.e. per capita output growth, is exogenous, and  $\gamma^l$  is endogenous. An equation in the model evaluates  $\lambda^f$  for labor based on the exogenous GDP growth target. In all subsequent simulations the labor productivity variable is fixed and GDP growth is endogenous.

Table 4.4: **Exogenous technology variables**

Variable	GAMS	Description
$\zeta^{mg}$	tmarg	Bilateral trade margins
$A^{xpv}$	axghg	Uniform shifter in production bundle
$\lambda^{xp}$	lambdaxp	Output shifter in production bundle
$\lambda^{ghg}$	lambdaghg	GHG shifter in production bundle
$\lambda^f$	lambdaf	Productivity parameter for primary factors
$\lambda^k$	lambdak	Capital efficiency shifter
$\lambda^{h2ob}$	lambdah2obnd	Water efficiency shifter in water bundle use
$\lambda^{io}$	lambdaio	Efficiency shifter in intermediate demand
$\lambda^e$	lambdae	lambdae
$\lambda^{ce}$	lambdace	lambdace
$g^y$	grrgdppc	Growth rate of real per capita GDP
$\gamma^l$	gl	Economywide labor productivity parameter

#### 4.9.5 Preference variables

Preference variables are all fixed, but may depend on which utility function is implemented. They could also be adjusted between solution periods to target some trend growth, for example the food budget, as incomes change.

Table 4.5: **Exogenous preferences variables**

Variable	GAMS	Description
$e^h$	eh	CDE expansion parameter
$b^h$	bh	CDE substitution parameter
$\alpha^h$	alphah	CDE shift parameter
$\gamma$	gamma	AIDADS/E/LES marginal propensity to consume parameter
$\mu^c$	muc	AIDADS/E/LES Subsistence minima

#### 4.9.6 Emission variables

The default specification for emissions is that they are endogenous and the emissions tax is exogenous (and in most cases set to zero<sup>14</sup>). One could target a trend in global emissions for any particular gas or pollutant and allow a global shifter—attached to the emission equations—be endogenous ( $\chi^{Emi}$ ). The model also allows for non-model based emissions, either at a regional (e.g. deforestation) or global level. The emission regime variables are described in the emissions section.

Table 4.6: **Emission variables**

Variable	GAMS	Description
$\chi^{Emi}$	chiEmi	Global shifter in emissions
$EmiOth$	emi0th	Other non-modeled emissions (by region)
$EmiOthGbl$	emi0thGbl	Other non-modeled global emissions
$\tau^{Emi}$	emitax	Emission tax
$Emi^{Cap}$	emiCap	Emission cap
$EmiQ$	emiQuota	Emission quota
$EmiQY$	emiQuotaY	Emission quota revenues
$\tau^{EmiQ}$	emiRegTax	Regional emission tax
$\chi^{Cap}$	chiCap	Emissions cap shifter

<sup>14</sup> The exceptions could include the price on carbon emissions in the European Union’s Emission Trading Scheme, or an interpretation of the Nationally Determined Commitments (NDCs) that are part of the Paris Agreement

# Appendix A

## The CES and CET functions

This appendix describes in full detail the two functional forms most widely used in CGE models—the constant-elasticity-of-substitution (CES) and constant-elasticity-of-transformation (CET) functions.

CES functions are widely used in demand functions where substitutability across different products and/or factors is needed and where the main objective is to minimize cost. CET functions are broadly used to determine supply functions across different markets where the main objective is to maximize revenues. The two are very similar in many ways and the algebraic derivations below will be more detailed for the CES function.

### A.1 The CES function

#### A.1.1 Basic formulas

In production, the CES function is used to select an optimal combination of inputs (either goods and/or factors) subject to a CES production function. In consumer demand, the CES is used as a utility (or sub-utility) or preference function. In either case, the purpose is to minimize the cost of purchasing the 'inputs' subject to the production or utility function. In generic terms the system takes the following form:

$$\min_{X_i} \sum_i P_i X_i$$

subject to the constraint:

$$V = A \left[ \sum_i a_i (\lambda_i X_i)^\rho \right]^{1/\rho}$$

The objective function represents aggregate expenditure. The constraint expression will be referred to as the CES primal function. The parameter  $A$  is an aggregate shifter that can be used to shift the overall production function (or utility function). Each input,  $X_i$ , is multiplied by an input-specific shifter,  $\lambda_i$ , that can be used to implement input-specific productivity increases (for example biased technological change), or specific changes in consumer preferences. The (primal) share coefficients,  $a_i$ , are typically calibrated to some base year data and held fixed. The CES exponent,  $\rho$ , is linked to the curvature of the CES function (and will be explained further below). For given component prices,  $P_i$ , and a given level of production or utility  $V$ , solving the optimization program above will yield optimal demand functions for the inputs,  $X_i$ .



The Lagrangian can be set up as:

$$\mathcal{L} = \sum_i P_i X_i + \Lambda \left( V - A \left[ \sum_i a_i (\lambda_i X_i)^\rho \right]^{1/\rho} \right)$$

Taking the partial derivative with respect to  $X_i$  and the Lagrange multiplier  $\Lambda$  yields the following system of equations:

$$P_i = \Lambda a_i \lambda_i^\rho X_i^{\rho-1} A \left[ \sum_i a_i (\lambda_i X_i)^\rho \right]^{(1-\rho)/\rho} = \Lambda a_i A^\rho \lambda_i^\rho X_i^{\rho-1} V^{1-\rho}$$

$$V = A \left[ \sum_i a_i (\lambda_i X_i)^\rho \right]^{1/\rho}$$

Taking the first expression, it can be multiplied by  $X_i$ , and then summed. This of course is equal to the value of the bundle, i.e.  $P.V$ , where  $P$  is the aggregate price:

$$P.V = \sum_i P_i X_i = \Lambda V^{1-\rho} A^\rho \sum_i a_i \lambda_i^\rho X_i^\rho = \Lambda V^{1-\rho} V^\rho = \Lambda V$$

This shows that  $\Lambda$ , the Lagrange multiplier is the same as the aggregate price,  $P$ . We can re-arrange expression above to get an expression for optimal input demand, where  $\Lambda$  is replaced by  $P$ :

$$X_i = a_i^{1/(1-\rho)} A^{\rho/(1-\rho)} \left( \frac{P}{P_i} \right)^{1/(1-\rho)} \lambda_i^{\rho/(1-\rho)} V$$

We finally end up with the following expression, where the CES primal exponent,  $\rho$ , is replaced by the so-called CES elasticity of substitution,  $\sigma$ :

$$X_i = \alpha_i (A \lambda_i)^{\sigma-1} \left( \frac{P}{P_i} \right)^\sigma V \quad (\text{A-1})$$

where we made the following substitutions:

$$\sigma = \frac{1}{1-\rho} \Leftrightarrow \rho = \frac{\sigma-1}{\sigma} \Leftrightarrow \frac{\rho}{1-\rho} = \sigma-1 \Leftrightarrow \rho \cdot \sigma = \sigma-1$$

and

$$\alpha_i = a_i^{1/(1-\rho)} = a_i^\sigma \Leftrightarrow a_i = \alpha_i^{1/\sigma}$$

Abstracting from the technology parameters, the demand equation implies that demand for 'input'  $X_i$  is a (volume) share of total demand  $V$ . The share, with equal prices is simply equal to  $\alpha_i$ . With a positive elasticity of substitution, the share is sensitive to the ratio of prices relative to the aggregate price index. Since the component price is in the denominator, the demand for that component declines if its price rises relative to the average and vice versa if its price declines vis-à-vis the average price. The  $\alpha$  parameters will be referred to as the CES dual share parameters (for reasons described below), and the  $a$  parameters are the primal CES share parameters. Notice that expression (A-1) simplifies if it is expressed in terms of efficiency inputs,  $X^e$  and efficiency prices,  $P^e$ :

$$X_i^e = \alpha_i A^{\sigma-1} \left( \frac{P}{P_i^e} \right)^\sigma V$$

where

$$X_i^e = \lambda_i X_i$$

and

$$P_i^e = \frac{P_i}{\lambda_i}$$

The aggregate price  $P$  can be determined using two expressions. The first is the zero profit condition:

$$P = \frac{\sum_i P_i X_i}{V}$$

The other is by inserting the optimal demand relation  $X_i$  (equation A-1) in the zero profit condition :

$$P.V = \sum_i P_i X_i = A^{\sigma-1} \sum_i P_i \alpha_i \left( \frac{P}{P_i} \right)^\sigma \lambda_i^{\sigma-1} V = P^\sigma A^{\sigma-1} V \sum_i \alpha_i \left( \frac{P_i}{\lambda_i} \right)^{1-\sigma}$$

The  $V$ 's cancel out, and the aggregate price can then be expressed by the following formula:

$$P = \frac{1}{A} \left[ \sum_i \alpha_i \left( \frac{P_i}{\lambda_i} \right)^{1-\sigma} \right]^{1/(1-\sigma)} = \frac{1}{A} \left[ \sum_i \alpha_i (P_i^e)^{1-\sigma} \right]^{1/(1-\sigma)} \quad (\text{A-2})$$

This is sometimes referred to as the dual price expression. It has virtually the same functional form as the CES primal, which is a CES aggregation of the input volumes using the primal share parameters as weights. The CES dual price formula is a CES aggregation of the input prices using the CES dual share parameters as weights and a different exponent. In a CGE model, the zero-profit condition or the dual price formula can be used interchangeably (with the proviso that the substitution elasticity differs from 1).<sup>1</sup> There is a simple formula for the budget shares given by:

$$s_i = \frac{P_i X_i}{P.V} = \alpha_i (A \lambda_i)^{\sigma-1} \left( \frac{P}{P_i} \right)^\sigma V \left( \frac{P_i}{P} \right) \frac{1}{V} = \alpha_i (A \lambda_i)^{\sigma-1} \left( \frac{P}{P_i} \right)^{\sigma-1} \quad (\text{A-3})$$

Notice that this expression for the budget shares is only a function of prices. With the technology parameters set to 1, this simplifies further to:

$$s_i = \alpha_i \left( \frac{P}{P_i} \right)^{\sigma-1}$$

It turns out that the parameter  $\sigma$  measures the elasticity of substitution for the CES function and is constant over the entire domain. The elasticity of substitution is an indication of the curvature of an isoquant, see Varian (1992), i.e. it measures the rate of change of the ratio of inputs (in a 2-input case), relative to the change in their relative prices. For example, if the CES

<sup>1</sup> We shall see below that when the substitution elasticity is 1, both primal and dual expressions take a different functional form.

combines capital and labor to form output, a large substitution elasticity suggests that the factor proportions will change rapidly as one of the inputs becomes cheaper relative to the other. There are two limiting cases of interest. If the substitution elasticity is zero, then there is no substitution across inputs and the optimal choice is to use them in fixed proportion. At the other extreme, if the substitution elasticity is infinite, this is equivalent to saying the inputs are identical, and in this case, in equilibrium, the two inputs would have the same price. This could potentially be the case for electricity production. If there is a regional or national buyer of electricity, the buyer is most likely indifferent about how the electricity is produced and thus will purchase from the lowest cost producer (a perhaps somewhat simplified view of electricity markets.) This implies that the cost of the electricity inputs, from all sources (e.g. thermal, nuclear, etc.) would be (nearly) identical.

The elasticity of substitution across inputs is defined by the following formula:

$$\sigma = \frac{\partial \left( \frac{X_i}{X_j} \right) \left( \frac{P_i}{P_j} \right)}{\partial \left( \frac{P_i}{P_j} \right) \left( \frac{X_i}{X_j} \right)}$$

The ratio of the optimal inputs using expression (A-1) is:

$$\frac{\alpha_i}{\alpha_j} \left( \frac{P_i}{P_j} \right)^{-\sigma} \left( \frac{\lambda_i}{\lambda_j} \right)^{\sigma-1}$$

Taking the partial derivative of the expression with respect to the ratio  $P_i/P_j$  and multiplying it by the second term of the elasticity of substitution yields the conclusion that the substitution elasticity is  $-\sigma$ . It is logical that it is negative. If the price of one input increases, say labor, relative to the other, say capital, producers would substitute away from labor towards capital, i.e. the ratio of labor to capital would drop as the price of labor increases relative to capital. Varian (1992) in fact defines the elasticity of substitution in terms of the absolute value of the technical rate of substitution, that measures the slope of the budget line. Numerically what it represents is the relative change in the ratios. If  $\sigma$  is 1, for example, and the price of labor increases by 10 percent relative to capital, the labor to capital ratio would decrease by (around) 10 percent.<sup>2</sup> The higher is  $\sigma$ , the more the proportion changes.

### A.1.2 Special cases

There are three special cases that require additional derivations due to numerical restrictions on the primal and dual exponents. A substitution elasticity of 0 is clearly a special case and is referred to as a Leontief technology. From the dual price formula, it is clear that  $\sigma$  equal to 1 is a special case and is known as a Cobb-Douglas technology (or utility function). Finally, a value of  $\rho$  equal to 1 corresponds to infinite substitution elasticity and a linear primal aggregation function. This is also referred to as a case of perfect substitution.

#### The Leontief case

The first special case is for the so-called Leontief functional form.<sup>3</sup> In this case the substitution elasticity is 0 and corresponds to a value for  $\rho$  that is  $-\infty$ . In this case the optimization program takes the following form:<sup>4</sup>

<sup>2</sup> The elasticity is a marginal concept that holds only approximately for large changes.

<sup>3</sup> Leontief, winner of the 1973 Nobel prize in Economics, is renowned for his work on input-output tables, much of which focused on fixed input technologies (!!!! reference).

<sup>4</sup> !!!! need a reference

$$\min_{X_i} \sum_i P_i X_i$$

subject to the constraint:

$$V = \min \left( \frac{a_i}{\lambda_i X_i} \right)$$

The visual implementation has L-shaped isoquants. The Leontief technology constraint or production/utility function is discontinuous. Fortunately, the optimal demand functions are easy to implement and are just special cases of expression (A-1):

$$X_i = \frac{\alpha_i}{\lambda_i} \frac{V}{A}$$

$$P = \frac{1}{A} \sum_i \alpha_i \left( \frac{P_i}{\lambda_i} \right)$$

Thus the Leontief specification implies that inputs are always in fixed proportion relative to output and the aggregate price is simply the linear weighted aggregation of the input prices, where the weights are given by the input-output coefficients, adjusted by changes in efficiency. The efficiency parameter has a nice intuitive interpretation in this case. Say  $\lambda$  increases by 10 percent, then demand for the input declines by 10 percent.

### The Cobb-Douglas function

Another special case is the so-called Cobb-Douglas function, very frequently used in introductory text books in microeconomics. The Cobb-Douglas function has a substitution elasticity of 1 implying that  $\rho$  is equal to 0. Clearly, this creates a problem for specifying the CES primal function as well as the CES dual price function. As with the Leontief, the optimal demand conditions are given by expression (A-1), with  $\sigma$  set to 1:

$$X_i = \alpha_i \left( \frac{P}{P_i} \right) V \Leftrightarrow s_i = \frac{P_i X_i}{P \cdot V} = \alpha_i$$

The Cobb-Douglas specification has constant budget shares irrespective of relative prices (and changes in technology). Another implication of the Cobb-Douglas specification is that the dual shares must add up to 1 as they are equivalent to the budget shares. By definition, as well, the primal and dual shares are the same. The Cobb-Douglas primal and dual price functions have the following expressions:

$$V = A \prod_i (\lambda_i X_i)^{\alpha_i}$$

$$P = \frac{1}{A} \prod_i \left( \frac{P_i}{\alpha_i \lambda_i} \right)^{\alpha_i}$$

Rather than code the Cobb-Douglas function as a special case, many modelers choose to replace the elasticity of 1 with a value close to 1 such as 1.01. This would have only marginal repercussions on the results.

### Perfect substitution

The third special case is for a substitution elasticity of infinity. In this case  $\rho$  takes the value of 1 and the primal function is a straight linear aggregation of the inputs. The optimal demand conditions cannot be used in the case of an infinite substitution elasticity. In its stead, the optimal demand condition is replaced with the law-of-one-price, adjusted by efficiency differentials, and the zero profit condition is replaced with the CES primal function, i.e. the linear weighted aggregation of the inputs:

$$\frac{P_i}{\alpha_i \lambda_i} = P$$

$$V = \sum_i \alpha_i \lambda_i X_i$$

The aggregation function can be replaced by the zero profit condition:<sup>5</sup>

$$P.V = \sum_i P_i X_i$$

#### A.1.3 Calibration of the CES function

Calibration typically involves inverting functional forms to evaluate the value of a parameter given initial values for variables. Prices and volumes,  $P_i$ ,  $X_i$ ,  $V$  and  $P$ , are normally initialized to a given database or SAM. This may or may not include actual price/volume splits. If not, prices will typically be initialized at unit value—potentially adjusted for a price wedge such as a tax or a margin. The substitution elasticities are also normally inputs—either derived from econometric estimation, other data bases or models, or from a literature review. This leaves the parameters  $\lambda_i$ ,  $\alpha_i$  and  $A$  to calibrate. The technology parameters are normally associated with dynamics, so there is little reason not to initialize them to unit value as they can be incorporated in the initial share parameter value without any loss in generality. Thus, the only parameters left to calibrate are the  $\alpha_i$  from which it is possible to derive the primal share parameters,  $a_i$ , if needed. The calibration formula is derived from the inversion of equation (A-1):

$$\alpha_i = \left( \frac{X_i}{V} \right) \left( \frac{P_i}{P} \right)^\sigma (A \cdot \lambda_i)^{1-\sigma} = \left( \frac{X_i}{V} \right) \left( \frac{P_i}{P} \right)^\sigma$$

The right-most term is the most used formula where the technology parameters are explicitly set to 1.<sup>6</sup>

<sup>5</sup> Modelers have the choice of using the primal aggregation function or the revenue function. The latter holds in all three special cases for the substitution elasticity.

<sup>6</sup> In many introductions to CGE models, the calibration formulas explicitly exclude the price term. This is a dangerous practice that can lead to model bugs that can be hard to detect. It is best to explicitly initialize prices to 1 and use the correct calibration formula. In fact, one way to test model calibration and specification is to initialize prices to an arbitrary value and initialize volumes subject to these prices. Simulating a counterfactual with no shocks should replicate the initial data solution. If not, there is an error in initialization, calibration and/or specification.

### A.1.4 Comparative statics

#### Elasticities

This section will derive some of the key elasticities of the CES function. The first relationship is the elasticity of the aggregate price with respect to a component price:

$$\frac{\partial P}{\partial P_i} \frac{P_i}{P} = s_i = \frac{P_i X_i}{P.V}$$

The elasticity of the aggregate price relative to an input price is equal to the budget share, irrespective of the substitution elasticity. The matrix of own- and cross-price elasticities, holding the aggregate volume constant is given by the following formula:

$$\varepsilon_{ij} = \frac{\partial X_i}{\partial P_j} \frac{P_j}{X_i} = \sigma(s_j - \delta_{ij})$$

where  $\delta_{ij}$  is the so-called Kronecker's delta that takes the value 1 for  $i$  equal to  $j$ , else it takes the value 0. Since  $\sigma$  is positive, all components are gross substitutes in the CES.

#### Example: Trade elasticities

In this section we derive the elasticity of demand for a good sourced in region  $s$ , using the standard dual nested Armington formulation. Let  $XA$  be the top level Armington demand for a commodity that is derived from a generic demand function—for example an LES, with an associated price  $PA$ . The Armington demand is split into two components,  $XD$ , a domestic good, and  $XMT$ , an aggregate (or composite import demand), with prices  $PD$  and  $PMT$  respectively. The aggregate import bundle,  $XMT$ , is broken out by region of origin to determine  $XM_s$ , import from region  $s$ , with a domestic import price of  $PM_s$  (i.e. the landed or CIF price tariff inclusive).

Given  $XA$ , the following equations determine the top-level demand system, where  $\sigma^m$  is the top-level substitution elasticity (i.e. between the domestic goods and aggregate imports):

$$XD = \alpha^d \left( \frac{PA}{PD} \right)^{\sigma^m} XA \quad (\text{A-4})$$

$$XMT = \alpha^m \left( \frac{PA}{PMT} \right)^{\sigma^m} XA \quad (\text{A-5})$$

$$PA = \left[ \alpha^d PD^{1-\sigma^m} + \alpha^m PMT^{1-\sigma^m} \right]^{1/(1-\sigma^m)} \quad (\text{A-6})$$

The next set of equations determines imports by region of origin, where  $\sigma^w$  represents the elasticity of substitution for imports across region of origin:

$$XM_s = \alpha_s^w \left( \frac{PMT}{PM_s} \right)^{\sigma^w} XMT \quad (\text{A-7})$$

$$PMT = \left[ \sum_s \alpha_s^w PM_s^{1-\sigma^w} \right]^{1/(1-\sigma^w)} \quad (\text{A-8})$$

The elasticity we would like to derive is the following:

$$\varepsilon_s^w = - \frac{\partial XM_s}{\partial PM_s} \frac{PM_s}{XM_s} \quad (\text{A-9})$$

In other words, by how much would demand for imports from region  $s$  change relative to a change in its own price (including possibly a reduction (or increase) in the applied tariff).

The starting point for the derivation is equation (A-7) from which it is possible to derive the following expression:

$$\begin{aligned}\varepsilon_s^w &= -\frac{\partial XM_s}{\partial PM_s} \frac{PM_s}{XM_s} = -\left[(1-s_s)\sigma^w - \frac{\partial XMT}{\partial PM_s} \frac{PM_s}{XMT}\right] \\ &= -\left[(1-s_s)\sigma^w + s_s \left(-\frac{\partial XMT}{\partial PMT} \frac{PMT}{XMT}\right)\right]\end{aligned}\quad (\text{A-10})$$

Equation (A-10) shows that the trade elasticity is equal to the weighted average of the second-level Armington elasticity and the elasticity of demand for aggregate imports relative to the aggregate price of imports, where the weights are given by the import share of region  $s$  in total imports:

$$s_s = \frac{PM_s \cdot XM_s}{PMT \cdot XMT}$$

The final term in equation (A-10), not surprisingly takes the following form as derived from equation (A-5):

$$\begin{aligned}-\frac{\partial XMT}{\partial PMT} \frac{PMT}{XMT} &= -\left[(1-s_m)\sigma^m - \frac{\partial XA}{\partial PMT} \frac{PMT}{XA}\right] \\ &= -\left[(1-s_m)\sigma^m + s_m \left(-\frac{\partial XA}{\partial PA} \frac{PA}{XA}\right)\right]\end{aligned}\quad (\text{A-11})$$

Similar to equation (A-10), equation (A-11) shows that the aggregate elasticity of import demand is the weighted average of the top-level Armington demand and the overall price elasticity of the (Armington) demand with respect to the (Armington) price, where the weight is given by the (aggregate) import share in total demand:

$$s_m = \frac{PMT \cdot XMT}{PA \cdot XA}$$

These equations use the following expressions that provide the elasticity of the composite CES price with respect to its individual price components:

$$\begin{aligned}\frac{\partial PMT}{\partial PM_s} \frac{PM_s}{PMT} &= s_s \\ \frac{\partial PA}{\partial PMT} \frac{PMT}{PA} &= s_m\end{aligned}$$

Finally, the trade elasticity is given by the following expression:

$$\varepsilon_s^w = -\left[(1-s_s)\sigma^w + s_s \left((1-s_m)\sigma^m + s_m \epsilon^d\right)\right] \quad (\text{A-12})$$

where

$$\epsilon^d = -\frac{\partial XA}{\partial PA} \frac{PA}{XA}$$

### Formulas in percent differences

It is useful in terms of comparative static analyses to convert the basic equations into percent differences. It is easy to trace out the impacts of a change in one of the 'exogenous' variables on demand and the overall price index. This is also the form of the equations used for models implemented in GEMPACK such as MONASH-style models.

The following expressions convey expressions (A-1) and (A-2) into their percent difference form:

$$\frac{\partial X_i}{X_i} = \dot{X}_i = \dot{V} + \sigma (\dot{P} - \dot{P}_i) + (\sigma - 1) (\dot{A} + \dot{\lambda}_i)$$

$$\frac{\partial P}{P} = \dot{P} = -\dot{A} + \sum_i s_i \dot{P}_i - \sum_i s_i \dot{\lambda}_i = -\dot{A} + \sum_i s_i (\dot{P}_i - \dot{\lambda}_i)$$

Thus the percent change in the unit cost,  $P$ , for a change in the input price,  $P_i$ , all else equal, is (approximately) the value share of component  $i$ —as already noted above.

### A.1.5 Growth Accounting

Use can be made of the linearization above to derive the linearized growth accounting formula:

$$\frac{\Delta V}{V} = \frac{\Delta A}{A} + \sum_i s_i \frac{\Delta x_i}{x_i} + \sum_i s_i \frac{\Delta \lambda_i}{\lambda_i}$$

### A.1.6 Parameter twists

#### The basic analytics

This final section on the CES describes how to adjust the share parameters in a dynamic scenario under a specific assumption—this is called the twist adjustment and is a core feature of the dynamic MONASH model, see [Dixon and Rimmer \(2002\)](#). The basic idea is to alter the share parameter, in a two-component CES, to target a given change in the ratio of the two components, however, with neutral impacts on the aggregate cost. For example, the target may be a cost-neutral increase in the capital/labor ratio by  $x\%$ , or an increase in the import to domestic ratio of  $y\%$ .

The ratio of the two components is given by the following expression using equation (A-1) as the starting point:

$$R = \frac{\alpha_1 \lambda_1^{\sigma-1} P_2^\sigma}{\alpha_2 \lambda_2^{\sigma-1} P_1^\sigma}$$

The idea is to move the initial ratio,  $R_{t-1}$  to  $R_t$  by  $tw$  percent.

$$\frac{R_t}{R_{t-1}} = (1 + tw_t)$$

Using the formulas above, we have:

$$\frac{R_t}{R_{t-1}} = (1 + tw_t) = \frac{\left(\frac{\lambda_{1,t}}{\lambda_{1,t-1}}\right)^{\sigma-1}}{\left(\frac{\lambda_{2,t}}{\lambda_{2,t-1}}\right)^{\sigma-1}} = \frac{(1 + \pi_{1,t})^{\sigma-1}}{(1 + \pi_{2,t})^{\sigma-1}}$$



The  $\pi$  variables represent the growth (either positive or negative) that will be applied to the technology parameters under the assumption of cost-neutral technological change. We can start with the dual cost function for year  $t$ , but with year  $t - 1$  prices:

$$\begin{aligned} P_{t-1}^{1-\sigma} &= \alpha_1 \left( \frac{P_{1,t-1}}{\lambda_{1,t}} \right)^{1-\sigma} + \alpha_2 \left( \frac{P_{2,t-1}}{\lambda_{2,t}} \right)^{1-\sigma} \\ &= \alpha_1 (1 + \pi_{1,t})^{\sigma-1} \left( \frac{P_{1,t-1}}{\lambda_{1,t}} \right)^{1-\sigma} + \alpha_2 (1 + \pi_{2,t})^{\sigma-1} \left( \frac{P_{2,t-1}}{\lambda_{2,t}} \right)^{1-\sigma} \end{aligned}$$

Recall that the share equation is given by:

$$s_{i,t-1} = \alpha_i \lambda_{i,t-1}^{\sigma-1} \left( \frac{P_t}{P_{i,t-1}} \right)^{\sigma-1}$$

Dividing through the expression above by  $P_t^{1-\sigma}$  and inserting the share expressions for year  $t - 1$ , we end up with:

$$1 = s_{1,t-1} (1 + \pi_{1,t})^{\sigma-1} + s_{2,t-1} (1 + \pi_{2,t})^{\sigma-1}$$

Solving in terms of  $\pi_1$ , we have:

$$(1 + \pi_{1,t})^{\sigma-1} = \frac{1 - s_{2,t-1} (1 + \pi_{2,t})^{\sigma-1}}{s_{1,t-1}}$$

and this can be inserted into the twist target formula to get:

$$1 + tw_t = \frac{1 - s_{2,t-1} (1 + \pi_{2,t})^{\sigma-1}}{s_{1,t-1} (1 + \pi_{2,t})^{\sigma-1}} = \frac{(1 + \pi_{2,t})^{1-\sigma} - s_{2,t-1}}{s_{1,t-1}}$$

Finally,  $\pi_2$  can be isolated to yield:

$$1 + \pi_{2,t} = [s_{1,t-1} (1 + tw_t) + s_{2,t-1}]^{1/(1-\sigma)} = [1 + s_{1,t-1} tw_t]^{1/(1-\sigma)}$$

We can re-insert this into the expression above to derive an expression for  $\pi_1$ :

$$1 + \pi_{1,t} = \left[ \frac{1 + s_{1,t-1} tw_t}{1 + tw_t} \right]^{1/(1-\sigma)}$$

Finally, the productivity update formulas that incorporate the twist adjustment take the form:

$$\begin{aligned} \lambda_{1,t} &= (1 + \pi_{1,t}) \lambda_{1,t-1} = \left[ \frac{1 + s_{1,t-1} tw_t}{1 + tw_t} \right]^{1/(1-\sigma)} \lambda_{1,t-1} \\ \lambda_{2,t} &= (1 + \pi_{2,t}) \lambda_{2,t-1} = [1 + s_{1,t-1} tw_t]^{1/(1-\sigma)} \lambda_{2,t-1} \end{aligned}$$

It is possible to generalize these formulas by partitioning the set of CES components into two sets—a set indexed by 1 that is the target set, and a set indexed by 2 that is the complement. For example, think of a set of electricity technologies that includes conventional and advanced. It is possible then to provide the same twist to all of the new technologies relative to the conventional technologies. The only change in the formulas above is that the share variable for the single component is replaced by the sum of the shares for the bundle of components:

$$\lambda_{1,t} = (1 + \pi_{1,t}) \lambda_{1,t-1} = \left[ \frac{1 + tw_t \sum_{i \in 1} s_{i,t-1}}{1 + tw_t} \right]^{1/(1-\sigma)} \lambda_{1,t-1}$$

$$\lambda_{2,t} = (1 + \pi_{2,t})\lambda_{2,t-1} = \left[ 1 + tw_t \sum_{i \in 1} s_{i,t-1} \right]^{1/(1-\sigma)} \lambda_{2,t-1}$$

### Converting to percent differences

The  $\pi$  factors reflect a percentage change in the relevant productivity factors for each of the components. Using a Taylor series approximation, the formulas above can be converted to a linear equation that is used by the Monash-style models. For the first component, we have:

$$\pi_1 = F(tw) = \left[ \frac{1 + s_1 tw}{1 + tw} \right]^{1/(1-\sigma)} - 1 \approx F(0) + tw.F'(0) = -tw \frac{1 - s_1}{1 - \sigma}$$

For the second component we have:

$$\pi_2 = F(tw) = [1 + s_1 tw]^{1/(1-\sigma)} - 1 \approx F(0) + tw.F'(0) = tw \frac{s_1}{1 - \sigma}$$

Note that in the Monash models, the signs are reversed because the productivity factors divide the volume components whereas in the formulation above the productivity factors are multiplicative.

### Examples of twisting the share parameters

We demonstrate these concepts with two examples. The first is a CES production function of capital and labor, where the labor share is 60% and the capital/labor substitution elasticity (i.e.  $\sigma$ ) is set to 0.9. Prices are initialized at 1, therefore the original capital/labor ratio is 2/3. The target is to raise the capital/labor ratio 10% assuming cost neutrality. Table A.1 shows the key results. Labor efficiency would increase by 48% and capital efficiency would decline by 43%.

Table A.1: **Example of capital/labor twist**

	Labor	Capital	Capital/labor ratio
Initial	60.0	40.0	0.6667
After twist	57.7	42.3	0.7333
Percent change	-3.8	5.8	10.0
Growth factor	0.48	-0.43	

The second example comes from trade and the Armington assumption. Assume an 80/20 split between domestic goods and imports in value and volume implying a ratio of imports to demand of domestic goods of 0.25. Table A.2 shows the twist parameters needed to achieve an increase in this ratio of 10 percent with an Armington elasticity of 2. The preference parameter for imports increases by nearly 8 percent, while that for domestic goods decreases by 2 percent.

Table A.2: **Example of Armington import/domestic twist**

	Domestic	Import	Import/domestic ratio
Initial	80.0	20.0	0.250
After twist	78.4	21.6	0.275
Percent change	-2.0	7.8	10.0
Growth factor	-0.02	0.08	

### A.1.7 Summary

In summary, the CES functional form is often used as a production (or sub-production) function that combines two or more inputs to form output (or an intermediate composite bundle), under the assumption of cost minimization. It is also frequently used to maximize utility (or sub-utility) over a set of two or more goods, again with the assumption of cost minimization. Table A.3 highlights the two main expressions to emerge from the optimization—the derived demand functions,  $X_i$ , and the CES dual price expression,  $P$ . The top row shows the expression with all technology parameters initialized at 1, and the bottom row the most generic version.

Table A.3: Key equations for CES implementation

	Demand	Aggregate price
Basic	$X_i = \alpha_i V \left( \frac{P}{P_i} \right)^\sigma$	$P = \left[ \sum_i \alpha_i P_i^{1-\sigma} \right]^{\frac{1}{1-\sigma}}$
with full technology	$X_i = \alpha_i (A \lambda_i)^{\sigma-1} V \left( \frac{P}{P_i} \right)^\sigma$	$P = \frac{1}{A} \left[ \sum_i \alpha_i \left( \frac{P}{\lambda_i} \right)^{1-\sigma} \right]^{\frac{1}{1-\sigma}}$

## A.2 The CET Function

### A.2.1 The basic formulation

This section describes the constant-elasticity-of-transformation (CET) function. The CET function is often used to describe a transformation frontier between two or more outputs. For example, a producer may produce two or more products and decides how much of each to produce based on market conditions, i.e. relative prices. The CET is often used to represent a producer's decision on the allocation of output between domestic and foreign markets. Another example is land supply, where land will be allocated across different uses according to the relative returns. The transformation elasticity is assumed to be uniform between any pair of outputs and therefore is analogous to the demand-based CES function described in detail above. The exposition of the CET will be much more succinct than that of the CES because most of the derivations can be derived in a similar fashion.

The CET can be setup as a revenue maximization problem, subject to a transformation frontier:

$$\max_{X_i} \sum_i P_i X_i$$

subject to

$$V = A \left[ \sum_i g_i (\lambda_i X_i)^v \right]^{1/v}$$

where  $V$  is the aggregate volume (e.g. aggregate supply),  $X_i$  are the relevant components (sector-specific supply),  $P_i$  are the corresponding prices,  $g_i$  are the CET (primal) share parameters, and  $\nu$  is the CET exponent. The CET exponent is related to the CET transformation elasticity,  $\omega$  via the following relation:

$$\nu = \frac{\omega + 1}{\omega} \Leftrightarrow \omega = \frac{1}{\nu - 1}$$

The transformation elasticity is assumed to be positive. Solution of this maximization problem leads to the following first order conditions:

$$X_i = \gamma_i (A \lambda_i)^{-1-\omega} \left( \frac{P_i}{P} \right)^\omega V \quad (\text{A-13})$$

and

$$P = \frac{1}{A} \left[ \sum_i \gamma_i \left( \frac{P_i}{\lambda_i} \right)^{1+\omega} \right]^{1/(1+\omega)} \quad (\text{A-14})$$

where the  $\gamma_i$  parameters are related to the primal share parameters,  $g_i$ , by the following formula:

$$\gamma_i = g_i^{-\omega} \Leftrightarrow g_i = \left( \frac{1}{\gamma_i} \right)^{1/\omega}$$

From expression A-13, and ignoring the technology parameters for the moment, the clear difference with the CES expression for optimal demand (equation A-1) is that the component price is in the numerator and the aggregate price in the denominator. This is intuitively logical. If the supply price to a market goes up relative to the average market price, one would anticipate that supply would increase to that market. The greater the transformation elasticity the greater are the market shifts.

Calibration is similar to the CES case. Prices and volumes are initialized using base year data. Equation (A-13) can then be inverted to calculate the share parameters,  $\gamma_i$ , with typically the technology parameters initialized to the value 1. In most implementations, there is no need to carry around the primal share parameters, nor the primal exponent.

The main interesting case for the CET is the case of perfect transformation, i.e. the transformation elasticity is infinity. In this case the CET exponent is 0 and the aggregation function is a linear weighted aggregation of the components. The standard CET equations are then replaced by the law-of-one price and the linear aggregation function (or alternatively, the zero profit condition).

$$\begin{aligned} \frac{P_i}{A \lambda_i} &= P \forall i \\ A \sum_i \lambda_i X_i &= X \end{aligned}$$

### A.2.2 Converting to percent differences

It is easier to interpret or decompose the results of a simulation by looking at the CET equations in percent differences form—that is the standard form for MONASH-style models and implementation in GEMPACK. The following equations show the equations in percent difference form:

$$\begin{aligned} \frac{\partial X_i}{X_i} = \dot{X}_i &= \dot{V} + \omega \left( \dot{P}_i - \dot{P} \right) - (\omega + 1) \left( \dot{\lambda}_i + \dot{A} \right) \\ \frac{\partial P}{P} = \dot{P} &= -\dot{A} + \sum_i s_i \dot{P}_i - \sum_i s_i \dot{\lambda}_i = -\dot{A} + \sum_i s_i \left( \dot{P}_i - \dot{\lambda}_i \right) \end{aligned}$$

where the variable  $s_i$  is the value share of component  $i$  in total revenue:

$$s_i = \frac{P_i X_i}{P \cdot V} = \gamma_i \left( \frac{P_i}{A \cdot \lambda_i \cdot P} \right)^{\omega+1}$$

### A.2.3 Twists and the CET

It is relatively easy to show that the analogous twist formulas for the CET are:

$$1 + \pi_{1,t} = \left[ \frac{1 + tw_t}{1 + s_{1,t-1} tw_t} \right]^{1/(1+\omega)} \quad 1 + \pi_{2,t} = \left[ \frac{1}{1 + s_{1,t-1} tw_t} \right]^{1/(1+\omega)}$$

## A.3 Modified CES and CET that incorporate additivity

The standard CET supply allocation specification does not preserve physical additivity, i.e. the sum of the volume components is not necessarily equal to the total volume. There are a number of alternative specifications that do preserve volume homogeneity, for example the multinomial logit. One alternative, described below, uses a modified form of the CET preference function. This specification has been used for labor and land supply allocations (see respectively [Dixon and Rimmer \(2006\)](#) and [Giesecke et al. \(2013\)](#)).

### A.3.1 The CET implementation

The CET alternative involves solving the following optimization:

$$\max_{X_i} U = \left[ \sum_i g_i (\lambda_i P_i X_i)^\nu \right]^{1/\nu}$$

subject to the constraint:

$$V = \sum_i X_i$$

The variable definitions are similar to above,  $X_i$  are the volume components,  $P_i$  are the relevant component prices and  $V$  is aggregate volume. The  $\lambda_i$  parameters are preference parameters. The CET utility function is not simply a function of the volumes, but explicitly a function of the preference-adjusted revenues of the individual components. The closed-form solution to the above system is the following set of equations:

$$X_i = \gamma_i V \left( \frac{\lambda_i P_i}{P^c} \right)^\omega \tag{A-15}$$

$$P^c = \left[ \sum_i \gamma_i (\lambda_i P_i)^\omega \right]^{1/\omega} \tag{A-16}$$

Both equations are similar to their standard CET counterparts, but with some differences. First,  $P^c$  is a price index, but it is not the average price of the components, i.e.  $P^c X \neq \sum_i P_i X_i$ . Second,

this price index is based on  $\omega$  not  $1 + \omega$  as in the standard CET dual price expression. The revenue correct price index is defined by the following formula:

$$P = \frac{\sum_i \gamma_i \lambda_i^\omega P_i^{\omega+1}}{\sum_i \gamma_i \lambda_i^\omega P_i^\omega} = \frac{\sum_i \gamma_i \lambda_i^\omega P_i^{\omega+1}}{(P^c)^\omega} = \sum_i \gamma_i P_i \left( \frac{\lambda_i P_i}{P^c} \right)^\omega = \sum_i \frac{X_i}{V} P_i \quad (\text{A-17})$$

The other transformations include:

$$\gamma_i = g_i^{1+\omega}$$

$$\omega = \frac{\nu}{1 - \nu} \iff \nu = \frac{\omega}{1 + \omega}$$

It is worth noting that the relation between  $\omega$  and  $\nu$  differs from the standard CET relation as the respective formula is inverted. The implication of this is that  $\nu$  is bounded below by 0 instead of  $\infty$ , but is otherwise positive over the entire (positive) range of  $\omega$ . And, in both the standard and adjusted CET  $\nu$  converges to 1 as  $\omega$  converges to  $\infty$ . As regards calibration, there is an extra degree of freedom as the value for utility is not specified. It is easiest to simply set  $P^c$  to 1 as for given  $P_i$  and  $\lambda_i$  the calibration of the  $\gamma$  parameters is straightforward:

$$\gamma_i = \frac{X_i}{V} \left( \frac{\lambda_i P_i}{P^c} \right)^{-\omega}$$

If prices and technology or preference parameters are initialized at 1, the calibrated  $\gamma$  parameters are equal to the initial volume shares.

Converting this to a Monash-style equation in percent differences, the derived supply function is:

$$\dot{X}_i = \dot{V} + \omega \left[ \dot{P}_i + \dot{\lambda}_i - \sum_{j=1}^n \frac{X_j}{V} (\dot{P}_j + \dot{\lambda}_j) \right]$$

This equation uses volume shares as weights for cross-price (and cross-preference) effects. In the standard CET formulation, value shares are used as weights.

The standard specification needs some modifications for two special cases—perfect transformation and perfect immobility. The case of perfect transformation, i.e. a transformation elasticity of  $\infty$ , leads to all prices moving in unison with the aggregate price index. Thus equation (A-15) is replaced with the following expression:

$$\lambda_i P_i = P^c$$

where  $\lambda_i$  is calibrated to the initial price ratios. The price index expression, equation (A-16) is replaced with the volume constraint:

$$V = \sum_i X_i$$

In the model implementation of the adjusted CET, this latter expression can be used in all cases and can replace equation (A-16).

The case of zero mobility is readily implemented by dropping completely equation (A-16) (or its equivalent, i.e. the volume adding up constraint). With a transformation elasticity of 0, the

price composite index in equation (A-15) simply drops out and the volume components are in strict proportion to the aggregate volume.

### A.3.2 ACET and twists

It is relatively easy to show that the analogous twist formulas for the ACET are:

$$1 + \pi_{1,t} = \left[ \frac{1 + tw_t}{1 + s_{1,t-1}tw_t} \right]^{1/\omega} \quad 1 + \pi_{2,t} = \left[ \frac{1}{1 + s_{1,t-1}tw_t} \right]^{1/\omega}$$

where the volume shares replace the value shares from the standard CET twists and the exponent differs.<sup>7</sup>

### A.3.3 The CES implementation

The adjusted CET (and CRETH) functions replace their counterparts for the allocation problem that preserves additivity. Analogous specifications exist for the CES and CRESH functions that emulate the implementation of their standard counterparts but also allow for additivity.

The CES alternative involves solving the following optimization:

$$\min_{X_i} U = \left[ \sum_i a_i (\lambda_i P_i X_i)^\rho \right]^{1/\rho}$$

subject to the constraint:

$$V = \sum_i X_i$$

As in the case of the adjusted CET, the adjusted CES utility function is a function of the preference adjusted cost components. The closed-form solution to the above system is the following set of equations:

$$X_i = \alpha_i V \left( \frac{P^c}{\lambda_i P_i} \right)^\sigma \quad (\text{A-18})$$

$$P^c = \left[ \sum_i \alpha_i (\lambda_i P_i)^{-\sigma} \right]^{-1/\sigma} \quad (\text{A-19})$$

Both equations are similar to their standard CES counterparts, but with some differences. First,  $P^c$  is a price index, but it is not the average price of the components, i.e.  $P^c X \neq \sum_i P_i X_i$ . Second, this price index is based on  $-\sigma$  not  $1-\sigma$  as in the standard CES dual price expression. The revenue correct price index is defined by the following formula:

$$P = \frac{\sum_i \alpha_i \lambda_i^\sigma P_i^{1-\sigma}}{\sum_i \alpha_i \lambda_i^\sigma P_i^\sigma} = \frac{\sum_i \alpha_i \lambda_i^\sigma P_i^{1-\sigma}}{(P^c)^{-\sigma}} = \sum_i \alpha_i P_i \left( \frac{\lambda_i P_i}{P^c} \right)^{-\sigma} = \sum_i \frac{X_i}{V} P_i \quad (\text{A-20})$$

The other transformations include:

<sup>7</sup> These formulas are consistent with no change in the ACET composite price. However, the average price may vary as well as utility.

$$\alpha_i = a_i^{1-\sigma}$$

$$\sigma = \frac{\rho}{\rho - 1} \iff \rho = \frac{\sigma}{\sigma - 1}$$

It is worth noting that the relation between  $\sigma$  and  $\rho$  differs from the standard CES relation as the respective formula is inverted. The implication of this is that  $\rho$  is bounded below by 0 instead of  $-\infty$ . It decreases towards  $-\infty$ , as  $\sigma$  increases towards 1, which is a discontinuity point. It decreases from  $\infty$  towards 1 as  $\sigma$  increases from 1 to  $\infty$ .

It is relatively easy to show that the following simple expression holds for the utility function:

$$U = P^c V \tag{A-21}$$

As regards calibration, there is an extra degree of freedom as the value for utility is not specified. It is easiest to simply set  $P^c$  to 1 as for given  $P_i$  and  $\lambda_i$  the calibration of the  $\alpha$  parameters is straightforward:

$$\alpha_i = \frac{X_i}{V} \left( \frac{\lambda_i P_i}{P^c} \right)^\sigma$$

If prices and technology or preference parameters are initialized at 1, the calibrated  $\alpha$  parameters are equal to the initial volume shares.

Converting this to a Monash-style equation in percent differences, the derived demand function is:

$$\dot{X}_i = \dot{V} - \sigma \left[ \dot{P}_i + \dot{\lambda}_i - \sum_{j=1}^n \frac{X_j}{V} (\dot{P}_j + \dot{\lambda}_j) \right]$$

This equation uses volume shares as weights for cross-price (and cross-preference) effects. In the standard CES formulation, value shares are used as weights.

### A.3.4 Using twists with the adjusted CES

The 'twist' idea described for the normal CES can be applied to the adjusted CES. The concept is somewhat different given the type of optimization problem posed. Rather than change the share parameters in a given direction with cost neutrality, the idea is to change the share parameters with utility neutrality. The problem posed, therefore, is to change the ratio of demand for two goods by a specified amount, while maintaining the same level of utility.

The ratio of the two components is given by the following expression using equation (A-18) as the starting point:

$$R = \frac{\alpha_1 \lambda_2 P_2^\sigma}{\alpha_2 \lambda_1 P_1^\sigma}$$

The idea is to move the initial ratio,  $R_{t-1}$  to  $R_t$  by  $tw$  percent.

$$\frac{R_t}{R_{t-1}} = (1 + tw_t)$$



while holding  $U$  constant. The two expressions above imply that the preference shifters, given by the  $\pi$  parameters, are linked via the following expression:

$$1 + \pi_2 = (1 + \pi_1) (1 + tw)^{1/\sigma} \quad (\text{A-22})$$

Given equation (A-21), holding  $U$  constant is equivalent to holding the price index,  $P^c$ , constant as well (for a fixed aggregate volume). Thus we can solve the following equation for the parameter  $\pi_1$ :

$$\begin{aligned} (P_{t-1}^c)^{-\sigma} &= \alpha_1 (P_{1,t-1} \lambda_{1,t-1})^{-\sigma} + \alpha_2 (P_{2,t-1} \lambda_{2,t-1})^{-\sigma} \\ &= \alpha_1 (P_{1,t-1} \lambda_{1,t-1} (1 + \pi_{1,t}))^{-\sigma} + \alpha_2 (P_{2,t-1} \lambda_{2,t-1} (1 + \pi_{2,t}))^{-\sigma} \\ &= \alpha_1 (P_{1,t-1} \lambda_{1,t-1} (1 + \pi_{1,t}))^{-\sigma} + \alpha_2 \left( P_{2,t-1} \lambda_{2,t-1} (1 + \pi_{1,t}) (1 + tw)^{1/\sigma} \right)^{-\sigma} \\ &= (P_t^c)^{-\sigma} \end{aligned}$$

The  $\pi$  variables represent the growth (either positive or negative) that will be applied to the preference parameters under the assumption of utility-preserving preference shifts. This formula can be written in terms of the initial volume shares,  $r_i = X_i/V$ , simplified and re-arranged to yield:

$$(1 + \pi_1)^\sigma = r_1 + \frac{r_2}{1 + tw}$$

and when re-inserted in equation (A-22) we get:

$$(1 + \pi_2)^\sigma = \left( r_1 + \frac{r_2}{1 + tw} \right) (1 + tw)$$

The final formulas for the two twist parameters only depend on the initial volume shares, the substitution elasticity and the level of the 'twist':

$$\pi_1 = \left[ \frac{1 + r_1 \tau}{1 + \tau} \right]^{1/\sigma} - 1 \quad (\text{A-23})$$

$$\pi_2 = [1 + r_1 \tau]^{1/\sigma} - 1 \quad (\text{A-24})$$

It is possible to generalize these formulas by partitioning the set of CES components into two sets—a set indexed by 1 that is the target set, and a set indexed by 2 that is the complement. For example, think of a set of electricity technologies that includes conventional and advanced. It is possible then to provide the same twist to all of the new technologies relative to the conventional technologies. The only change in the formulas above is that the volume share variable for the single component is replaced by the sum of the volume shares for the bundle of components:

$$\begin{aligned} \lambda_{1,t} &= (1 + \pi_{1,t}) \lambda_{1,t-1} = \left[ \sum_{i \in 1} r_{i,t-1} + \frac{\sum_{i \in 2} r_{i,t-1}}{1 + tw_t} \right]^{1/\sigma} \lambda_{1,t-1} \\ \lambda_{2,t} &= (1 + \pi_{2,t}) \lambda_{2,t-1} = \left[ \sum_{i \in 1} (1 + tw_t) r_{i,t-1} + \sum_{i \in 2} r_{i,t-1} \right]^{1/\sigma} \lambda_{2,t-1} \end{aligned}$$

### Converting to percent differences

The  $\pi$  factors reflect a percentage change in the relevant productivity factors for each of the components. Using a Taylor series approximation, the formulas above can be converted to a linear equation that is used by the Monash-style models. For the first component, we have:

$$\pi_1 = F(tw) = \left[ r_1 + \frac{r_2}{1 + tw} \right]^{1/\sigma} - 1 \approx F(0) + tw.F'(0) = -tw \frac{r_2}{\sigma}$$

For the second component we have:

$$\pi_2 = F(tw) = [r_1 (1 + tw) + r_2]^{1/\sigma} - 1 \approx F(0) + tw.F'(0) = tw \frac{r_1}{\sigma}$$

# Appendix B

## The demand systems

This appendix describes in full detail the consumer demand systems available in the ENVISAGE Model. The first section derives the main relations for the extended linear expenditure system (ELES). The second section describes the constant differences in elasticity (CDE) demand system.

### B.1 The extended linear expenditure system (ELES)

Many models assume separability in household decision making between saving and current consumption. Lluch and Howe<sup>1</sup> introduced a relatively straightforward extension of the LES consumer demand function to include the saving decision simultaneously with the allocation of income on goods and services, this has become known as the extended linear expenditure system or the ELES. The ELES is based on consumers maximizing their intertemporal utility between a bundle of current consumption and an expected future consumption bundle represented in the form of savings.

#### B.1.1 Basic formulation

The utility function of the ELES has the following form:

$$u = \prod_i (x_i - \gamma_i)^{\mu_i} \left( \frac{S}{P^s} \right)^{\mu_s} \quad (\text{B-1})$$

with

$$\sum_i \mu_i + \mu_s = 1 \quad (\text{B-2})$$

where  $u$  is utility,  $x$  is the vector of consumption goods,  $S$  is household saving (in value),  $P^s$  is the price of saving, and  $\mu$  and  $\gamma$  are ELES parameters.

The consumer solves the following problem:

$$\max \prod_i (x_i - \gamma_i)^{\mu_i} \left( \frac{S}{P^s} \right)^{\mu_s}$$

subject to

---

<sup>1</sup> See [Lluch \(1973\)](#) and [Howe \(1975\)](#).

$$\sum_{i=1}^n p_i x_i + S = Y$$

where  $p$  is the vector of consumer prices, and  $Y$  is disposable income. The demand functions are:

$$x_i = \gamma_i + \frac{\mu_i}{p_i} \left( Y - \sum_{j=1}^n p_j \gamma_j \right) \quad (\text{B-3})$$

$$S = \mu_s \left( Y - \sum_{j=1}^n p_j \gamma_j \right) = Y - \sum_{j=1}^n p_j x_j \quad (\text{B-4})$$

The term in parentheses is sometimes called supernumerary income, i.e. it is the income that remains after subtracting total expenditures on the so-called subsistence (or floor) expenditures as represented by the  $\gamma$  parameter. The parameter  $\mu$  then represents the marginal budget share out of supernumerary income.

### B.1.2 ELES elasticities

From the demand equation we can derive the income and price elasticities:

$$\eta_i = \frac{\mu_i Y}{p_i x_i} = \frac{\mu_i}{s_i} \quad \eta_s = \frac{\mu_s Y}{S} = \frac{\mu_s}{s} \quad (\text{B-5})$$

$$\varepsilon_i = \frac{\gamma_i (1 - \mu_i)}{x_i} - 1 \quad \varepsilon_s = -1 \quad (\text{B-6})$$

$$\varepsilon_{ij} = -\frac{\mu_i p_j \gamma_j}{p_i x_i} = -\frac{\mu_i p_j \gamma_j}{s_i Y} \quad \varepsilon_{sj} = -\frac{\mu_s p_j \gamma_j}{s Y} = -\frac{p_j \gamma_j}{Y^*} \quad (\text{B-7})$$

where  $s$  is the average propensity to save. Note that the matrix of elasticities can be collapsed to a single formula using the Kronecker factor:

$$\varepsilon_{ij} = -\frac{\mu_i p_j \gamma_j}{p_i x_i} - \delta_{ij} \frac{p_i x_i - p_i \gamma_i}{p_i x_i} = -\frac{\mu_i}{s_i Y} [\delta_{ij} Y^* + p_j \gamma_j] = -\eta_i \left[ \delta_{ij} \frac{Y^*}{Y} + \frac{p_j \gamma_j}{Y} \right] \quad (\text{B-8})$$

The last expression shows that there is clear linkage between the income and price elasticities. At the limit, when income is much larger than supernumerary income, the two are virtually identical in levels (with opposite signs).

### B.1.3 Welfare

With the addition of saving, the indirect utility function is given by:

$$v(p, Y) = \prod_i \left( \frac{\mu_i}{p_i} Y^* \right)^{\mu_i} \left( \frac{\mu_s}{P^s} Y^* \right)^{\mu_s} \quad (\text{B-9})$$

or

$$v(p, Y) = \frac{Y^*}{P} \quad (\text{B-10})$$

where

$$P = \prod_i \left( \frac{p_i}{\mu_i} \right)^{\mu_i} \left( \frac{P^s}{\mu_s} \right)^{\mu_s}$$

The expenditure function is derived by minimizing the cost of achieving a given level of utility,  $u$ . It is set-up as:

$$\min \sum_{i=1}^n p_i x_i + S$$

subject to

$$\prod_i (x_i - \gamma_i)^{\mu_i} \left( \frac{S}{P^s} \right)^{\mu_s} = u$$

The final expression for the expenditure function is:

$$E(p, u) = \sum_{i=1}^n p_i \gamma_i + uP \quad (\text{B-11})$$

where  $P$ , the aggregate price index (including the price of savings) is defined as above.

#### B.1.4 Calibration

Calibration of the ELES uses the budget share information from the base SAM, including the household saving share. Typically, calibration uses income elasticities for all of the  $n$  commodities represented in the demand system and uses equation (B-5) to derive the marginal budget shares,  $\mu_i$ . This procedure leads to a residual income elasticity, which in this case is the income elasticity of saving. The derived savings income elasticity may be implausible, in which case adjustments need to be made to individual income elasticities for the goods, or adjustments can be made on the group of goods, assuming some target for the savings income elasticity.

The first step is therefore to calculate the marginal budget shares using the average budget shares and the initial income elasticity estimates.

$$\mu_i = \frac{\eta_i p_i x_i}{Y} = \eta_i s_i$$

The savings marginal budget share is derived from the consistency requirement that the marginal budget shares sum to 1:

$$\mu_s = 1 - \sum_{i=1}^n \mu_i$$

Assuming this procedure leads to a plausible estimate for the savings income elasticity, the next step is to calibrate the subsistence minima,  $\gamma$ . This can be done by seeing that the demand equations, (B-3), are linear in the  $\gamma$  parameters. Note that in the case of the ELES the system of equation are of full rank because the  $\mu_i$  parameters do not sum to 1 (over the  $n$  commodities)—they only sum to 1 including the marginal saving share.<sup>2</sup> This may lead to calibration problems if the propensity to save is 0, which may be the case in some SAMs with poor households. The linear system can be written as:

<sup>2</sup> Note that the calibration and the setup of the ELES assume explicitly that the minimal expenditure on savings is zero.

$$C = I\gamma + MY - M\Pi\gamma$$

where  $I$  is an  $n \times n$  identity matrix,  $M$  is a diagonal matrix with  $\mu_i/P_i$  on the diagonal, and  $\Pi$  is a matrix, where each row is identical, each row being the transpose of the price vector. The above system of linear equations can be solved via matrix inversion for the parameter  $\gamma$ :

$$\gamma = A^{-1}C^*$$

where

$$A = I - M\Pi$$

$$C^* = C - MY$$

The matrices  $A$  and  $C^*$  are defined by:

$$A = [a_{ij}] = \begin{bmatrix} \delta_{ij} - \mu_i \frac{p_j}{p_i} \end{bmatrix} = \begin{cases} 1 - \mu_i & \text{if } i = j \\ -\mu_i \frac{p_j}{p_i} & \text{if } i \neq j \end{cases}$$

$$C^* = [c_i] = x_i - \frac{\mu_i Y}{p_i}$$

The  $A$  and  $C^*$  matrices are simplified if the price vector is initialized at 1:

$$A = [a_{ij}] = \begin{cases} 1 - \mu_i & \text{if } i = j \\ -\mu_i & \text{if } i \neq j \end{cases}$$

$$C^* = [c_i] = x_i - \mu_i Y$$

In GAMS one could invert the system of equations embodied in equation (B-3) directly by solving for the endogenous  $\gamma$  while holding all of the other variables and parameters fixed.

## B.2 An Implicitly Direct Additive Demand System (AIDADS)

Many commonly used utility functions typically exhibit poor Engel behavior—particularly in a dynamic framework. The CDE utility function, popularized by the GTAP model (see Hertel (1997)), has relatively constant income elasticities. The LES utility function has even worse behavior, as in the absence of any shifts in the underlying parameters, the LES converges relatively quickly to a Cobb-Douglas utility function as rapidly rising consumption tends to dominate the floor consumption parameters, even when adjusting the latter to take into account population growth. Rimmer and Powell (see Rimmer and Powell (1992b), Rimmer and Powell (1992a) and Rimmer and Powell (1996)) examine an extension to the standard LES demand system that in effect allows the marginal propensity to consumer parameter to be driven by changes in utility. Their utility function has been called An Implicitly Direct Additive Demand System, or AIDADS. The LES function is a special case of the AIDADS system where the marginal propensity variable is constant. This extension allows for more complex demand behavior, as well as providing better validation for observed changes in consumption patterns.<sup>3</sup>

<sup>3</sup> AIDADS has also been explored in the context of the GTAP model, see for example Yu et al. (2003).

### B.2.1 Basic formulation

AIDADS starts with the implicitly additive utility function given by:

$$\sum_i U_i(x_i, u) \equiv 1 \quad (\text{B-12})$$

Assume the following functional form for the utility function:

$$U_i = \mu_i \ln \left( \frac{x_i - \gamma_i}{Ae^u} \right) \quad (\text{B-13})$$

where

$$\mu_i = \frac{\alpha_i + \beta_i G(u)}{1 + G(u)} \quad (\text{B-14})$$

with the restrictions

$$\begin{aligned} \sum_i \alpha_i &= \sum_i \beta_i = 1 \\ 0 &\leq \alpha_i \leq 1 \\ 0 &\leq \beta_i \leq 1 \\ \gamma_i &< x_i \end{aligned}$$

Cost minimization implies the following:

$$\min \sum_i p_i x_i$$

subject to

$$\sum_i \mu_i \ln \left( \frac{x_i - \gamma_i}{Ae^u} \right) \equiv 1 \quad (\text{B-15})$$

The first order conditions lead to:

$$\lambda \frac{\partial U_i}{\partial x_i} = p_i = \lambda \frac{\mu_i}{x_i - \gamma_i} \Rightarrow \lambda \mu_i = p_i x_i - p_i \gamma_i \quad (\text{B-16})$$

Taking the sum over  $i$  and using the fact that the  $\mu_i$  sum to unity implies:

$$\lambda = \sum_i p_i x_i - \sum_i p_i \gamma_i = Y - \sum_i p_i \gamma_i = Y^* \quad (\text{B-17})$$

where  $Y$  is aggregate expenditure, and  $Y^*$ , sometimes referred to as supernumerary income, is residual expenditure after subtracting total expenditure on the so-called subsistence minima,  $\gamma$ .

Re-inserting equation (B-17) into (B-16) yields the consumer demand equations:

$$x_i = \gamma_i + \frac{\mu_i}{p_i} Y^* = \gamma_i + \frac{\mu_i}{p_i} \left[ Y - \sum_j p_j \gamma_j \right] \quad (\text{B-18})$$

Equation (B-18) is virtually identical to the LES demand equation. Similar to the linear expenditure system (LES), demand is the sum of two components—a subsistence minimum,  $\gamma$ , and

a share,  $\mu$ , of supernumerary income. Unlike the LES, the share parameter,  $\mu$ , is not constant, but depends on the level of utility itself. AIDADS collapses to the LES if each  $\alpha$  parameter is equal to the corresponding  $\beta$  parameter, with the ensuing function of utility,  $G(u)$ , dropping from equation (B-14).

### B.2.2 Elasticities

This section develops the main expressions for the income and price elasticities. These formulas will be needed to calibrate the initial parameters of the AIDADS function.

#### Income elasticities

To derive further properties of AIDADS requires specifying a functional form for  $G(u)$ . [Rimmer and Powell \(1996\)](#) propose the following:

$$G(u) = e^u \quad (\text{B-19})$$

The first step is to calculate the marginal budget share,  $\rho$ , defined as:

$$\rho_i = p_i \frac{\partial x_i}{\partial Y}$$

The following expression can be derived from equation (B-18):

$$\frac{\partial x_i}{\partial Y} = \frac{Y^*}{p_i} \frac{\partial \mu_i}{\partial Y} + \frac{\mu_i}{p_i} \frac{\partial Y^*}{\partial Y} = \frac{Y^*}{p_i} \frac{\partial \mu_i}{\partial u} \frac{\partial u}{\partial Y} + \frac{\mu_i}{p_i}$$

Thus:

$$\rho_i = \mu_i + Y^* \frac{\partial \mu_i}{\partial u} \frac{\partial u}{\partial Y} \quad (\text{B-20})$$

Expression (B-20) can be expanded in two steps—first evaluating the partial derivative of the share variable,  $\mu$ , with respect to utility, and then the more difficult calculation of the partial derivative of  $u$  with respect to income. The marginal share formula is:

$$\mu_i = \frac{\alpha_i + \beta_i e^u}{1 + e^u}$$

Its derivative is:

$$\frac{\partial \mu_i}{\partial u} = \frac{(1 + e^u)(\beta_i e^u) - (\alpha_i + \beta_i e^u)e^u}{(1 + e^u)^2} = \frac{e^u(\beta_i - \alpha_i)}{(1 + e^u)^2} \quad (\text{B-21})$$

Utility and income are combined in implicit form and thus we will invoke the implicit function theorem to calculate the partial derivative of  $u$  with respect to  $Y$ . First, insert equation (B-18) into equation (B-15):

$$\sum_i \mu_i \ln \left( \frac{x_i - \gamma_i}{A e^u} \right) = \sum_i \mu_i \ln \left( \frac{\mu_i Y^*}{A e^u p_i} \right) = 1$$

Expanding the latter expression yields:

$$f(u, Y) = \sum_i \mu_i \ln \left( \frac{\mu_i}{p_i} \right) + \ln(Y^*) - \ln(A) - u = 1 \quad (\text{B-22})$$



which provides the implicit relation between  $Y$  and  $u$ . The implicit function theorem states the following:

$$\frac{\partial u}{\partial Y} = -\frac{\partial f}{\partial Y} \left[ \frac{\partial f}{\partial u} \right]^{-1} \quad (\text{B-23})$$

The partial derivative of  $f$  with respect to  $Y$  is simply:

$$\frac{\partial f}{\partial Y} = \frac{1}{Y^*} \quad (\text{B-24})$$

The partial derivative of  $f$  with respect to  $u$  is:

$$\begin{aligned} \frac{\partial f}{\partial u} &= -1 + \sum_i \left[ \frac{\partial \mu_i}{\partial u} \ln \left( \frac{\mu_i}{p_i} \right) + \mu_i \frac{p_i}{\mu_i} \frac{\partial \mu_i}{\partial u} \right] \\ &= -1 + \frac{e^u}{(1+e^u)^2} \sum_i \left[ \left( \ln \left( \frac{\mu_i}{p_i} \right) + 1 \right) (\beta_i - \alpha_i) \right] \\ &= \frac{e^u}{(1+e^u)^2} \left[ \sum_i (\beta_i - \alpha_i) \ln (x_i - \gamma_i) - \frac{(1+e^u)^2}{e^u} \right] \\ &= \frac{e^u}{(1+e^u)^2} \Omega^{-1} \end{aligned} \quad (\text{B-25})$$

where

$$\Omega = \left[ \sum_i (\beta_i - \alpha_i) \ln (x_i - \gamma_i) - \frac{(1+e^u)^2}{e^u} \right]^{-1} \quad (\text{B-26})$$

The second line uses equation (B-21). In the third line, equation (B-18) substitutes for the expression in the logarithm, and the adding up constraint allows for the deletion of non-indexed variables. Substituting equations (B-24) and (B-25) into equation (B-23) yields:

$$\frac{\partial u}{\partial Y} = -\frac{\Omega}{Y^*} \frac{(1+e^u)^2}{e^u} \quad (\text{B-27})$$

Substituting equations (B-21) and (B-27) into equation (B-20) yields the following expression for  $\rho$ :

$$\rho_i = \mu_i - (\beta_i - \alpha_i) \Omega$$

The income elasticities are derived from the following expression:

$$\eta_i = \frac{\partial x_i}{\partial Y} \frac{Y}{x_i} = \frac{\partial x_i}{\partial Y} \frac{Y}{x_i} \frac{p_i}{p_i} = \frac{\rho_i}{s_i}$$

where  $s_i$  is the average budget share:

$$s_i = \frac{p_i x_i}{Y} = \frac{p_i \gamma_i}{Y} + \mu_i \frac{Y^*}{Y} = \mu_i + \left( \frac{p_i \gamma_i - \mu_i \sum_j p_j \gamma_j}{Y} \right)$$

Thus the income elasticity,  $\eta$ , is equal to the ratio of the marginal budget share,  $\rho$ , and the average budget share,  $s$ . Finally, equation (B-28) describes one formulation of the income elasticity:

$$\eta_i = \frac{\mu_i - (\beta_i - \alpha_i) \Omega}{s_i} \quad (\text{B-28})$$

### Price elasticity

The matrix of substitution elasticities is identical to the expression for the LES and has the form:

$$\sigma_{ij} = [\mu_j - \delta_{ij}] \frac{\mu_i Y^*}{s_i s_j Y} \quad (\text{B-29})$$

where  $\delta$  is the Kronecker delta:

$$\delta_{ij} = \begin{cases} 1 & i = j \\ 0 & i \neq j \end{cases}$$

It is clear that the matrix is symmetric. The matrix of substitution elasticities is also equal to:

$$\sigma_{ij} = [\mu_j - \delta_{ij}] \frac{\mu_i Y^*}{s_i s_j Y} = \frac{(x_i - \gamma_i)}{x_i} \frac{(x_j - \gamma_j)}{x_j} \frac{Y}{Y^*} - \frac{\delta_{ij}}{s_j} \frac{(x_i - \gamma_i)}{x_i}$$

The compensated demand elasticities derive from the following:

$$\xi_{ij} = s_j \sigma_{ij} = [\mu_j - \delta_{ij}] \frac{\mu_i Y^*}{s_i Y} \quad (\text{B-30})$$

Finally, the matrix of uncompensated demand elasticities is given by:

$$\varepsilon_{ij} = \xi_{ij} - s_j \eta_i = [\mu_j - \delta_{ij}] \frac{\mu_i Y^*}{s_i Y} - s_j \eta_i \quad (\text{B-31})$$

The uncompensated demand elasticities can also be written as:

$$\varepsilon_{ij} = -\frac{\mu_i}{s_i Y} [p_j \gamma_j + \delta_{ij} Y^*] + \frac{s_j}{s_i} (\beta_i - \alpha_i) \Omega \quad (\text{B-32})$$

The first term on the right-hand side is always negative. The second term differs from the LES expression for the uncompensated demand elasticities.<sup>4</sup> We can see from expression (B-32) that the AIDADS specification allows for both gross complementarity and substitution. As well, it allows for luxury goods, i.e. positive own-price demand elasticities should the second term be positive and greater than the first term.

### B.2.3 Implementation

Implementation of AIDADS is somewhat more complicated than the LES since the marginal propensity to consume out of supernumerary income is endogenous, and utility is defined implicitly. The following four equations are needed for model implementation:

$$Y^* = Y - \sum_i p_i \gamma_i \quad (\text{B-33})$$

$$x_i = \gamma_i + \frac{\mu_i}{p_i} Y^* \quad (\text{B-34})$$

$$\mu_i = \frac{\alpha_i + \beta_i e^u}{1 + e^u} \quad (\text{B-35})$$

$$u = \sum_i \mu_i \ln (x_i - \gamma_i) - 1 - \ln (A) \quad (\text{B-36})$$

<sup>4</sup> Recall that for the LES, the  $\alpha$  and  $\beta$  terms are equal and thus the second term drops.

Equations (B-33) and (B-34) are identical to their LES and ELES counterparts.<sup>5</sup> Equation (B-35) determines the level of the marginal propensity to consume out of supernumerary income,  $\mu$ , which is a constant in the case of the LES (ELES). It requires however the calculation of the utility level,  $u$ , which is defined in equation (B-36).

#### B.2.4 Calibration

[To be updated] Calibration requires more information than the LES. Where the LES has  $2n$  parameters to calibrate (subject to consistency constraints), AIDADS has  $3n$  parameters (less the consistency requirements)— $\alpha$ ,  $\beta$  and  $\gamma$ . The calibration system includes equations (B-33) through (B-36) which have  $2 + 2n$  endogenous variables ( $Y^*$ ,  $\gamma$ ,  $\mu$ , and  $A$ ). There are no equations for calibrating the  $\alpha$  and  $\beta$  parameters. If we have knowledge of the income elasticities, we can add the following equations:

$$\Psi = \frac{1}{\Omega} = \left[ \sum_i (\beta_i - \alpha_i) \ln(x_i - \gamma_i) - \frac{(1 + e^u)^2}{e^u} \right] \quad (\text{B-37})$$

$$\eta_i = \frac{\rho_i}{s_i} = \frac{\mu_i - (\beta_i - \alpha_i)\Omega}{s_i} = \frac{\mu_i}{s_i} - \frac{(\beta_i - \alpha_i)}{s_i \Psi} \quad (\text{B-38})$$

There are an additional  $1 + n$  equations, solving for  $\Psi$  and  $\alpha$ . There is need for an additional  $n$  equations. Assuming we have knowledge of at least  $n$  price elasticities, for example the own-price elasticities, we can add the following equation:

$$\varepsilon_{ii} = -\frac{\mu_i}{s_i Y} [p_i \gamma_i + Y^*] + (\beta_i - \alpha_i) \Omega \quad (\text{B-39})$$

The  $\alpha$  and  $\beta$  parameters are not independent, the following restrictions must hold:

$$\sum_i \alpha_i = 1 \quad (\text{B-40})$$

$$\sum_i \beta_i = 1 \quad (\text{B-41})$$

The system is under-determined, there are  $5 + 4n$  equations and  $3 + 4n$  variables. One solution, is to make the own-price elasticities endogenous. In this case, we are adding  $n$  variables, but then the system is over-determined. We can minimize a loss function with respect to the price elasticities:

$$L = \sum_i (\varepsilon_i - \varepsilon_i^0)^2$$

where  $\varepsilon^0$  represents an initial guess of the own-price elasticities and the calibration algorithm will calculate the endogenous  $\varepsilon$  in order to minimize the loss function, subject to constraints (B-37) through (B-41) and the model equations (B-33) through (B-36). The exogenous parameters in the calibration procedure include  $p$ ,  $x$ ,  $s$ ,  $Y$ ,  $\eta$ ,  $\varepsilon^0$  and  $u$ .

<sup>5</sup> Though the definition of  $Y$  includes savings in the case of the ELES.

### B.3 Constant differences in elasticity (CDE) demand system

The *Constant Difference of Elasticities* (CDE) function is a generalization of the CES function, but it allows for more flexibility in terms of substitution effects across goods and for non-homotheticity.<sup>6</sup> The starting point is an implicitly additive indirect utility function (see [Hanoch \(1975\)](#)) from which we can derive demand using Roy's identity (and the implicit function theorem).

#### B.3.1 General form

A dual approach is used to determine the properties of the CDE function. The indirect utility function is defined implicitly via the following expression:

$$V(p, u, Y) = \sum_{i=1}^n \alpha_i u^{e_i b_i} \left( \frac{p_i}{y} \right)^{b_i} \equiv 1 \quad (\text{B-42})$$

where  $p$  is the vector of commodity prices,  $u$  is (per capita) utility and  $y$  is per capita income. Using Roy's identity and the implicit function theorem<sup>7</sup> we can derive demand,  $x$ , where  $v$  is the indirect utility function (defined implicitly):

$$x_i = -\frac{\partial v}{\partial p_i} / \frac{\partial v}{\partial Y} = -\left( \frac{\partial V}{\partial p_i} / \frac{\partial V}{\partial u} \right) / \left( \frac{\partial V}{\partial Y} / \frac{\partial V}{\partial u} \right) = -\left( \frac{\partial V}{\partial p_i} / \frac{\partial V}{\partial Y} \right) \quad (\text{B-43})$$

This then leads to the following demand function:

$$x_i = \frac{\alpha_i b_i u^{e_i b_i} \left( \frac{p_i}{y} \right)^{b_i - 1}}{\sum_j \alpha_j b_j u^{e_j b_j} \left( \frac{p_j}{y} \right)^{b_j}} \quad (\text{B-44})$$

Implementation is easier if we define the following variable:

$$ZC_i = \alpha_i b_i u^{e_i b_i} \left( \frac{p_i}{y} \right)^{b_i} \quad (\text{B-45})$$

Then the budget shares can be expressed as:

$$s_i = \frac{ZC_i}{\sum_j ZC_j} \quad (\text{B-46})$$

and the demand expression is:

$$x_i = \frac{s_i}{p_i} y \quad (\text{B-47})$$

Implementation also requires evaluating  $u$ . This can be done by implementing equation (B-42) and inserting the expression for  $ZC$ :

$$\sum_{i=1}^n \frac{ZC_i}{b_i} \equiv 1 \quad (\text{B-48})$$

<sup>6</sup> More detailed descriptions of the CDE can be found in [Hertel et al. \(1991\)](#), [Surry \(1993\)](#) and [Hertel \(1997\)](#).

<sup>7</sup> See [Varian \(1992\)](#), p. 109.

### B.3.2 Elasticities

In order to calibrate the CDE system, it is necessary to derive the demand and income elasticities of the CDE. The algebra is tedious, but straightforward.

The own-price elasticity is given by the following:

$$\varepsilon_i = \frac{\partial x_i}{\partial p_i} \frac{p_i}{x_i} = \frac{s_i \left[ \sum_j s_j e_j b_j - e_i b_i \right]}{\sum_j s_j b_j} + b_i (1 - s_i) - 1 \quad (\text{B-49})$$

In deriving the elasticity, we make use of the following formula that defines the elasticity of utility with respect to price (and again makes use of the implicit function theorem):

$$\frac{\partial u}{\partial p_i} \frac{p_i}{u} = -\frac{p_i}{u} \left( \frac{\partial V}{\partial p_i} \right) / \left( \frac{\partial V}{\partial u} \right) = -\frac{s_i}{\sum_j s_j e_j} \quad (\text{B-50})$$

The price elasticity of utility is approximately the value share of the respective demand component as long as the weighted sum of the expansion parameters,  $e$ , is close to unity. The value (or budget) share is defined in the next equation:

$$s_i = \frac{p_i x_i}{y} \quad (\text{B-51})$$

Letting  $\sigma_i = 1 - b_i$  (or  $b_i = 1 - \sigma_i$ ), we can also write:

$$\varepsilon_i = s_i \left[ \sigma_i - \frac{e_i(1 - \sigma_i)}{\sum_j s_j e_j} - \frac{\sum_j s_j e_j \sigma_j}{\sum_j s_j e_j} \right] - \sigma_i \quad (\text{B-52})$$

With  $\sigma$  uniform, we also have:

$$\varepsilon_i = -\frac{s_i e_i (1 - \sigma)}{\sum_j s_j e_j} - \sigma \quad (\text{B-53})$$

With both  $e$  and  $\sigma$  uniform, the formula simplifies to:

$$\varepsilon_i = -s_i(1 - \sigma) - \sigma = \sigma(s_i - 1) - s_i \quad (\text{B-54})$$

Equation (B-54) reflects the own-price elasticity for the standard CES utility function. Finally, with  $e$  uniform but not  $\sigma$ , we have:

$$\varepsilon_i = s_i \left[ 2\sigma_i - 1 - \sum_j s_j \sigma_j \right] - \sigma_i \quad (\text{B-55})$$

The derivation of the cross elasticities is almost identical and will not be carried out here. Combining both the own-and cross price elasticities, the matrix of substitution elasticities takes

the following form where we use the Kronecker product,  $\delta$ :<sup>8</sup>

$$\varepsilon_{ij} = s_j \left[ -b_j - \frac{e_i b_i}{\sum_k s_k e_k} + \frac{\sum_k s_k e_k b_k}{\sum_k s_k e_k} \right] + \delta_{ij}(b_i - 1) \quad (\text{B-56})$$

Again, we replace  $b$  by  $1 - \sigma$ , to get:

$$\varepsilon_{ij} = s_j \left[ \sigma_j - \frac{e_i(1 - \sigma_i)}{\sum_k s_k e_k} - \frac{\sum_k s_k e_k \sigma_k}{\sum_k s_k e_k} \right] - \delta_{ij} \sigma_i \quad (\text{B-57})$$

For uniform  $\sigma$ , equation (B-57) takes the form:

$$\varepsilon_{ij} = -\frac{e_i s_j (1 - \sigma)}{\sum_k s_k e_k} - \delta_{ij} \sigma \quad (\text{B-58})$$

And with a uniform  $\sigma$  and  $e$ , i.e. the CES assumption, we have:

$$\varepsilon_{ij} = -s_j(1 - \sigma) - \delta_{ij} \sigma = \sigma(s_j - \delta_{ij}) - s_j \quad (\text{B-59})$$

Finally, for a uniform  $e$  only, the matrix of elasticities is:

$$\varepsilon_{ij} = s_j \left[ \sigma_j - (1 - \sigma_i) - \sum_k s_k \sigma_k \right] - \delta_{ij} \sigma_i \quad (\text{B-60})$$

The income elasticities are derived in a similar fashion:

$$\eta_i = \frac{\partial x_i}{\partial Y} \frac{Y}{x_i} = \frac{1}{\sum_k s_k e_k} \left[ e_i b_i - \sum_k s_k e_k b_k \right] - (b_i - 1) + \sum_k b_k s_k \quad (\text{B-61})$$

For this, we need the elasticity of utility with respect to income:

$$\frac{\partial u}{\partial Y} \frac{Y}{u} = -\frac{Y}{u} \left( \frac{\partial V}{\partial Y} \right) / \left( \frac{\partial V}{\partial u} \right) = \frac{1}{\sum_k s_k e_k} \quad (\text{B-62})$$

Note that for a uniform and unitary  $e$ , the income elasticity of utility is 1.

Replacing  $b$  with  $1 - \sigma$ , equation (B-61) can be re-written to be:

$$\eta_i = \frac{1}{\sum_k s_k e_k} \left[ e_i(1 - \sigma_i) + \sum_k s_k e_k \sigma_k \right] + \sigma_i - \sum_k s_k \sigma_k \quad (\text{B-63})$$

With a uniform  $\sigma$ , the income elasticity becomes:

$$\eta_i = \frac{1}{\sum_k s_k e_k} \left[ e_i(1 - \sigma) + \sigma \sum_k s_k e_k \right] = \frac{e_i(1 - \sigma)}{\sum_k s_k e_k} + \sigma \quad (\text{B-64})$$

With  $e$  uniform, the income elasticity is unitary, irrespective of the values of the  $\sigma$  parameters.

<sup>8</sup>  $\delta$  takes the value of 1 along the diagonal (i.e. when  $i = j$ ) and the value 0 off-diagonal (i.e. when  $i \neq j$ ).

From the Slutsky equation, we can calculate the compensated demand elasticities:

$$\xi_{ij} = \varepsilon_{ij} + s_j \eta_i = -\delta_{ij} \sigma_i + s_j \left[ \sigma_j + \sigma_i - \sum_k s_k \sigma_k \right] \quad (\text{B-65})$$

The cross-Allen partial elasticities are equal to the compensated demand elasticities divided by the share:

$$\sigma_{ij}^a = \sigma_j + \sigma_i - \sum_k s_k \sigma_k - \delta_{ij} \sigma_i / s_j \quad (\text{B-66})$$

It can be readily seen that the difference of the partial elasticities is constant, hence the name of *constant difference in elasticities*.

$$\sigma_{ij}^a - \sigma_{il}^a = \sigma_j - \sigma_l \quad (\text{B-67})$$

With a uniform  $\sigma$ , we revert back to the standard CES where there is equivalence between the CES substitution elasticity and the cross-Allen partial elasticity:

$$\sigma_{ij}^a = \sigma \quad (\text{B-68})$$

### B.3.3 Calibration of the CDE

Calibration assumes that we know the budget shares, the own uncompensated demand elasticities and the income elasticities. The weighted sum of the income elasticities must equal 1, so the first step in the calibration procedure is to make sure Engel's law holds. One alternative is to fix some (or none) of the income elasticities and re-scale the others using least squares. The problem is to minimize the following objective function:

$$\sum_{i \in \Omega} (\eta_i - \eta_i^0)^2$$

subject to

$$\sum_{i \in \Omega} s_i \eta_i = 1 - \sum_{i \notin \Omega} s_i \eta_i$$

where the set  $\Omega$  contains all sectors where the income elasticity is not fixed, i.e. its complement contains those sectors with fixed income elasticities. The solution is:

$$\eta_i = \eta_i^0 + s_i \frac{1 - \sum_{j \notin \Omega} s_j \eta_j - \sum_{j \in \Omega} s_j \eta_j^0}{\sum_{j \in \Omega} s_j^2} \quad \forall i \in \Omega$$

Calibration of the  $\sigma$  parameters is straightforward given the own elasticities and the input value shares. The first step is to calculate the Allen partial elasticities, these are simply the income elasticity adjusted by the own elasticities divided by the budget shares:

$$\sigma_{ii}^a = \eta_i + \frac{\varepsilon_{ii}}{s_i} \quad (\text{B-69})$$

Next, equation (B-66) is setup in matrix form:

$$\sigma_{ii}^a = A\sigma_i \quad (\text{B-70})$$

where the matrix  $A$  has the form:

$$A = \begin{bmatrix} 2 - \frac{1}{s_1} - s_1 & -s_2 & \dots & -s_n \\ -s_1 & 2 - \frac{1}{s_2} - s_2 & \dots & -s_n \\ \vdots & \vdots & \ddots & \vdots \\ -s_1 & -s_2 & \dots & 2 - \frac{1}{s_n} - s_n \end{bmatrix} \quad (\text{B-71})$$

or each element of  $A$  has the following formula:

$$a_{ij} = \delta_{ij}(2 - 1/s_i) - s_j$$

We can then solve for  $\sigma$  (and back-out the  $b$  parameters):

$$\sigma_i = A^{-1}\sigma_{ii}^a \quad (\text{B-72})$$

There is nothing which guarantees the consistency of the calibrated  $\sigma$  parameters, which are meant to be positive. The calculation of the  $\sigma$  parameters depends only on the budget shares and the own-price uncompensated elasticities. If the calibrated  $\sigma$  parameters are not all positive, one could modify the elasticities until consistency is achieved. In practice, problems have occurred when a sector's budget share dominates total expenditure.

The  $e$  parameters are derived from Equation (B-63) and normalizing them so that their share weighted sum is equal to 1. Equation (B-63) can then be converted to matrix form and inverted:

$$B = \begin{bmatrix} s_1\sigma_1 + (1 - \sigma_1) & s_2\sigma_2 & \dots & s_n\sigma_n \\ s_1\sigma_1 & s_2\sigma_2 + (1 - \sigma_2) & \dots & s_n\sigma_n \\ \vdots & \vdots & \ddots & \vdots \\ s_1\sigma_1 & s_2\sigma_2 & \dots & s_n\sigma_n + (1 - \sigma_n) \end{bmatrix} \quad (\text{B-73})$$

or

$$b_{ij} = s_j\sigma_j + \delta_{ij}(1 - \sigma_i)$$

Then the  $e$  parameters are derived from matrix inversion:

$$e_i = B^{-1}C_i = B^{-1} \left( \eta_i - \sigma_i + \sum_k s_k\sigma_k \right) \quad (\text{B-74})$$

Calibration of the  $\alpha$  parameters is based on equations (B-42) and (B-44). Start first with equation (B-44) and write it in terms relative to consumption of good 1, i.e.:

$$\frac{x_i}{x_1} = \frac{\alpha_i b_i u^{e_i b_i} \left( \frac{p_i}{Y} \right)^{b_i - 1}}{\alpha_1 b_1 u^{e_1 b_1} \left( \frac{p_1}{Y} \right)^{b_1 - 1}} \quad (\text{B-75})$$



This equation can be used to isolate  $\alpha_i$ :

$$\alpha_i = \frac{x_i}{x_1} \frac{\alpha_1 b_1 u^{e_1 b_1} \left(\frac{p_1}{Y}\right)^{b_1-1}}{b_i u^{e_i b_i} \left(\frac{p_i}{Y}\right)^{b_i-1}} \quad (\text{B-76})$$

and then inserted into equation (B-44):

$$\sum_{i=1}^n \alpha_i u^{e_i b_i} \left(\frac{p_i}{Y}\right)^{b_i} = \alpha_1 u^{e_1 b_1} \frac{b_1}{s_1} \left(\frac{p_1}{Y}\right)^{b_1} \left[ \sum_{i=1}^n \frac{s_i}{b_i} \right] \equiv 1 \quad (\text{B-77})$$

The final expression in equation (B-77) can be used to solve for  $\alpha_1$  since the formula must equal 1 by definition:

$$\alpha_1 = u^{-e_1 b_1} \frac{s_1}{b_1} \left(\frac{Y}{p_1}\right)^{b_1} \left[ \sum_{i=1}^n \frac{s_i}{b_i} \right]^{-1} \quad (\text{B-78})$$

Substituting back into equation (B-77) we get:

$$\alpha_i = \frac{x_i}{b_i} u^{-e_i b_i} \left(\frac{Y}{p_i}\right)^{b_i-1} \left[ \sum_{j=1}^n \frac{s_j}{b_j} \right]^{-1} \quad (\text{B-79})$$

The final calibration expression is then the following:

$$\alpha_i = \frac{s_i}{b_i} \left(\frac{Y}{p_i}\right)^{b_i} \frac{u^{-e_i b_i}}{\sum_{j=1}^n \frac{s_j}{b_j}} \quad (\text{B-80})$$

Utility is undefined in the base data and it is easiest to simply set it to 1.

In conclusion, for calibration we need the budget shares, initial prices, total expenditure, income elasticities and the own-price uncompensated elasticities. From this, we can derive base year consumption volumes, the Allen partial substitution elasticities through equation (B-69),  $\sigma$  (and therefore  $b$ ) through equation (B-72) and the inversion of the  $A$ -matrix,  $e$  through equation (B-74) and inversion of the  $B$ -matrix, and finally  $\alpha$  through equation (B-80).

It is possible that the initial shares and elasticities lead to inconsistent calibrated values for the  $b$  or  $e$  parameters. One solution, modified from Hertel (1997), is to implement some sort of maximum entropy method—explicitly imposing the constraints on the parameters. Step 1 is to calibrate the  $b$ -parameters using the following minimization problem:

$$\min L = \sum_i s_i (\varepsilon_{ii} - \varepsilon_{ii}^0)^2$$

subject to

$$\varepsilon_{ii} = (1 - b_i) (s_i - 1) - s_i \left[ b_i + \eta_i - \sum_j s_j b_j \right]$$

$$0 < b_i < 1$$

The loss function is a weighted sum of square errors where  $\varepsilon^0$  represents the initial or target own-price elasticity and  $\varepsilon$  will be the estimated elasticity with the constraints holding. The first constraint is a transformation of equation (B-49) where the income elasticity is substituted into the definition of the own-price elasticity (swapping out for the yet unknown  $e$ -coefficients). One critical issue is to ascertain what income elasticities to use in the formula above. One could use the target income elasticities, or an initial transformation of the target elasticities such as described above.

The next step calibrates the  $e$ -parameters with some target income elasticities as given as well as the now calibrated  $b$ -parameters. The minimization problem is formulated as the following:

$$\min L = \sum_i s_i (\eta_i - \eta_i^0)^2$$

subject to

$$\begin{aligned} \eta_i &= \frac{1}{\sum_k s_k e_k} \left[ e_i b_i - \sum_k s_k e_k b_k \right] - (b_i - 1) + \sum_k b_k s_k \\ \sum_i s_i \eta_i &\equiv 1 \\ (\eta_i - 1) (\eta_i^0 - 1) &> 0 \end{aligned}$$

The final constraint insures that the estimated income elasticities preserve their relationship relative to 1, i.e. target elasticities lower than 1 remain lower than 1 in the estimation procedure.

### B.3.4 CDE in first differences

It is useful to decompose changes in demand using a linearized version of the demand function, and that which is used in the standard GEMPACK version of the CDE function. The CDE implicit utility function can be used to derive a relation between changes in income, utility and prices (all in per capita terms). The first step in the differentiation of the utility function, equation (B-42), leads to the following expression:

$$\begin{aligned} 0 &= \sum_i \alpha_i e_i b_i u^{e_i b_i - 1} \left( \frac{p_i}{Y} \right)^{b_i} du \\ &\quad - \sum_i \alpha_i b_i u^{e_i b_i} \left( \frac{p_i}{Y} \right)^{b_i - 1} \frac{p_i}{Y^2} dY \\ &\quad + \sum_i \alpha_i b_i u^{e_i b_i} \left( \frac{p_i}{Y} \right)^{b_i - 1} \frac{1}{Y} dp_i \end{aligned}$$

This can be simplified by inserting the expression for the demand equation, equation (B-44), and replacing demand with the budget shares ( $s_i$ ):

$$0 = \frac{du}{u} \sum_i e_i s_i - \frac{dY}{Y} \sum_i s_i + \sum_i s_i \frac{dp_i}{p_i}$$

And the final expression can be written as:

$$\dot{Y} = \sum_i e_i s_i \dot{u} + \sum_i s_i \dot{p}_i \quad (\text{B-81})$$

where the dotted variables represent the percent change (and noting that the sum of the budget shares is equal to 1).

The differentiation of the demand function, equation (B-44) is somewhat more tedious. The first step leads to the following expression:

$$\begin{aligned}
 dx_i &= \alpha_i b_i e_i b_i u^{e_i b_i - 1} \left( \frac{p_i}{Y} \right)^{b_i - 1} \frac{du}{D} \\
 &+ \alpha_i b_i u^{e_i b_i} (b_i - 1) \left( \frac{p_i}{Y} \right)^{b_i - 2} \frac{1}{Y} \frac{dp_i}{D} \\
 &- \alpha_i b_i u^{e_i b_i} (b_i - 1) \left( \frac{p_i}{Y} \right)^{b_i - 2} \frac{p_i}{Y^2} \frac{dY}{D} \\
 &- \alpha_i b_i u^{e_i b_i} \left( \frac{p_i}{Y} \right)^{b_i - 1} D^{-2} \sum_j \alpha_j b_j e_j b_j u^{e_j b_j - 1} \left( \frac{p_j}{Y} \right)^{b_j} du \\
 &- \alpha_i b_i u^{e_i b_i} \left( \frac{p_i}{Y} \right)^{b_i - 1} D^{-2} \sum_j \alpha_j b_j b_j u^{e_j b_j} \left( \frac{p_j}{Y} \right)^{b_j - 1} \frac{1}{Y} dp_j \\
 &+ \alpha_i b_i u^{e_i b_i} \left( \frac{p_i}{Y} \right)^{b_i - 1} D^{-2} \sum_j \alpha_j b_j b_j u^{e_j b_j} \left( \frac{p_j}{Y} \right)^{b_j - 1} \frac{p_j}{Y^2} dY
 \end{aligned}$$

where  $D$  is the denominator in the demand equation. This can be simplified to the following expression in terms of the percent changes:

$$\begin{aligned}
 \dot{x}_i &= e_i b_i \dot{u} + (b_i - 1) \dot{p}_i - (b_i - 1) \dot{Y} \\
 &- \sum_j e_j b_j s_j \dot{u} - \sum_j b_j s_j \dot{p}_j + \sum_j b_j s_j \dot{Y}
 \end{aligned}$$

Re-grouping terms, the expression becomes:

$$\begin{aligned}
 \dot{x}_i &= (b_i - 1) \dot{p}_i - \sum_j b_j s_j \dot{p}_j \\
 &+ \dot{u} \left[ e_i b_i - \sum_j e_j b_j s_j \right] \\
 &+ \dot{Y} \left[ \sum_j b_j s_j - (b_i - 1) \right]
 \end{aligned}$$

The percent change in  $u$  can be replaced with the expression above, equation (B-81), to yield the following after re-arrangement:

$$\begin{aligned}
 \dot{x}_i &= (b_i - 1) \dot{p}_i - \sum_j b_j s_j \dot{p}_j - \frac{1}{\sum_k e_k s_k} \sum_j s_j \dot{p}_j \left[ e_i b_i - \sum_k e_k b_k s_k \right] \\
 &+ \dot{Y} \left[ \sum_k b_k s_k - (b_i - 1) + \frac{1}{\sum_k e_k s_k} \left( e_i b_i - \sum_k e_k b_k s_k \right) \right]
 \end{aligned}$$

The final formula inserts the formulas for the income and price elasticities from above to simplify further to the following expression:

$$\dot{x}_i = \sum_j \varepsilon_{ij} \dot{p}_j + \eta_i \dot{Y} \tag{B-82}$$

# Appendix C

## Analytics of dynamics

### C.1 Capital stock

When moving to flexible time steps, cumulative investment over intermediate years needs to be accounted for. One simple way to do this is to assume a constant growth rate for investment between a past year and the current year:

$$I_t = (1 + g_t^I)^n I_{t-n}$$

After some algebra, this leads to a motion equation for the capital stock that takes the following form:<sup>1</sup>

$$K_t^s = (1 - \delta_t)^n K_{t-n}^s + \frac{(1 + g_t^I)^n - (1 - \delta_t)^n}{g_t^I + \delta_t} I_{t-n}$$

The resulting motion equation is no longer pre-determined as the contemporaneous level of investment is needed to measure the annual growth in investment. The equation is also quite sensitive to the relative values of  $g^I$  and  $\delta$ . In the model implementation, a somewhat transformed version of the motion equation is used that includes an additional equation that captures investment growth. Equation (C-1) is the investment growth factor. The value of the expression is equal to the inverse of the annual rate of growth of investment plus the depreciation rate. For example, when both are at 5 percent,  $\Psi$  takes the value 10. Equation (C-2) is the capital stock accumulation function for multi-period time steps. Both formulas depend on the contemporaneous level of investment. Equation (G-6) defines the normalized capital stock. It is a fixed scalar of the non-normalized level where the scaling factor is initialized using base year data, i.e.  $\chi_r^k = TK_{r,0}^s / K_{r,0}^s$

$$\Psi_{r,t} = \left[ \left( \frac{XFD_{r,inv,t}}{XFD_{r,inv,t-n}} \right)^{1/n} - 1 + \delta_{r,t} \right]^{-1} \quad (C-1)$$

$$K_{r,t}^s = [K_{r,t-n}^s - \Psi_{r,t} XFD_{r,inv,t-n}] (1 - \delta_{r,t})^n + \Psi_{r,t} XFD_{r,inv,t} \quad (C-2)$$

At times it is useful to measure sectoral investment needs. For a single time step, this is relatively straightforward using the following expression:

$$I_{r,a,t-1} = \frac{1}{\chi_r^k} \left[ \sum_v K_{r,a,v,t}^v - (1 - \delta_{r,t}) \sum_v K_{r,a,v,t-1}^v \right]$$

---

<sup>1</sup> It is easy to see that the equation collapses to the one-period motion equation when  $n$  is equal to 1.

where it is assumed that the factor  $\chi^k$  applies uniformly across all activities and not just for the aggregate capital stock. The value of investment is equal to the volume times the price of investment given by the variable  $PFD$ . In a multi-time step simulation, the formula can be generalized to the following:

$$I_{r,a,t-1} = \frac{1}{\chi_r^k} \frac{\delta_{r,t}}{1 - (1 - \delta_{r,t})^n} \left[ \sum_v K_{r,a,v,t}^v - (1 - \delta_{r,t})^n \sum_v K_{r,a,v,t-n}^v \right]$$

under the assumption of equal investment in each year. Cumulative investment is then  $n$  times the annual level. The formula is derived from the following:

$$\sum_v K_{r,a,v,t}^v = (1 - \delta_{r,t})^n \sum_v K_{r,a,v,t-n}^v + \chi_r^k \sum_{i=0}^{n-1} (1 - \delta_{r,t})^{n-i-1} I_{r,a,t-n+i}$$

and under the assumption that investment is equal in each year.

## C.2 Knowledge stock

The starting point for the model implementation is the 1-step motion equation for the knowledge stock<sup>2</sup>:

$$KN_t = (1 - \delta) KN_{t-1} + \sum_{k=0, i=t}^{N, t-N} \gamma_k RD_i = (1 - \delta) KN_{t-1} + G' RDL_{t,t}$$

where the last term is the summation expression in vector form (the double-indexed subscript will become clearer below).  $G' RDL$  represents the gamma-weighted sum of the lags of R&D expenditures. From this, we can use induction and convert the one-step expression to an  $n$ -step expression, equation (C-3). The first term on the right-hand side is intuitive. The second term adds all of the relevant cumulative weighted-lags and includes the appropriate depreciation for the intermediate years. Say for example we are evaluating knowledge stock in year 2025 with respect to 2020. The sum will be over all years between 2021 and 2025 inclusive. Equation (C-4) calculates the cumulative weighted lags for each year between  $t - n + 1$  and  $t$ , e.g. between 2021 and 2025 for the example above. It does require, nonetheless, the level of R&D expenditures for the intermediate years. Since these have differing weights over time, we need to keep track of these explicitly. In the GAMS code, most equations are indexed over  $t$ , which has variable step sizes. The R&D variable is calculated for each year, assuming constant growth in the intermediate years. Equation (C-5) determines the multi-step growth rate. Equation (C-6) defines R&D expenditures for each year, irrespective of the model's time step.

$$KN_t = (1 - \delta)^n KN_{t-n} + \sum_{i=t-n+1}^t (1 - \delta)^{t-i} RDL_{i,t} \quad (C-3)$$

$$RDL_{i,t} = \sum_{k=0, j=i-k}^{N, i-N} \gamma_k RD_j \quad \text{for } t - n < i \leq t \quad (C-4)$$

$$RD_t = (1 + gr_t)^n RD_{t-n} \quad (C-5)$$

<sup>2</sup> In the model the variable  $RD$  is given by  $XFD_{rd}$ .

$$RD_i = (1 + gr_t)^{n-(t-i)} RD_{t-n} \quad \text{for } t - n < i < t \quad (\text{C-6})$$

In the model implementation, a macro is used to substitute equation (C-4) into equation (C-3). And equations (C-5) and (C-6) can be collapsed to equation (C-7) thereby eliminating the growth rate of R&D expenditures, where the calculations are done for each year between  $t - n$  and  $t$ .

$$RD_i = RD_{t-n}^{(t-i)/n} RD_t^{(n-t-i)/n} \quad \text{for } t - n < i < t \quad (\text{C-7})$$

If R&D is growing at a steady rate of  $g$ , we can then write:

$$\begin{aligned} KN_t &= (1 - \delta) KN_{t-1} + \sum_{k=0}^N \gamma_k RD_{t-N} (1 + g)^{N-k} \\ &= (1 - \delta) KN_{t-1} + RD_t \sum_{k=0}^N \frac{\gamma_k}{(1 + g)^k} \end{aligned} \quad (\text{C-8})$$

In steady state, the summation term is a constant and we can thus write:

$$KN_t = (1 - \delta) KN_{t-1} + \beta RD_t \quad \text{where } \beta = \sum_{k=0}^N \frac{\gamma_k}{(1 + g)^k}$$

The formula can also be converted to work with multi-year time steps.

$$\begin{aligned} KN_t &= (1 - \delta) KN_{t-1} + \beta RD_t \\ &= (1 - \delta) [(1 - \delta) KN_{t-2} + \beta RD_{t-1}] + \beta RD_t \\ &= (1 - \delta)^2 KN_{t-2} + (1 - \delta) \beta RD_{t-1} + \beta RD_t \\ &= (1 - \delta)^3 KN_{t-3} + (1 - \delta)^2 \beta RD_{t-2} + (1 - \delta) \beta RD_{t-1} + \beta RD_t \\ &\vdots \\ &= (1 - \delta)^n KN_{t-n} + \beta \sum_{i=0}^{n-1} (1 - \delta)^i RD_{t-i} \end{aligned}$$

Assuming a constant growth rate for R&D expenditures, the last expression becomes:

$$KN_t = (1 - \delta)^n KN_{t-n} + \beta \sum_{i=0}^{n-1} (1 - \delta)^i RD_t (1 + g)^{-i}$$

where

$$RD_t = RD_{t-i} (1 + g)^i$$

After a bit of algebra, it is straightforward to show that the summation expression simplifies to:

$$\sum_{i=0}^{n-1} \frac{(1 - \delta)^i}{(1 + g)^i} = \frac{1}{(1 + g)^{n-1}} \frac{(1 + g)^n - (1 - \delta)^n}{g + \delta}$$

The final expression can thus be written as:

$$\begin{aligned} KN_t &= (1 - \delta)^n KN_{t-n} + \beta \frac{RD_t}{(1 + g)^{n-1}} \frac{(1 + g)^n - (1 - \delta)^n}{g + \delta} \\ &= (1 - \delta)^n KN_{t-n} + \beta (1 + g) RD_{t-n} \frac{(1 + g)^n - (1 - \delta)^n}{g + \delta} \end{aligned}$$

These formulas rely on  $n > N$ , or that the growth rate is uniform back through period  $t - N$ .

From the last formula, we can show that in steady-state the following relation must hold for the ratio of knowledge stocks relative to R&D expenditures:

$$\frac{KN}{RD} = \beta \frac{1+g}{g+\delta}$$

## Implementation

Initializing the stock of knowledge and the pre-base year R&D expenditures is an issue. The strategy followed so far is based on the following steps:

- Assume a given base year level of R&D. Calculate the knowledge stock ( $KN$ ) expenditure consistent with some assumption about the steady-state growth rate around the base year and the rate of knowledge depreciation. From the expression above, we have:

$$KN_{t_0} = \beta RD_{t_0} \frac{1+g}{g+\delta} = RD_{t_0} \frac{1+g}{g+\delta} \sum_{k=0}^N \frac{\gamma_k}{(1+g)^k}$$

Alternatively, one could make an assumption about the base year knowledge stock and invert the expression to calculate the R&D expenditures consistent with the steady-state assumption.

- Back-cast the R&D expenditures to the earliest period, that should be at least  $N$  years back or more, for the largest  $N$ :

$$RD_t = RD_{t+1} / (1+g) \quad \text{for } t_f \leq t < t_0 - 1$$

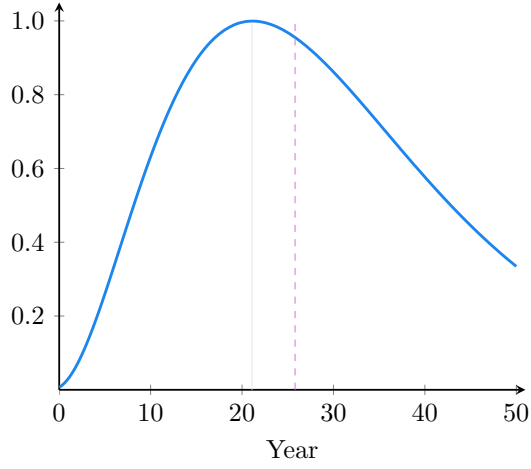
where  $t_f$  is the first year defined, e.g. 1960, and  $t_0$  is the base year for the simulations, e.g. 2014.

- Back-cast the cumulative lag structure,  $RDL$  using the following formula:

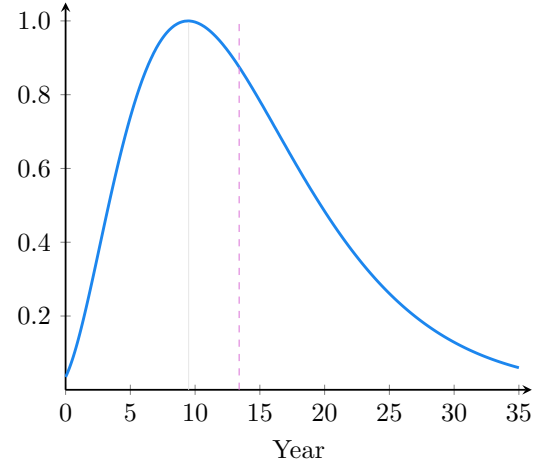
$$RDL_{t,t_0} = \sum_{k=0, i=t-k}^{N, t-N} \gamma_k RD_i \quad \text{for } t_f + N \leq t \leq t_0$$

- Back-cast the stock of knowledge. This step is not strictly necessary.

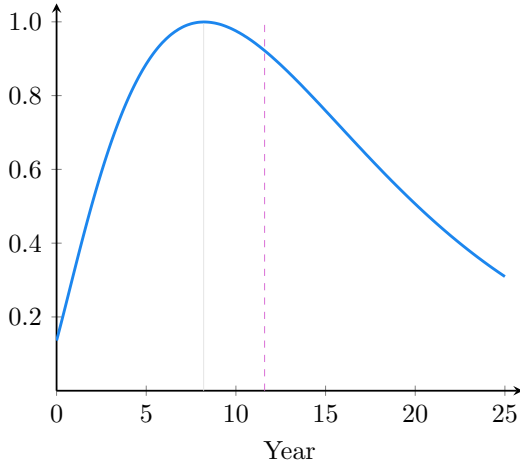
$$KN_t = \frac{KN_{t+1} - RDL_{t+1,t_0}}{1-\delta} \quad \text{for } t_f + N - 1 \leq t < t_0$$



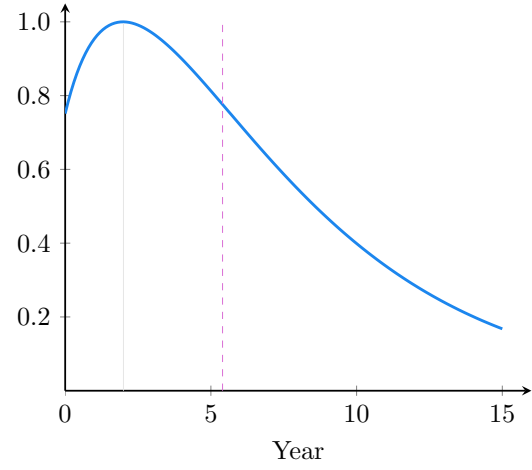
(a)  $\delta = 0.7$ ,  $\lambda = 0.9$ ,  $N = 50$ , **max=21.1**, **mean=25.8**



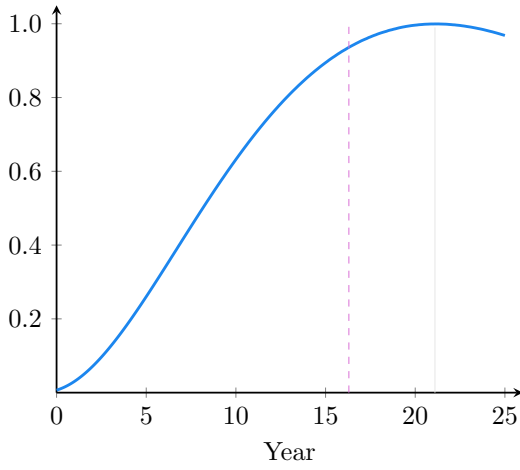
(b)  $\delta = 0.7$ ,  $\lambda = 0.8$ ,  $N = 35$ , **max=9.5**, **mean=13.4**



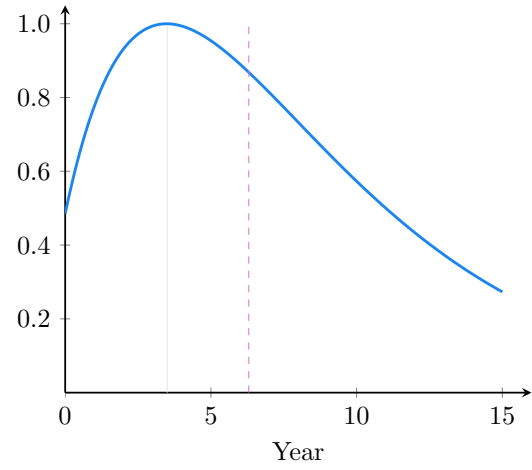
(c)  $\delta = 0.6$ ,  $\lambda = 0.85$ ,  $N = 25$ , **max=8.2**, **mean=11.6**



(d)  $\delta = 0.4$ ,  $\lambda = 0.8$ ,  $N = 15$ , **max=2.0**, **mean=5.4**



(e)  $\delta = 0.7$ ,  $\lambda = 0.9$ ,  $N = 25$ , **max=21.1**, **mean=16.3**



(f)  $\delta = 0.5$ ,  $\lambda = 0.8$ ,  $N = 15$ , **max=3.5**, **mean=6.3**

Figure 16: **Gamma distribution examples**



## Appendix D

# Analytical SAM and GTAP correspondence

This section describes the accounting framework used by the the model. Two different accounting frameworks will be described. The first describes the analytical SAM underlying the model, i.e. how the SAM can be constructed using the variables of the model. The second describes the links between the model variables and the initial database as provided by GTAP.<sup>1</sup> The SAM is not a standard SAM and does not represent the full functionality of the model nor the underlying database. For example, demand is specified at the Armington level and the sales tax are collapsed to a single row dimension.

### D.1 The analytical SAM

The analytical SAM is the accounting framework that is derived from the variables of the model. In the absence of a shock, the analytical SAM should reproduce the input SAM. Table D.1 re-produces the analytical SAM. Below are a few explanatory notes:

1. The regional index ( $r$ ) is dropped for most of the expressions except when needed, for example in the bilateral trade flows.
2. The SAM reflects the differentiation between activities and commodities. The intersection of the activity rows with the commodity columns represents the transformation of domestic production into commodities. In the standard version of the model, the matrix is diagonal and there is a one-to-one mapping between activities and commodities.
3. The trade rows under the commodity columns represents the bilateral imports for region  $r$ . These are priced at border, i.e. CIF, prices. The commodity rows under the trade columns represent bilateral exports from region  $r$  and are valued at border, i.e. FOB, prices. At the global level, the difference between the two matrices represents the aggregate value of the trade and transport margins.
4. Outbound remittances are calculated post-tax. Inbound remittances accrue to households.
5. In the standard GTAP database, most direct taxes are levied on factor remuneration. There is no balancing between private and public savings. Direct household taxation is calibrated

---

<sup>1</sup> This refers to the new standard database, not the 'classic' GTAP database.

as a balancing item to clear the government accounts under the assumption that the deficit is zero. This can readily be modified with supplemental data. Direct household taxation can be negative, reflecting that net transfers between the government and households is positive.

6. The BOP row under the trade columns represents aggregate imports from each region (at border prices). Similarly, the trade rows under the BOP column represents the aggregate exports of region  $r$  towards each region. The BOP accounts reflect the full balance of payments. For each region, the sum of exports (including the export of trade and transport services) less the sum of imports, adjusted by remittances and net capital income, equals the net flow of foreign savings. At the global level, the foreign savings cell should be zero.

Table D.1: Analytical SAM

	$ACT(a)$	$COMM(i)$	$LAB(l)$
<i>Activities(a)</i>		$P_{a,i} X_{a,i}$	
<i>Commodities(i)</i>	$\gamma_{i,a}^{eda} PAT_i XA_{i,a}$		
<i>Labor(l)</i>	$PF_{l,a} XF_{l,a}$		
<i>Capital(cap)</i>	$PF_{cap,a} XF_{cap,a}$		
<i>Land(lnd)</i>	$PF_{lnd,a} XF_{lnd,a}$		
<i>Nat. Res.(nrs)</i>	$PF_{nrs,a} XF_{nrs,a}$		
<i>Nat. Res.(wat)</i>	$PF_{wat,a} XF_{nrs,a}$		
<i>Indirect tax</i>	$\sum_i \tau_{i,a}^a \gamma_{i,a}^{eda} PAT_i XA_{i,a}$		
<i>Production tax</i>		$\sum_a \tau_{a,i}^p P_{a,i} X_{a,i}$	
<i>Factor taxes</i>	$\sum_f \tau_{f,a}^v PF_{l,a} XF_{f,a}$		
<i>Import tax</i>		$\sum_s \tau_{s,i,r}^m PWM_{s,i,r} XW_{s,i,r}^d$	
<i>Export tax</i>		$\sum_d \tau_{r,i,d}^e PE_{r,i,d} XW_{r,i,d}^s$	
<i>Direct tax</i>			$\sum_a \kappa_{l,a}^f PF_{l,a} XF_{l,a}$
<i>Household</i>			$\sum_a (1 - \kappa_{l,a}^f) PF_{l,a} XF_{l,a} - \sum_d Remit_{d,l,r}$
<i>Government</i>			
<i>Investment</i>			
<i>Depreciation</i>			
<i>Trade margins</i>			
<i>Trade(s)</i>		$PWM_{s,i,r} XW_{s,i,r}^d$	
<i>Balance of payments</i>			$\sum_d Remit_{d,l,r}$

Table D.1 Analytical SAM, ctd.

	<i>CAP</i>	<i>LND</i>	<i>NRS</i>	<i>WAT</i>
<i>Activities(a)</i>				
<i>Commodities(i)</i>				
<i>Labor(l)</i>				
<i>Capital(cap)</i>				
<i>Land(lnd)</i>				
<i>Nat. Res.(nrs)</i>				
<i>Nat. Res.(wat)</i>				
<i>Indirect tax</i>				
<i>Production tax</i>				
<i>Factor taxes</i>				
<i>Import tax</i>				
<i>Export tax</i>				
<i>Direct tax</i>	$\sum_a \kappa_{cap,a}^f PF_{cap,a} XF_{cap,a}$	$\sum_a \kappa_{lnd,a}^f PF_{lnd,a} XF_{lnd,a}$	$\sum_a \kappa_{nrs,a}^f PF_{nrs,a} XF_{nrs,a}$	$\sum_a \kappa_{wat,a}^f PF_{wat,a} XF_{wat,a}$
<i>Household</i>	$\sum_a (1 - \kappa_{cap,a}^f) PF_{cap,a} XF_{cap,a} - YQTF$	$\sum_a (1 - \kappa_{lnd,a}^f) PF_{lnd,a} XF_{lnd,a}$	$\sum_a (1 - \kappa_{nrs,a}^f) PF_{nrs,a} XF_{nrs,a}$	$\sum_a (1 - \kappa_{wat,a}^f) PF_{wat,a} XF_{wat,a}$
<i>Government</i>				
<i>Investment</i>				
<i>Depreciation</i>				
<i>Trade margins</i>				
<i>Trade(s)</i>				
<i>Balance of payments</i>	$YQTF_r$			

Table D.1 Analytical SAM, ctd.

	<i>ITX</i>	<i>PTX</i>	<i>VTX</i>	<i>MTX</i>	<i>ETX</i>	<i>DTX</i>
<i>Activities(a)</i>						
<i>Commodities(i)</i>						
<i>Labor(l)</i>						
<i>Capital(cap)</i>						
<i>Land(lnd)</i>						
<i>Nat. Res.(nrs)</i>						
<i>Nat. Res.(wat)</i>						
<i>Indirect tax</i>						
<i>Production tax</i>						
<i>Factor taxes</i>						
<i>Import tax</i>						
<i>Export tax</i>						
<i>Direct tax</i>						
<i>Household</i>						
<i>Government</i>	$YGOV_{itx}$	$YGOV_{ptx}$	$YGOV_{vtx}$	$YGOV_{mtx}$	$YGOV_{etx}$	$YGOV_{dtx}$
<i>Investment</i>						
<i>Depreciation</i>						
<i>Trade margins</i>						
<i>Trade(s)</i>						
<i>Balance of payments</i>						

Table D.1 Analytical SAM, ctd.

	<i>HHD</i>	<i>GOV</i>	<i>INV</i>	<i>DEPR</i>
<i>Activities(a)</i>				
<i>Commodities(i)</i>	$\gamma_{i,h}^{eda} PAT_i^a XA_{i,h}$	$\gamma_{i,gov}^{eda} PAT_i^a XA_{i,gov}$	$\gamma_{i,inv}^{eda} PA_i^a XA_{i,inv}$	
<i>Labor(l)</i>				
<i>Capital(cap)</i>				
<i>Land(lnd)</i>				
<i>Nat. Res.(nrs)</i>				
<i>Nat. Res.(wat)</i>				
<i>Indirect tax</i>	$\sum_i \tau_{i,h}^a \gamma_{i,h}^{eda} PAT_i XA_{i,h}$	$\sum_i \tau_{i,gov}^a \gamma_{i,gov}^{eda} PAT_i XA_{i,gov}$	$\sum_i \tau_{i,inv}^a \gamma_{i,inv}^{eda} PAT_i XA_{i,inv}$	
<i>Production tax</i>				
<i>Factor taxes</i>				
<i>Import tax</i>				
<i>Export tax</i>				
<i>Direct tax</i>	$\kappa^h YH$			
<i>Household</i>				
<i>Government</i>				
<i>Investment</i>	$S^h$	$S^g$		$DeprY$
<i>Depreciation</i>	$DeprY$			
<i>Trade margins</i>				
<i>Trade(s)</i>				
<i>Balance of payments</i>		$ODAOut$		

Table D.1 Analytical SAM, ctd.

	$TMG(fd)$	$TRD(d)$	$BOP$
<i>Activities(a)</i>			
<i>Commodities(i)</i>	$PDT_i XTT_i$	$\sum_d PWE_{r,i,d} XW_{r,i,d}$	
<i>Labor(l)</i>			
<i>Capital(cap)</i>			
<i>Land(lnd)</i>			
<i>Nat. Res.(nrs)</i>			
<i>Nat. Res.(wat)</i>			
<i>Indirect tax</i>			
<i>Production tax</i>			
<i>Factor taxes</i>			
<i>Import tax</i>			
<i>Export tax</i>			
<i>Direct tax</i>			
<i>Household</i>			$\sum_s Remit_{r,l,s} + YQHT_r$
<i>Government</i>			$ODAO_{out}$
<i>Investment</i>			$PW^{sav} S_r^f$
<i>Depreciation</i>			
<i>Trade margins</i>			$\sum_i PDT_{r,i} XTT_{r,i}$
<i>Trade(s)</i>			$\sum_d \sum_i PWE_{r,i,d} XW_{r,i,d}$
<i>Balance of payments</i>		$\sum_s \sum_i PWM_{s,i,r} XW_{s,i,r}$	

## D.2 Correspondence to GTAP database

Table [D.2](#) provides the correspondence between the variables in the GAMS model and the standard GTAP database that is used as the starting basis for the GTAP model. Note that The GTAP database reflects an Armington structure that is agent-specific (at the top nest). The model allows for either a national or agent-specific decomposition of domestic absorption across sources. The correspondence herein reflects national sourcing. Hence, the domestic/import split by agent in the GTAP database is aggregated to provide demand at the Armington level.



Table D.2: Correspondence between GAMS model variables and GTAP database

<i>GTAP</i>	<i>GAMS</i>	<i>Description</i>
VDFB+VIFB	$PAT_{r,i}XA_{r,i,a}$	Firms' (Armington) purchases at basic prices
VDFP+VIFP	$PA_{r,i,a}XA_{r,i,a}$	Firms' (Armington) purchases at purchaser prices
VDPB+VIPB	$PAT_{r,i}XA_{r,i,h}$	Household (Armington) purchases at basic prices
VDPP+VIIP	$PA_{r,i,h}XA_{r,i,h}$	Household (Armington) purchases at purchaser prices
VDGB+VIGB	$PAT_{r,i}XA_{r,i,gov}$	Government (Armington) purchases at basic prices
VDGP+VIGP	$PA_{r,i,gov}XA_{r,i,gov}$	Government (Armington) purchases at purchasers prices
VDIB+VIIB	$PAT_{r,i}XA_{r,i,inv}$	Investment (Armington) purchases at basic prices
VDIP+VIIP	$PA_{r,i,inv}XA_{r,i,inv}$	Investment (Armington) purchases at purchaser prices
EVFB(f)	$PF_{r,f,a}XF_{r,f,a}$	Firms' purchases of factors at basic prices
EVFP(f)	$PF_{r,f,a}^pXF_{r,f,a}$	Firms' purchases of factors at purchaser prices
VOS	$PX_{r,a}XP_{r,a}$	Value of production at producer cost
MAKB	$P_{r,a,i}X_{r,a,i}$	Value of make matrix at basic prices
MAKS	$PP_{r,a,i}X_{r,a,i}$	Value of make matrix at suppliers' prices
EVOS(f)	$(1 - \kappa_{r,f,a}^f)PF_{r,f,a}XF_{r,f,a}$	After tax factor remuneration
VKB	$K_r^s$	Initial capital stock
VDEP	$DeprY$	Value of depreciation
VIMB	$PDM_{s,i,r}XW_{s,i,r}^d$	Value of bilateral imports tariff inclusive
VCIF	$PWM_{s,i,r}XW_{s,i,r}^d$	Value of bilateral imports at border prices
VFOB	$PWE_{r,i,d}XW_{r,i,d}^s$	Value of bilateral exports at border prices
VXSB	$PE_{r,i,d}XW_{r,i,d}^s$	Value of bilateral exports at producer prices
VST	$PDT_{r,m}XTT_{r,m}$	Domestic exports of trade & transport services
VTWR	$PTMG_mXMG_{m,r,i,d}^m$	Value of margins from $r$ to $d$ for good $i$ using mode $m$
POP	$Pop_r$	Population—millions

## Appendix E

# Dimensions of the GTAP database, release 9.2

Table E.1: **Regional dimension of GTAP database**

1	AUS	<b>Australia</b>
2	NZL	<b>New Zealand</b>
3	XOC	<b>Rest of Oceania</b>
		American Samoa (asm), Cook Islands (cok), Fiji (fji), French Polynesia (pyf), Guam (gum), Kiribati (kir), Marshall Islands (mhl), Federated States of Micronesia (fsm), Nauru (nau), New Caledonia (ncl), Norfolk Island (nfl), Northern Mariana Islands (mnp), Niue (niu), Palau (plw), Papua New Guinea (png), Samoa (wsm), Solomon Islands (slb), Tokelau (tkl), Tonga (ton), Tuvalu (tuv), Vanuatu (vut), Wallis and Futura Islands (wlf)
4	CHN	<b>China</b>
5	HKG	<b>Hong Kong (China)</b>
6	JPN	<b>Japan</b>
7	KOR	<b>Republic of Korea</b>
8	MNG	<b>Mongolia</b>
9	TWN	<b>Taiwan (China)</b>
10	XEA	<b>Rest of East Asia</b>
		Macao (mac), North Korea (prk)
12	KHM	<b>Cambodia</b>
13	IDN	<b>Indonesia</b>
13	IDN	<b>Indonesia</b>
14	LAO	<b>Lao, PDR</b>
15	MYS	<b>Malaysia</b>
16	PHL	<b>Philippines</b>
17	SGP	<b>Singapore</b>
18	THA	<b>Thailand</b>
19	VNM	<b>Vietnam</b>
20	XSE	<b>Rest of Southeast Asia</b>
		Myanmar (mmr), Timor-Leste (tmp)

Table E.0: Regional dimension of GTAP database (cont.)

21	BGD	<b>Bangladesh</b>
22	IND	<b>India</b>
23	LKA	<b>Sri Lanka</b>
24	NPL	<b>Nepal</b>
25	PAK	<b>Pakistan</b>
26	XSA	<b>Rest of South Asia</b>
		Afghanistan (afg), Bhutan (btn), Maldives (mdv)
27	CAN	<b>Canada</b>
28	USA	<b>United States</b>
29	MEX	<b>Mexico</b>
30	XNA	<b>Rest of North America</b>
		Bermuda (bmu), Greenland (grl), Saint Pierre & Miquelon (spm)
31	ARG	<b>Argentina</b>
32	BOL	<b>Bolivia</b>
33	BRA	<b>Brazil</b>
34	CHL	<b>Chile</b>
35	COL	<b>Colombia</b>
36	ECU	<b>Ecuador</b>
37	PRY	<b>Paraguay</b>
38	PER	<b>Peru</b>
39	URY	<b>Uruguay</b>
40	VEN	<b>Venezuela, Republica Bolivariana de</b>
41	XSM	<b>Rest of South America</b>
		Falkland Islands (flk), French Guiana (guf), Guyana (guy), Suriname (sur)
42	CRI	<b>Costa Rica</b>
43	GTM	<b>Guatemala</b>
44	HND	<b>Honduras</b>
45	NIC	<b>Nicaragua</b>
46	PAN	<b>Panama</b>
47	SLV	<b>El Salvador</b>
48	XCA	<b>Rest of Central America</b>
		Belize (blz)
49	DOM	<b>Dominican Republic</b>
50	JAM	<b>Jamaica</b>
51	PRI	<b>Puerto Rico</b>
52	TTO	<b>Trinidad &amp; Tobago</b>
53	XCB	<b>Caribbean</b>
		Anguilla (aia), Antigua & Barbuda (atg), Aruba (abw), Bahamas (bhs), Barbados (brb), Cayman Islands (cym), Cuba (cub), Dominica (dma), Grenada (grd), Guadeloupe (glp), Haiti (hti), Martinique (mtq), Montserrat (msr), Netherlands Antilles (ant), Saint Kitts & Nevis (kna), Saint Lucia (lca), Saint Vincent & the Grenadines (vct), Turks and Caicos Islands (tca), British Virgin Islands (vgb), United States Virgin Islands (vir)

Table E.-1: Regional dimension of GTAP database (cont.)

54	AUT	<b>Austria</b>
55	BEL	<b>Belgium</b>
56	BGR	<b>Bulgaria</b>
57	CYP	<b>Cyprus</b>
58	CZE	<b>Czech Republic</b>
59	DNK	<b>Denmark</b>
60	EST	<b>Estonia</b>
61	FIN	<b>Finland</b>
62	FRA	<b>France</b>
63	DEU	<b>Germany</b>
64	GRC	<b>Greece</b>
65	HUN	<b>Hungary</b>
66	IRL	<b>Ireland</b>
67	ITA	<b>Italy</b>
68	LVA	<b>Latvia</b>
69	LTU	<b>Lithuania</b>
70	LUX	<b>Luxembourg</b>
71	MLT	<b>Malta</b>
72	NLD	<b>Netherlands</b>
73	POL	<b>Poland</b>
74	PRT	<b>Portugal</b>
75	ROU	<b>Romania</b>
76	SVK	<b>Slovakia</b>
77	SVN	<b>Slovenia</b>
78	ESP	<b>Spain</b>
79	SWE	<b>Sweden</b>
80	GBR	<b>United Kingdom</b>
81	NOR	<b>Norway</b>
82	CHE	<b>Switzerland</b>
83	XEF	<b>Rest of European Free Trade Area (EFTA)</b>
		Iceland (isl), Liechtenstein (lie)
84	ALB	<b>Albania</b>
85	BLR	<b>Belarus</b>
86	HRV	<b>Croatia</b>
87	RUS	<b>Russian Federation</b>
88	UKR	<b>Ukraine</b>
89	XEE	<b>Rest of Eastern Europe</b>
		Moldova (mda)
90	XER	<b>Rest of Europe</b>
		Andorra (and), Bosnia and Herzegovina (bih), Faroe Islands (fro), Gibraltar (gib), Macedonia (mkd, former Yugoslav Republic of), Monaco (mco), San Marino (smr), Serbia and Montenegro (scg)

Table E.-2: Regional dimension of GTAP database (cont.)

91	KAZ	<b>Kazakhstan</b>
92	KGZ	<b>Kyrgyz Republic</b>
93	TJK	<b>Tajikistan</b>
94	XSU	<b>Rest of Former Soviet Union</b>
		Turkmenistan (tkm), Uzbekistan (uzb)
95	ARM	<b>Armenia</b>
96	AZE	<b>Azerbaijan</b>
97	GEO	<b>Georgia</b>
98	BHR	<b>Bahrain</b>
99	IRN	<b>Iran</b>
100	ISR	<b>Israel</b>
101	JOR	<b>Jordan</b>
102	KWT	<b>Kuwait</b>
103	OMN	<b>Oman</b>
104	QAT	<b>Qatar</b>
105	SAU	<b>Saudi Arabia</b>
106	TUR	<b>Turkey</b>
107	ARE	<b>United Arab Emirates</b>
108	XWS	<b>Rest of Western Asia</b>
		Iraq (irq), Lebanon (lbn), West Bank and Gaza (pse), Syrian Arab Republic (syr), Republic of Yemen (yem)
109	EGY	<b>Egypt</b>
110	MAR	<b>Morocco</b>
111	TUN	<b>Tunisia</b>
112	XNF	<b>Rest of North Africa</b>
		Algeria (dza), Libya (lby)

Table E.-3: Regional dimension of GTAP database (cont.)

113	BEN	<b>Benin</b>
114	BFA	<b>Burkina Faso</b>
115	CMR	<b>Cameroon</b>
116	CIV	<b>Côte d'Ivoire</b>
117	GHA	<b>Ghana</b>
118	GIN	<b>Guinea</b>
119	NGA	<b>Nigeria</b>
120	SEN	<b>Senegal</b>
121	TGO	<b>Togo</b>
122	XWF	<b>Rest of Western Africa</b>
		Cape Verde (cpv), Gambia, The (gmb), Guinea-Bissau (gnb), Liberia (lbr), Mali (mli), Mauritania (mrt), Niger (ner), Saint Helena (shn), Sierra Leone (sle)
123	XCF	<b>Central Africa</b>
		Central African Republic (caf), Chad (tcd), Congo (cog), Equatorial Guinea (gnq), Gabon (gab), Sao Tome & Principe (stp)
124	XAC	<b>South-Central Africa</b>
		Angola (ago), Democratic Republic of the Congo (cod)
125	ETH	<b>Ethiopia</b>
126	KEN	<b>Kenya</b>
127	MDG	<b>Madagascar</b>
128	MWI	<b>Malawi</b>
129	MUS	<b>Mauritius</b>
130	MOZ	<b>Mozambique</b>
131	RWA	<b>Rwanda</b>
132	TZA	<b>Tanzania</b>
133	UGA	<b>Uganda</b>
134	ZMB	<b>Zambia</b>
135	ZWE	<b>Zimbabwe</b>
136	XEC	<b>Rest of Eastern Africa</b>
		Burundi (bdi), Comoros (com), Djibouti (dji), Eritrea (eri), Mayotte (myt), Réunion (reu), Seychelles Islands (syc), Somalia (som), Sudan (sdn)
137	BWA	<b>Botswana</b>
138	NAM	<b>Namibia</b>
139	ZAF	<b>South Africa</b>
140	XSC	<b>Rest of South African Customs Union</b>
		Lesotho (lso), Swaziland (swz)
141	XTW	<b>Rest of the World</b>
		Antarctica (ata), Bouvet Island (bvt), British Indian Ocean Territory (iot), French Southern Territories (atf)

Table E.-2: Commodity dimension of GTAP database

1	PDR	Paddy rice
2	WHT	Wheat
3	GRO	Cereal grains, n.e.s.
4	V.F	Vegetables and fruits
5	OSD	Oil seeds
6	C.B	Sugar cane and sugar beet
7	PFB	Plant-based fibers
8	OCR	Crops, n.e.s.
9	CTL	Bovine cattle, sheep and goats, horses
10	OAP	Animal products n.e.s.
11	RMK	Raw milk
12	WOL	Wool, silk-worm cocoons
13	FRS	Forestry
14	FSH	Fishing
15	COA	Coal
16	OIL	Oil
17	GAS	Gas
18	OMN	Minerals n.e.s.
19	CMT	Bovine cattle, sheep and goat, horse meat products
20	OMT	Meat products n.e.s.
21	VOL	Vegetable oils and fats
22	MIL	Dairy products
23	PCR	Processed rice
24	SGR	Sugar
25	OFD	Food products n.e.s.
26	B.T	Beverages and tobacco products
27	TEX	Textiles
28	WAP	Wearing apparel
29	LEA	Leather products
30	LUM	Wood products
31	PPP	Paper products, publishing
32	P.C	Petroleum, coal products
33	CRP	Chemical, rubber, plastic products
34	NMM	Mineral products n.e.s.
35	LS	Ferrous metals
36	NFM	Metals n.e.s.
37	FMP	Metal products
38	MVH	Motor vehicles and parts
39	OTN	Transport equipment n.e.s.
40	ELE	Electronic equipment

Table E.-3: Commodity dimension of GTAP database (cont.)

41	OME	<b>Machinery and equipment n.e.s.</b>
42	OMF	<b>Manufactures n.e.s.</b>
43	ELY	<b>Electricity</b>
44	GDT	<b>Gas manufacture, distribution</b>
45	WTR	<b>Water</b>
46	CNS	<b>Construction</b>
47	TRD	<b>Trade</b>
48	OTP	<b>Transport n.e.s.</b>
49	WTP	<b>Sea transport</b>
50	ATP	<b>Air transport</b>
51	CMN	<b>Communication</b>
52	OFI	<b>Financial services n.e.s.</b>
53	ISR	<b>Insurance</b>
54	OBS	<b>Business services n.e.s.</b>
55	ROS	<b>Recreation and other services</b>
56	OSG	<b>Public administration and defense, education, health services</b>
57	DWE	<b>Dwellings</b>



# Appendix F

## User Guide

### F.1 Data preparation facility

#### F.1.1 Introduction

The data facility is composed of three modules.

- The first is the aggregation facility. It reads a user-prepared map file that contains the aggregation mappings. For example, a user may want to aggregate from GTAP's 140 regions and 57 sectors to some 25 regions and 10 sectors. The aggregation facility aggregates all relevant GTAP data files, the projection file for dynamic scenarios and the elasticities.
- The second module is called the 'filter' module. First developed by Tom Rutherford for his 'GTAPinGAMS' project<sup>1</sup> and integrated into Wolfgang Britz's CGEBox framework<sup>2</sup>, filtering is designed to remove small numeric values from the database that have minimal impact on data balances and model results. The 'Filter' program insures that the resulting database is nonetheless balanced.
- The third module is called the 'AlterTax' module. It is described in [Malcolm \(1998\)](#) and is a procedure to modify an initial GTAP database while minimizing changes to the structure of the original database. Though designed to change tax rates, it can be used to modify other features of a database. This module relies on the standard GTAP model itself, which is provided as part of the data preparation facility.

All three modules are designed to work with the 'geometry' of the standard GTAP database (albeit with the aggregated dimensions). In other words, the output of each of the three modules is identical in structure to the standard GTAP database. Users need to provide two mappings. One mapping is necessary for the data preparation facility, where the geometry is preserved. A second mapping is needed for the model aggregation where it is possible to move from the implicit diagonal make matrix of the standard GTAP geometry to a non-diagonal make matrix for the model. This allows for activities that can produce more than one commodity—say for instance all crop production is collapsed into a single activity, but crop demand is disaggregated. And it allows for a single commodity to be produced by more than one activity—for example electricity.

The outputs are contained in different containers. The output from the aggregation facility is found in the 'Agg' folder. The output of the filtering procedure is found in the 'Flt' folder.

---

<sup>1</sup> [Lanz and Rutherford \(2016\)](#)

<sup>2</sup> [http://www.ilr.uni-bonn.de/em/rsrch/cgebox/cgebox\\_GUI.pdf](http://www.ilr.uni-bonn.de/em/rsrch/cgebox/cgebox_GUI.pdf)

And the output of AlterTax is found in the 'Alt' folder. The end results, assuming all has worked appropriately, will be found in the 'Fnl' folder. Note that the filtering and AlterTax modules are optional.

Each data project is associated with a code name, for example '10x10'. This code name will be used for all output files and the aggregation facility will prepare a folder in the 'Data' directory that is named using the code name. All output folders will be contained in that directory.

### F.1.2 Preliminaries

There are no proscribed ways to setup the directory structures for the various components of Env10 so the setup below is just a suggestion. All of the components rely on the user providing full path names so it should be flexible enough to handle most user preferences. Figure 17 depicts a fairly standard layout. The root directory is *Env10* and it contains three obligatory folders: *Data*, *SatAcct* and *Model*. The first contains the routines for aggregating and preparing the data for the model. It also contains some sub-folders that will be further described below. The *SatAcct* contains miscellaneous data that is used by the aggregation facility and eventually the model. Use of most of the satellite data is optional and is user-specified. The *Model* folder contains the core code for the ENVISAGE Model. The remaining folders are user folders and will contain the files for single projects. Single projects can be identified with a specific aggregation, different dynamic assumptions and/or a different focus.

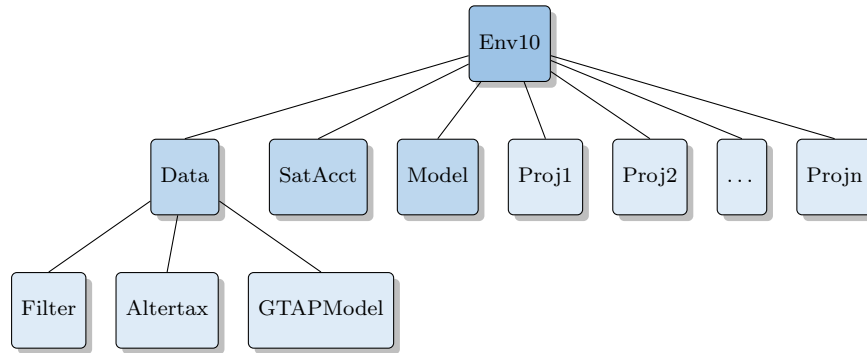


Figure 17: Possible directory structure

### Aggregation files

Table F.1 provides a list of files distributed with the data aggregation facility. In addition, the distribution will have a few sample 'MAP' files that users are likely to use as a starting point for new or modified aggregations. It is assumed that users will have access and appropriate licenses for the GTAP database files (see below for list of GTAP files). Most files need no user intervention. The exceptions would be the two base 'sets' definition files that would need to be updated with new database releases.<sup>3</sup>

<sup>3</sup> The 'sets' definition files are now configured for the new GTAP database configuration that accompanies Version 7 of the GTAP model. These 'sets' definition files have been given the suffix 'F'. The aggregation facility is not backward compatible and only works with the new database format.

Table F.1: Files distributed with the aggregation facility

<i>File name</i>	<i>Description</i>
<b>Core code</b>	
AggGTAP.gms	Main aggregation program
aggSAM.gms	Code to output aggregate SAM in a CSV file
aggNRG.gms	Code to output aggregate energy and emissions in a CSV file
SSPSets.gms	Set definitions for reading SSP data file
saveMap.gms	Code to save aggregation mappings to be loaded into either Word/Excel or L <sup>A</sup> T <sub>E</sub> X
makeset.gms	Code to write out core set definitions for models, both GTAP and ENVISAGE
<b>Additional code for Envisage Model</b>	
makesetEnv.gms	Code to write additional sets for ENVISAGE
AggEnvElast.gms	Code to aggregate ENVISAGE elasticities
aggra.gms	Batinclude code for aggregating elasticities
aggrav.gms	Batinclude code for aggregating elasticities
aggrave.gms	Batinclude code for aggregating elasticities
makeAggSets.gms	Code to write final sets with appropriate suffixes
ConvertLabel.gms	Code to convert labels in aggregated parameter data file
<b>Base data set definitions</b>	
GTAPSets9_2F.gms	Set definitions for standard GTAP database V9.2 (141 × 57)
GTAPSets9_2PF.gms	Set definitions for power+water GTAP database V9.0
<b>Miscellaneous</b>	
makedata.cmd	Windows command file for facilitating all data processing
nipa.gms	Optional GAMS code that will output (in CSV format) the national income and product accounts for the base regions/countries in the GTAP database

The aggregation facility assumes that all of the GTAP-based data files have the same base name and are located in the same folder.<sup>4</sup> Note that some of the files downloaded from GAMS may not follow this convention, in which case they need to be renamed. All original GTAP files are delivered in the GEMPACK 'HAR' format. They need to be converted to GAMS' 'GDX' format for use by the aggregation facility. The program 'HAR2GDX.EXE' is a utility that will convert from 'HAR' format to 'GDX' format.<sup>5</sup>

Table F.2 provides a list of all possible data files that can be used with the data aggregation facility. Some are required, for obvious reasons. Others are optional and will depend on user-based options. Note that the power+water version of the GTAP database includes three extra headers compared to the standard GTAP database. It has total power generated by electricity in Gwhr, water withdrawals in irrigated crops and aggregate water withdrawals from other sectors of the economy.

<sup>4</sup> Referred to as '%GTAPBASE%' and '%gtpDir%', respectively, in the GAMS code.

<sup>5</sup> At times it can be the case that the header array in the HAR file does not correspond to what is expected in the GDX interface. The GTAP Center is working on ensuring consistency. One possible solution is to use the '/H' option when running HAR2GDX.

Table F.2: Base Data Files

<i>File name</i>	<i>Description</i>
<b>Required GTAP files</b>	
%GTAPBASE%DAT.gdx	File containing base SAM data
%GTAPBASE%PAR.gdx	File containing base parameter values for GTAP model
%GTAPBASE%VOLE.gdx	File containing base energy volume data
%GTAPBASE%EMISS.gdx	File containing base CO <sub>2</sub> emissions data
<b>Optional data files</b>	
%GTAPBASE%NCO2.gdx	File containing base non-CO <sub>2</sub> emissions data
%GTAPBASE%BOP.gdx	File containing remittances and cross-border profit flows. The former is sourced from the 'GMIG' database and the latter from the 'GDYN' database.
%GTAPBASE%ELAST.gdx	Data file containing key land and natural resource supply parameters. This file is not part of the standard GTAP suite and is described in more detail below [TBD]. If the file does not exist, these parameters will be set to default values.
%GTAPBASE%WAGES.gdx	File containing wage/employment splits for labor remuneration. This file is part of the 'GMIG' database. Its use is described below [TBD].
%GTAPBASE%MRIO.gdx	File containing the split of import demand by end-user. This data is used for the Multi-Regional Input-Output version of the model, also referred to as the Global Value Chain (GVC) version. Note that the MRIO database has only three end-users: 'INT' for all intermediate demand; 'CONS' for private and public expenditures; and 'CGDS' for investment expenditures. The initialization module of the CGE models assumes proportionality across end-users within each of the broad categories. Use of this module requires setting the <code>ifMRIO</code> flag (see below). <sup>6</sup>
<b>Other data files</b>	
%SSPFILE%	Data file containing projections for population and GDP for 230 countries for all 5 Shared Socio-Economic Pathways (SSPs)—further described below [TBD]. In addition, the file contains the UN's population projections for the 2010, 2012, 2015 and 2017 revisions.
%GIDDLAB%	Data file containing an alternative wage/employment split. The World Bank's Global Income Distribution Dynamics (GIDD) Model provides an alternative set for employment levels. Note that the GIDD data is only provided for two types of labor—unskilled and skilled.
%GIDDPROJ%	Data file containing population projections (including education levels) from the GIDD.
<b>Parameter files</b>	
%EnvElast%	File containing disaggregated elasticities for use with the ENVISAGE Model. Two files are distributed. One for the standard GTAP database and the other for the power+water version.

### Filter module files

The filter module files are scattered in two folders—see Table F.3. The *Data* folder contains the calling `filter` routine and two user prepared files. The core files for the Filter module are all in the *Filter* sub-folder.

Table F.3: Files distributed with the Filter module

<i>File name</i>	<i>Description</i>
<b><i>Files in core 'Data' directory</i></b>	
<code>Filter.gms</code>	GAMS file invoking core filter code
<code>basenameFlt.gms</code>	User prepared file with filter options
<code>basenameMap.gms</code>	User prepared file with key mappings—same file used for Aggregation facility
<b><i>Additional code for Filter module in the 'Filter' directory</i></b>	
<code>filter.gms</code>	GAMS code containing core filter algorithm
<code>remTinyValues.gms</code>	Code to remove tiny value flows
<code>itrlog.gms</code>	Code that creates diagnostics for the Filter module
<code>title.gms</code>	Code mostly intended for the CGEBox GUI

### Altertax module files

The Altertax files are scattered in three folders, see Table F.4. The main *Data* folder contains `Altertax.gms` and the user options file `[basename]Alt.gms`. The sub-folder *GTAPModel* contains the core code for the GAMS-based GTAP model. The sub-folder *Altertax* contains a single file with the parameter definitions needed to run the GTAP model. Note that many of the parameters will be over-ridden by Altertax to make all CES functions (and the utility function) Cobb-Douglas, further described below.

Table F.4: Files distributed with the Alvertax module

<i>File name</i>	<i>Description</i>
<b>Files in core 'Data' directory</b>	
Alvertax.gms	GAMS file invoking core GTAP model code
basenameAlt.gms	User prepared file with the Alvertax shock
<b>Additional code for Alvertax module in the 'GTAPModel' directory</b>	
model.gms	GAMS code containing core GTAP model specification
getData.gms	GAMS code to read core database(s)
cal.gms	GAMS code containing variable and parameter initialization and calibration
iterloop.gms	Code containing instructions to prepare a new solve statement
solve.gms	Code invoking a GAMS solver
postsim.gms	Code saving model results in a CSV-formatted file
emiCSV.gms	Code to save emission results in a CSV-formatted file
saveData.gms	Code saving model results in GDX-formatted files with the same structure as the input GTAP-based data files
<b>Additional code for Alvertax module in the 'AlterTax' directory</b>	
AlterTaxPrm.gms	GAMS code containing default elasticities for GTAP model.

### Data sub-folders

Each aggregation project, with its code, will be associated with a data folder that has the same name as the project. So if the code for the project is '10x10', for example, the data from the data preparation modules will be contained in the folder '10x10'. The output from each module is contained in its own sub-folder of which there are four named **Agg**, **Flt**, **Alt** and **Fn1**. The last folder contains the final results from running the three modules. If a module is skipped, all the data files from the previous module should be copied over to the next folder. Users can create these folders, or use the automated Windows command file called `makedata.cmd` that is further described below. The `makeData` command file also automates the moving of files across folders. Figure 18 provides a visual guide to the directory structure of the output data.

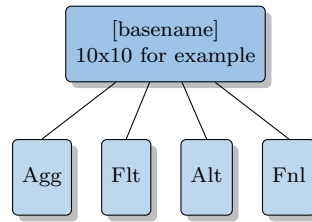


Figure 18: Directory structure of output data

### F.1.3 Aggregation

From the user perspective aggregation involves the preparation of the map file that will contain the aggregation mappings as well as miscellaneous other information required for the data facility. The mapping file has two components. The first part relates to a generic aggregation of the GTAP

database and its components. It hones closely to the GTAPAgg and FlexAgg facilities available for GEMPACK users. The second component is specific to aggregation for the ENVISAGE Model.

### Global options

The first part of a map file contains global options for the aggregation. Listing F.1 provides an example of the preamble statements for a mapping file. The first line is an include file for the base data's set definitions. The distribution comes with two prepared set definition files—one for the standard GTAP database and the other for the power+water database. The following set of lines provide information for locating the input databases. The first is the location of all of the GTAP data files (DAT, PAR, VOLE, EMISS, NCO2, BOP, ELAST and WAGES). The second line provides the base name for the GTAP files, for example GSD or GSDPW—the latter stands for the power+water database. The remaining four lines refer to the full path name of the other input databases—the SSP projection file, the two GIDD databases and the file of disaggregated ENVISAGE elasticities.

Listing F.1: Global options for the MAP file

---

```

1  $include "GTAPSets9-2F.gms"

3  $setGlobal gtpDir      "V:\GTAP9\V9-2\"
4  $setGlobal GTAPBASE    "GSDF"
5  $setGlobal SSPFile     "V:\ClimateChange\AR5\SSP\sspScenV9\sspScenV9.gdx"
6  $setGlobal giddLab     "..\SatAcct\giddLab.gdx"
7  $setGlobal giddProj    "..\SatAcct\giddProj.gdx"
8  $setGlobal EnvElast    "V:\GTAP9\Elast\EnvLinkElast.gdx"

10 $setGlobal DYN         ON
11 $setGlobal ifPower     OFF
12 $setGlobal ifWater     OFF
13 $setGlobal NCO2        ON
14 $setGlobal ELAST       OFF
15 $setGlobal LAB         OFF
16 $setGlobal BoP         ON
17 $setGlobal SAVEMAP     TXT
18 $setGlobal ifMRIO      OFF

20 * Only used to override GTAP parameters for Env model

22 $setGlobal OVRRIDEGTAPARM 0
23 $setGlobal OVRRIDEGTAPINC 0

25 * Select a labor option
26 * Valid options are:
27 *   noLab — ignore employment volumes (all wages are set to 1)
28 *   agLab — calculate ag and non-ag employment (wages uniform within zones)
29 *   allLab — assume employment data is correct for each sector (wages differ for each sector)
30 *   giddLab — Use the GIDD labor data

32 $macro IFLABOR noLab

34 $onempty

```

---

The remaining set of options turn on or off various features of the aggregation. These are described in further detail in Table F.5.

Table F.5: Global options for the MAP file

<i>Option name</i>	<i>Description</i>
DYN	If this option is set to ON, the aggregation facility will read the SSP and GIDD projection files and create the scenario file with the appropriate aggregation. This is to be used for dynamic versions of the model.
ifPower	If this option is set to ON, the aggregation facility will retrieve the Gwhr information from the GTAP (power) database and aggregate.
ifWater	If this option is set to ON, the aggregation facility will retrieve the water volume information from the GTAP (water) database and aggregate.
NCO2	If this option is set to ON, the aggregation facility will retrieve the non-CO <sub>2</sub> information from the non-CO <sub>2</sub> database and aggregate.
ELAST	If this option is set to ON, the aggregation facility will retrieve the land and natural resource parameters from the relevant 'ELAST' file. It should be set to OFF for an ENVISAGE aggregation as the same parameters are contained in the parameter file.
LAB	If this option is set to ON, the aggregation facility will retrieve the employment data from one of the available databases—either GTAP or the GIDD—and aggregate. The exact type of aggregation will depend on the IFLABOR option described below.
BOP	If this option is set to ON, the aggregation facility will retrieve the additional balance of payments data from the 'BOP' database and aggregate.
SAVEMAP	This option takes two values: 'TXT' and 'LATEX'. The 'TXT' option will save the aggregation mappings in a file called '[baseBame]Sets.txt'. It is a comma delimited text file (CSV) that contains the aggregation mappings from the original GTAP database to the model-level aggregation. It can be loaded into Excel and the <code>Data Convert to Columns</code> command can be used to convert the data into a table with three columns, using the 'comma' delimiter option. It is then easy to copy the table into Word. The 'LATEX' option creates a file with the name '[baseName]Sets.tex' and creates three relatively generic tables that can be included in a L <sup>A</sup> T <sub>E</sub> X file. The user may have to edit the file to convert special characters—such as underscores ('_') and ampersands ('&').
ifMRIO	This option will request aggregation of the MRIO database. Options are ON and OFF. If set to ON, it requires the MRIO database associated with the given GTAP version.
OVERRIDEGTAPARM	Currently ignored and to be reviewed.
OVERRIDEGTAPINC	If this option is set to 1, the income elasticities will be taken from the input base parameter file. If the option is set to 0, the GTAP income elasticities prevail and override the income elasticities from the parameter file.
IFLABOR	This option takes four possible values. The first value is 'noLab'. In this case the employment volumes are ignored and in effect all wages are set to 1. The second value is 'agLab'. In this case the volume data is used to aggregate agricultural and non-agricultural employment. Within each of these two broad sectors wages are assumed to be uniform across sectors and in effect the implicit inter-sectoral wage differences are ignored. The purpose of this is to minimize potential welfare implications by assuming heterogeneous inter-sectoral wages (apart from the ag/non-ag distinction). The third option is 'allLab' in which case the employment volume data is fully utilized and explicit inter-sectoral wages are assumed to hold. The fourth option is 'giddLab' in which case the GIDD employment volumes are used to determine the wage/employment split. Note that the GIDD option implies a skill/non-skilled labor aggregation.



## Standard mappings

The second part of the mapping file contains the standard mappings—this is aimed for the GTAP model and hones very closely to GTAPAGG, the GEMPACK-based aggregation facility. Users define the aggregated sets for activities/commodities, regions and factors of production—with their respective mappings to the full GTAP database. Users have some flexibility in how to define these sets and mappings and the distribution comes with two examples.

Listing F.2 provides an example that is virtually identical to the 10x10 mapping example from GTAP. A first set  $i$  is defined that contains the 10 aggregate sectors. Note that since the current database is diagonal, as is the output from the aggregation facility, the set for activities ( $a$ ) is aliased with the set for sectors ( $i$ ). Similarly, the user defines set  $r$  for regions and  $fp$  for the factors of production. These basic sets are subsequently decomposed into a number of subsets—most of which are fairly obvious. Some do have implications for the model:

- The subset  $ul$  is described as the set of unskilled labor. In fact, it is the set of labor types that substitute for the capital/skilled bundle.<sup>7</sup> It is possible to assign all labor types to the subset  $ul$ , in which case the so-called skilled labor bundle will be empty. The subset  $sl$  will be determined as the complement of the subset  $ul$ .
- The subset  $lr$  determines the labor types used to calculate the skill premium. This became necessary with the new GTAP database that has 5 labor types.
- The  $wat$  subset should be empty in the absence of the water factor—only available with the power+water database.
- The  $mapz$  mapping defines the ag vs. non-ag activities. This is only used to determine ag vs. non-ag wages in the case the 'IFLABOR' option is set to  $agLab$ .
- The set  $mapt$  will instruct the aggregation facility to merge land and capital payments into capital payments only—for the selected activities. For example, it is possible to remove the land payment in other livestock (the  $oap$  sector).
- The set  $mapn$  has the same purpose, but for the natural resource factor. For example, one may wish to delete the natural resource factor payment if the  $omn$  activity is merged with manufacturing; as depending on the supply elasticity of the natural resource, this may overly restrict output expansion (or contraction).

Post-simulation analysis of the results allow for aggregation of activities, commodities, regions and labor. These are at the full discretion of the user. The aggregate sets are given respectively by  $aga$ ,  $ia$ ,  $ra$  and  $lagg$ . The aggregation facility will automatically append the suffixes '-a' and '-c' for the individual activities and commodities, respectively. So it would be usual practice to define the aggregation set items with the '-a' and '-c' suffixes for compatibility. For purposes of post-simulation, these sets will also include the individual set items. For example, the aggregation facility will write out the final set  $ra$  as the union of the set  $r$  and  $ra$ , but the user only needs to provide the definitions for the aggregate items as the aggregation facility will take care of merging the two sets. The definition of these aggregations (with their respective mappings) simplifies the output code. For example, the code to aggregate GDP across regions takes the form:

---

```
2 loop(ra, put ra.tl, (sum(r$mapra(ra,r), gdp(r)) / ; ) ;
```

---

<sup>7</sup> This determines which labor types are assigned to the  $LAB1$  and  $LAB2$  bundles.

where the mapping *mapra* defines the aggregation mapping for each region *ra*, including the individually modeled regions.

The aggregation facility requires the factor mobility parameters for the GTAP model. This is given in the parameter **etrae1**. In the sample map file, **etrae1** is initialized with **etrae0** that has no regional index and thus the key factor mobility elasticities in all regions are identical, but this is not a requirement.

The new GTAP standard model has an explicit make matrix. Multi-product activities are governed by a CET function with the transformation elasticity provided by **etraq1** with a default value of 5. The formation of supply from production of multiple activities is governed by a CES function with a substitution provided by **esubq1** with a default elasticity of  $\infty$ . Similarly, the new GTAP standard model has new flexibility regarding the CES expenditure elasticities for government, investment and the formation of transportation margins. These are represented respectively in the parameters **esubg1**, **esubi1** and **esubs1**. These are initialized to their default values.

The aggregation mappings are relatively self-explanatory. The exception may be the mapping *mapl* that will match the user-input labor types to the GIDD labor labels that are fixed.

### Listing F.2: Standard mappings

```

1  sets

3  i  "Commodities" /
4      GrainsCrops    "Grains and Crops"
5      MeatLstk       "Livestock and Meat Products"
6      Extraction     "Mining and Extraction"
7      ProcFood       "Processed Food"
8      TextWapp       "Textiles and Clothing"
9      LightMnfc      "Light Manufacturing"
10     HeavyMnfc      "Heavy Manufacturing"
11     Util.Cons      "Utilities and Construction"
12     TransComm      "Transport and Communication"
13     OthServices    "Other Services"
14     /

16  r  "Regions" /
17     Oceania        "Australia, New Zealand"
18     EastAsia       "East Asia"
19     SEAsia         "Southeast Asia"
20     SouthAsia      "South Asia"
21     NAmerica       "North America"
22     LatinAmer      "Latin America"
23     EU.28          "European Union 28"
24     MENA           "Middle East and North Africa"
25     SSA            "Sub-Saharan Africa"
26     RestofWorld    "Rest of World"
27     /

29  fp  "Factors of production" /
30     UnSkLab        "Unskilled labor"
31     SkLab          "Skilled labor"
32     Capital        "Capital"
33     Land           "Land"
34     NatRes         "Natural resource"
35     /

37  l(fp)  "Labor factors" /
38     UnSkLab        "Unskilled labor"
39     SkLab          "Skilled labor"
40     /
41  ul(l)  "Unskilled labor" /
42     UnSkLab        "Unskilled labor"
43     SkLab          "Skilled labor"
44     /

```

```

45   lr(l) "Reference labor for skill premium" /
46       SkLab      "Skilled labor"
47   /
48   cap(fp) "Capital" /
49       Capital    "Capital"
50   /
51   lnd(fp) "Land endowment" /
52       Land       "Land"
53   /
54   nrs(fp) "Natural resource" /
55       NatRes     "Natural resource"
56   /
57   wat(fp) "Water resource" /
58       /

60   ra "Aggregate regions for emission regimes and model output" /
61       hic      "High-income countries"
62       lmy      "Developing countries"
63       wld      "World Total"
64   /
65   ia "Aggregate commodities for model output" /
66       tagr-c    "Agriculture"
67       tman-c    "Manufacturing"
68       tsrv-c    "Services"
69       toth-c    "Other"
70       ttot-c    "Total"
71   /
72   aga "Aggregate activities for model output" /
73       tagr-a    "Agriculture"
74       tman-a    "Manufacturing"
75       tsrv-a    "Services"
76       toth-a    "Other"
77       ttot-a    "Total"
78   /
79   lagg "Aggregate labor for model output" /
80       tot       "Total labor"
81   /
82 ;

84 * !!!! Explicit assumption about diagonality

86 alias(i,a) ;

88 alias(m,i) ;

90 * User defined parameters (i.e. not aggregated by aggregation facility)
91 * NEW — New region specific

93 Parameter
94   etrael(fp,r) "CET transformation elasticities for factor allocation"
95 ;

97 parameter etrae0(fp) "CET transformation elasticities for factor allocation" /
98   UnSkLab    inf
99   SkLab      inf
100  Capital    inf
101  Land       1.00
102  NatRes     0.001
103 / ;

105 etrael(fp,r) = etrae0(fp) ;

107 * NEW — MAKE ELASTICITIES

109 Parameter
110   etraq1(a,r)      "MAKE CET Elasticity"
111   esubq1(i,r)      "MAKE CES Elasticity"
112 ;

```

```

113 etraq1(a,r) = 5 ;
114 esubq1(i,r) = inf ;

116 * NEW — EXPENDITURE ELASTICITIES

118 Parameter
119     esubg1(r)      "Government expenditure CES elasticity"
120     esubil1(r)     "Investment expenditure CES elasticity"
121     esubsl1(m)     "Transport margins CES elasticity"
122 ;

124 esubg1(r) = 1 ;
125 esubil1(r) = 0 ;
126 esubsl1(m) = 1 ;

128 * This zonal mapping is for labor/volume splits between agriculture and other

130 set mapz(z,a)  "Mapping of activities to zones" /
131     rur.(GrainsCrops,MeatLstk)
132 / ;

134 mapz("urb",a) = not mapz("rur",a) ;
135 mapz("nsg",a) = yes ;

137 * >>>> MUST INSERT RESIDUAL REGION (ONLY ONE)

139 set rres(r)  "Residual region" /

141     NAmerica

143 / ;

145 * >>>> MUST INSERT MUV REGIONS (ONE OR MORE)

147 set rmuv(r)  "MUV regions" /

149     Oceania
150     EastAsia
151     NAmerica
152     EU_28

154 / ;

156 set mapt(a)  "Merge land and capital payments in the following sectors" /

158 / ;

160 set mapn(a)  "Merge natl. res. and capital payments in the following sectors" /

162 / ;

164 * MAPPINGS TO GTAP

166 set mapa(acts,a) /
167     pdr . GrainsCrops
168     wht . GrainsCrops
169     gro . GrainsCrops
170     v_f . GrainsCrops
171     osd . GrainsCrops
172     c_b . GrainsCrops
173     pfb . GrainsCrops
174     ocr . GrainsCrops
175     ctl . MeatLstk
176     oap . MeatLstk
177     rmk . MeatLstk
178     wol . MeatLstk
179     frs . Extraction
180     fsh . Extraction

```

```

181 coa . Extraction
182 oil . Extraction
183 gas . Extraction
184 omn . Extraction
185 cmt . MeatLstk
186 omt . MeatLstk
187 vol . ProcFood
188 mil . ProcFood
189 pcr . GrainsCrops
190 sgr . ProcFood
191 ofd . ProcFood
192 b_t . ProcFood
193 tex . TextWapp
194 wap . TextWapp
195 lea . LightMnfc
196 lum . LightMnfc
197 ppp . LightMnfc
198 p_c . HeavyMnfc
199 crp . HeavyMnfc
200 nmm . HeavyMnfc
201 i_s . HeavyMnfc
202 nfm . HeavyMnfc
203 fmp . LightMnfc
204 mvh . LightMnfc
205 otn . LightMnfc
206 ele . HeavyMnfc
207 ome . HeavyMnfc
208 omf . LightMnfc
209 ely . Util.Cons
210 gdt . Util.Cons
211 wtr . Util.Cons
212 cns . Util.Cons
213 trd . TransComm
214 otp . TransComm
215 wtp . TransComm
216 atp . TransComm
217 cmn . TransComm
218 ofi . OthServices
219 isr . OthServices
220 obs . OthServices
221 ros . OthServices
222 osg . OthServices
223 dwe . OthServices
224 / ;

226 set mapr(reg,r) /
227 aus . Oceania
228 nzl . Oceania
229 xoc . Oceania
230 chn . EastAsia
231 hkg . EastAsia
232 jpn . EastAsia
233 kor . EastAsia
234 mng . EastAsia
235 twm . EastAsia
236 xea . EastAsia
237 brn . EastAsia
238 khm . SEAsia
239 idn . SEAsia
240 lao . SEAsia
241 mys . SEAsia
242 phl . SEAsia
243 sgp . SEAsia
244 tha . SEAsia
245 vnm . SEAsia
246 xse . SEAsia
247 bgd . SouthAsia
248 ind . SouthAsia

```

```
249 npl . SouthAsia
250 pak . SouthAsia
251 lka . SouthAsia
252 xsa . SouthAsia
253 can . NAmerica
254 usa . NAmerica
255 mex . NAmerica
256 xna . NAmerica
257 arg . LatinAmer
258 bol . LatinAmer
259 bra . LatinAmer
260 chl . LatinAmer
261 col . LatinAmer
262 ecu . LatinAmer
263 pry . LatinAmer
264 per . LatinAmer
265 ury . LatinAmer
266 ven . LatinAmer
267 xsm . LatinAmer
268 cri . LatinAmer
269 gtm . LatinAmer
270 hnd . LatinAmer
271 nic . LatinAmer
272 pan . LatinAmer
273 slv . LatinAmer
274 xca . LatinAmer
275 dom . LatinAmer
276 jam . LatinAmer
277 pri . LatinAmer
278 tto . LatinAmer
279 xcb . LatinAmer
280 aut . EU.28
281 bel . EU.28
282 cyp . EU.28
283 cze . EU.28
284 dnk . EU.28
285 est . EU.28
286 fin . EU.28
287 fra . EU.28
288 deu . EU.28
289 grc . EU.28
290 hun . EU.28
291 irl . EU.28
292 ita . EU.28
293 lva . EU.28
294 ltu . EU.28
295 lux . EU.28
296 mlt . EU.28
297 nld . EU.28
298 pol . EU.28
299 prt . EU.28
300 svk . EU.28
301 svn . EU.28
302 esp . EU.28
303 swe . EU.28
304 gbr . EU.28
305 che . RestofWorld
306 nor . RestofWorld
307 xef . RestofWorld
308 alb . RestofWorld
309 bgr . EU.28
310 blr . RestofWorld
311 hrv . EU.28
312 rou . EU.28
313 rus . RestofWorld
314 ukr . RestofWorld
315 xee . RestofWorld
316 xer . RestofWorld
```

```
317     kaz . RestofWorld
318     kgz . RestofWorld
319     tjk . RestofWorld
320     xsu . RestofWorld
321     arm . RestofWorld
322     aze . RestofWorld
323     geo . RestofWorld
324     bhr . MENA
325     irn . MENA
326     isr . MENA
327     jor . MENA
328     kwt . MENA
329     omn . MENA
330     qat . MENA
331     sau . MENA
332     tur . MENA
333     are . MENA
334     xws . MENA
335     egy . MENA
336     mar . MENA
337     tun . MENA
338     xnf . MENA
339     ben . SSA
340     bfa . SSA
341     cmr . SSA
342     civ . SSA
343     gha . SSA
344     gin . SSA
345     nga . SSA
346     sen . SSA
347     tgo . SSA
348     xwf . SSA
349     xcf . SSA
350     xac . SSA
351     eth . SSA
352     ken . SSA
353     mdg . SSA
354     mwi . SSA
355     mus . SSA
356     moz . SSA
357     rwa . SSA
358     tza . SSA
359     uga . SSA
360     zmb . SSA
361     zwe . SSA
362     xec . SSA
363     bwa . SSA
364     nam . SSA
365     zaf . SSA
366     xsc . SSA
367     xtw . RestofWorld
368 / ;

370 set mapf(endw, fp) /
371     ag_othlowsk . UnSkLab
372     service_shop . UnSkLab
373     clerks . UnSkLab
374     tech.aspros . SkLab
375     off_mgr_pros . SkLab
376     Capital . Capital
377     Land . Land
378     NatlRes . NatRes
379 / ;

381 set mapl(lg, 1) "Mapping to GIDD labor database" /
382     nsk.UnSkLab
383     skl.SkLab
384 / ;
```

---

## Sets and mappings for model aggregation

As noted above, the model aggregation can differ from the data aggregation. The data aggregation is built around the concept of a diagonal make matrix—this ensures that the output from any of the three modules always has the same ‘geometry’. The model aggregation is intended to allow for a non-diagonal make matrix—this can vastly simplify a model, for example in incorporating the recent GTAP power database. This section allows the user to input the sets and mappings for the model aggregation. Listing F.3 has an example from the ‘10x10’ map file.

Most of the entries in the listing file should be fairly self-explanatory. The user defines two sets, `actf` and `commf`, that will represent the model aggregation. They can replicate the set  $i$  from the first section, in which case the make matrix will be the standard diagonal. In the example from the listing file, the two agricultural activities from the data aggregation are merged into one activity called **Agriculture**. Thus the single activity **Agriculture** will produce two commodities: **GrainsCrops** and **MeatLstk**. This is reflected in the two mapping sets `mapaf` and `mapif`. Note that the mappings are based on the data aggregation and not the original GTAP database.<sup>8</sup>

The model needs an `imuvf` subset based on the final model aggregation. This is followed by the definition of the aggregations of commodities, activities and regions, i.e. the set mappings `mapia`, `mapaga` and `mapra`. Note that these mappings are ignored by the model and typically are only used for post-processing of results.

The mapping file also includes a user-determined sort order for regions, activities and commodities. GAMS does not preserve the order of set labels. The sort orders can be used upon output to guarantee that output will be sorted according to user preferences. The sort order can also be used to sort labels for Excel (for example in the creation of Pivot tables.).

Listing F.3: Sets for model aggregation

---

```

1  *
2  *
3  *      Section dealing with model aggregations (to handle non-diagonal make matrix)
4  *
5  *
6  *
7  *      Model aggregation(s)
8  *
9  set actf "Model activities" /
10      Agriculture      "Agriculture"
11      Extraction       "Mining and Extraction"
12      ProcFood         "Processed Food"
13      TextWapp         "Textiles and Clothing"
14      LightMnfc        "Light Manufacturing"
15      HeavyMnfc        "Heavy Manufacturing"
16      Util.Cons        "Utilities and Construction"
17      TransComm        "Transport and Communication"
18      OthServices      "Other Services"
19  / ;
20
21 set commf "Model commodities" /
22      GrainsCrops      "Grains and Crops"
23      MeatLstk         "Livestock and Meat Products"
24      Extraction       "Mining and Extraction"
25      ProcFood         "Processed Food"
26      TextWapp         "Textiles and Clothing"
27      LightMnfc        "Light Manufacturing"
28      HeavyMnfc        "Heavy Manufacturing"
29      Util.Cons        "Utilities and Construction"
30      TransComm        "Transport and Communication"

```

---

<sup>8</sup> The aggregation facility will automatically append the ‘-a’ and ‘-c’ suffixes to activities and commodities, respectively, except for the user-defined aggregations.



```

31      OthServices      "Other Services"
32  / ;

34  set mapaf(i, actf) "Mapping from original to modeled activities" /
35      GrainsCrops      .Agriculture
36      MeatLstk         .Agriculture
37      Extraction       .Extraction
38      ProcFood         .ProcFood
39      TextWapp         .TextWapp
40      LightMnfc        .LightMnfc
41      HeavyMnfc        .HeavyMnfc
42      Util.Cons        .Util.Cons
43      TransComm        .TransComm
44      OthServices      .OthServices
45  / ;

47  set mapif(i, commf) "Mapping from original to modeled commodities" /
48      GrainsCrops      .GrainsCrops
49      MeatLstk         .MeatLstk
50      Extraction       .Extraction
51      ProcFood         .ProcFood
52      TextWapp         .TextWapp
53      LightMnfc        .LightMnfc
54      HeavyMnfc        .HeavyMnfc
55      Util.Cons        .Util.Cons
56      TransComm        .TransComm
57      OthServices      .OthServices
58  / ;

60  * >>>> MUST INSERT MUV COMMODITIES (ONE OR MORE)
61  *      !!!! Be careful of compatibility with modeled imuv
62  *      This one is intended for AlterTax

64  set imuvf(commf) "MUV commodities" /

66      ProcFood
67      TextWapp
68      LightMnfc
69      HeavyMnfc

71  / ;

73  * >>>> Aggregation of modeled sectors and regions

75  set mapia(ia, commf) "mapping of individual comm to aggregate comm" /
76      tagr-c.GrainsCrops
77      tagr-c.MeatLstk
78      tman-c.ProcFood
79      tman-c.TextWapp
80      tman-c.LightMnfc
81      tman-c.HeavyMnfc
82      toth-c.Extraction
83      tsrv-c.Util.Cons
84      tsrv-c.TransComm
85      tsrv-c.OthServices
86  / ;
87  mapia("ttot-c", commf) = yes ;

89  set mapaga(aga, actf) "mapping of individual comm to aggregate comm" /
90      tagr-a.Agriculture
91      tman-a.ProcFood
92      tman-a.TextWapp
93      tman-a.LightMnfc
94      tman-a.HeavyMnfc
95      toth-a.Extraction
96      tsrv-a.Util.Cons
97      tsrv-a.TransComm
98      tsrv-a.OthServices

```

---

```

99 / ;
100 mapaga("ttot-a",actf) = yes ;

102 set mapra(ra,r) "Mapping of model regions to aggregate regions" /
103     hic.(Oceania, NAmerica, EU28)
104 / ;
105 mapra("lmy", r)$(not mapra("hic",r)) = yes ;
106 mapra("wld", r) = yes ;

108 set maplagg(lagg,l) "Mapping of model labor to aggregate labor" ;
109 maplagg("Tot",l) = yes ;

111 set sortOrder / sort1*sort500 / ;
112 set mapRegSort(sortOrder,r) /

114     sort1 . Oceania
115     sort2 . EastAsia
116     sort3 . SEAsia
117     sort4 . SouthAsia
118     sort5 . NAmerica
119     sort6 . LatinAmer
120     sort7 . EU_28
121     sort8 . MENA
122     sort9 . SSA
123     sort10. RestofWorld

125 / ;

127 set mapActSort(sortOrder,actf) /

129     sort1 .Agriculture
130     sort2 .Extraction
131     sort3 .ProcFood
132     sort4 .TextWapp
133     sort5 .LightMnfc
134     sort6 .HeavyMnfc
135     sort7 .Util.Cons
136     sort8 .TransComm
137     sort9 .OthServices

139 / ;

141 set mapCommSort(sortOrder,commf) /
142     sort1 .GrainsCrops
143     sort2 .MeatLstk
144     sort3 .Extraction
145     sort4 .ProcFood
146     sort5 .TextWapp
147     sort6 .LightMnfc
148     sort7 .HeavyMnfc
149     sort8 .Util.Cons
150     sort9 .TransComm
151     sort10 .OthServices
152 / ;

```

---

## Sets and mappings for the Envisage Model

The final part of the map file is specific to the ENVISAGE Model. Many are critical components of ENVISAGE as they often determine model specification. For example, the first two subsets, **acr** and **alv** determine which activities use the crop and livestock production structure, respectively. All other activities will be assigned to the subset **ax** that is the default production structure. Listing F.4 provides an example for the 10x10 map file, though we will provide additional snippets below for a different aggregation tuned to the power and water modules of the model.

The subsets `agr` and `man` are not formally part of the model. These subsets are currently being used to determine activity-specific productivity shifters in dynamic scenarios.<sup>9</sup> They could potentially be used as well in post-simulation processing. The subsets `aenergy`, `affl` and `aw` are currently not being used.

Listing F.4: Sets for the Envisage Model aggregation

---

```

1
2 *
3 *   Envisage section
4 *
5 *
6
7 * >>>> Activity related sets and subsets
8
9 set acr(actf)   "Crop activities" /
10 / ;
11
12 set alv(actf)   "Livestock activities" /
13 / ;
14
15 set agr(actf)   "Agricultural activities" /
16     Agriculture   "Agriculture"
17 / ;
18
19 set man(actf)   "Manufacturing activities" /
20     Extraction    "Mining and Extraction"
21     ProcFood      "Processed Food"
22     TextWapp      "Textiles and Clothing"
23     LightMnfc     "Light Manufacturing"
24     HeavyMnfc     "Heavy Manufacturing"
25 / ;
26
27 set aenergy(actf) "Energy activities" /
28     Extraction    "Mining and Extraction"
29     UtilCons      "Utilities and Construction"
30 / ;
31
32 set affl(actf)   "Fossil fuel activities" /
33     Extraction    "Mining and Extraction"
34 / ;
35
36 set aw(actf)     "Water services activities" /
37 / ;
38
39 set elya(actf)   "Power activities" /
40 / ;
41
42 set etd(actf)    "Electricity transmission and distribution activities" /
43 / ;
44
45 set primElya(actf) "Primary power activities" /
46 / ;
47
48 set pb          "Power bundles" /
49     othp        "Other power"
50 / ;
51
52 set mappow(pb,elya) "Mapping of power activities to power bundles" /
53 / ;
54
55 * >>>> Commodity sets and subsets
56
57 set frt(commf)   "Fertilizer commodities" /
58 / ;

```

---

<sup>9</sup> See the model file `initScen.gms`

```

60 set feed(commf) "Feed commodities" /
61 / ;

63 set iw(commf) "Water services commodities" /
64 / ;

66 set e(commf) "Energy commodities" /
67 $ontext
68     Extraction      "Mining and Extraction"
69     Util.Cons       "Utilities and Construction"
70 $offtext
71 / ;

73 set elyc(commf) "Electricity commodities" /
74     Util.Cons       "Utilities and Construction"
75 / ;

77 set f(commf) "Fuel commodities" /
78 *     Extraction    "Mining and Extraction"
79 / ;

81 * This zonal mapping is for labor market segmentation in final model

83 set mapzf(z,actf) "Mapping of activities to zones" /
84     rur.Agriculture
85 / ;

87 mapzf("urb",actf) = not mapzf("rur",actf) ;
88 mapzf("nsg",actf) = yes ;

90 * >>>> Household commodity section

92 set k "Household commodities" /
93     GrainsCrops     "Grains and Crops"
94     MeatLstk        "Livestock and Meat Products"
95     ProcFood        "Processed Food"
96     TextWapp        "Textiles and Clothing"
97     LightMnfc       "Light Manufacturing"
98     HeavyMnfc       "Heavy Manufacturing"
99     TransComm       "Transport and Communication"
100    OthServices      "Other Services"
101    Energy           "Energy"
102 / ;

104 set fud(k) "Household food commodities" /
105     GrainsCrops     "Grains and Crops"
106     MeatLstk        "Livestock and Meat Products"
107     ProcFood        "Processed Food"
108 / ;

110 set mapk(commf,k) "Mapping from i to k" /
111     GrainsCrops     .GrainsCrops
112     MeatLstk        .MeatLstk
113     Extraction      .Energy
114     ProcFood        .ProcFood
115     TextWapp        .TextWapp
116     LightMnfc       .LightMnfc
117     HeavyMnfc       .HeavyMnfc
118     Util.Cons       .Energy
119     TransComm       .TransComm
120     OthServices     .OthServices
121 / ;

123 set lb "Land bundles" /
124     agr             "Agriculture"
125 / ;

```

```

127 set lb1(lb) "First land bundle" /
128     agr          "Livestock"
129 / ;

131 set maplb(lb,actf) "Mapping of activities to land bundles" /
132     agr          .Agriculture
133 / ;

135 * !!!! TO BE REVIEWED

137 set lb0 "Default land bundles" / lb1*lb1 / ;
138 set maplb0(lb, lb0) "Mapping of land bundles to original" /
139     agr.lb1
140 / ;

142 set wbnd "Aggregate water markets" /
143     N_A          "N_A"
144 / ;

146 set wbnd1(wbnd) "Top level water markets" /
147 / ;

149 set wbnd2(wbnd) "Second level water markets" /
150 / ;

152 set wbndEx(wbnd) "Exogenous water markets" /
153 / ;

155 set mapw1(wbnd,wbnd) "Mapping of first level water bundles" /
156 / ;

158 set mapw2(wbnd,actf) "Mapping of second level water bundle" /
159 / ;

161 set wbnda(wbnd) "Water bundles mapped one-to-one to activities" /
162 / ;

164 set wbndi(wbnd) "Water bundles mapped to aggregate output" /
165 / ;

167 set NRG "Energy bundles used in model" /
168     coa          "Coal"
169     oil          "Oil"
170     gas          "Gas"
171     ely          "Electricity"
172 / ;

174 set coa(NRG) "Coal bundle used in model" /
175     coa          "Coal"
176 / ;

178 set oil(NRG) "Oil bundle used in model" /
179     oil          "Oil"
180 / ;

182 set gas(NRG) "Gas bundle used in model" /
183     gas          "Gas"
184 / ;

186 set ely(NRG) "Electricity bundle used in model" /
187     ely          "Electricity"
188 / ;

190 set mape(NRG,e) "Mapping of energy commodities to energy bundles" /
191 * oil          .Extraction
192 * ely          .Util.Cons
193 / ;

```

---

```

195 * >>>> Sets required for 'growing' labor by skill

197 set skl(l) "Skill types for labor growth assumptions" /
198     SkLab
199 / ;

201 set elev / elev0*elev3 / ;

203 set educMap(r,l,elev) "Mapping of skills to education levels" ;

205 * Use GIDD definitions (i.e. "elev3" has no meaning)

207 educMap(r,"UnSkLab","elev0")$mapra("lmy",r) = yes ;
208 educMap(r,"SkLab","elev1")$mapra("lmy",r) = yes ;
209 educMap(r,"SkLab","elev2")$mapra("lmy",r) = yes ;

211 educMap(r,"UnSkLab","elev0")$mapra("hic",r) = yes ;
212 educMap(r,"UnSkLab","elev1")$mapra("hic",r) = yes ;
213 educMap(r,"SkLab","elev2")$mapra("hic",r) = yes ;

215 $offempty

```

---

The subsets **elya** and **etd** are used by the power module. The first contains all of the (aggregate) power activities and the second contains the electricity transmission and distribution activity (normally a single activity). If these sets are empty, the distribution of power will be modeled using the standard 'make' specification, i.e. a single CES nest for aggregation. An additional subset, **primElya**, is used post-simulation to calculate primary energy demand. The latter is composed of all combusted fossil fuels and primary electricity production, which excludes thermal power plants to avoid double counting.

The power module allows for a multiple nested CES structure for power aggregation. The example below, Listing F.5, shows the power nesting for a different map file. There are eight power activities (coal, oil, gas, nuclear, hydro, wind, solar, and other renewable), plus the transmission and distribution activity. The eight activities will be mapped to five power bundles (coal, gas, oil, nuclear and renewables), see set **pb**. The power aggregation is based on three nests, see figure 6—that is fully defined with the subsets in the power module of the map file. The subset **elyc** is designed to contain the single electricity commodity that is the output of aggregating electricity output across all electricity activities.

Listing F.5: A power bundle example

---

```

1 set elya(actf) "Power activities" /
2     clp
3     olp
4     gsp
5     nuc
6     hyd
7     wnd
8     sol
9     xel
10    etd
11 / ;

13 set etd(actf) "Electricity transmission and distribution activities" /
14     etd
15 / ;

17 set primElya(elya) "Primary energy power activities" /
18     nuc
19     hyd
20     wnd
21     sol

```

---

```

22     xel
23     etd
24 / ;

26 set pb "Power bundles" /
27     coap "Coal power"
28     gasp "Gas power"
29     oilp "Oil power"
30     nucp "Nuclear power"
31     othp "Other power"
32 / ;

34 set mappow(pb,elya) "Mapping of power activities to power bundles" /
35     coap.clp
36     gasp.gsp
37     oilp.olp
38     nucp.nuc
39     othp.(hyd, wnd, sol, xel)
40 / ;

42 set elyc(commf) "Electricity commodities" /
43     ely
44 / ;

```

---

The subsets **ftr** and **feed** are used respectively by the **acr** and **lvs** activities to define the *ND2* bundle.<sup>10</sup> The former contains the fertilizer commodities and the latter the feed commodities. The subset **iw** defines the commodities that can be bundled with the water factor in the water module. It could contain, for example, the water services commodity (**wtr**) from the GTAP database.

The subset **e** is a critical component of the energy module, see figures 5 and 8. The subset is also linked to the **NRG** set, the subsets **coa**, **oil**, **gas** and **ely** and the mapping set **mape**. Listing F.6 shows a potential configuration using the standard energy commodities in the GTAP database.<sup>11</sup> In the current configuration the set **NRG** and the name of the subsets **coa**, **oil**, **gas** and **ely** cannot be modified by the user. The subsets, however, can be empty. The subset **f** is not used by the model itself [TBC]. It is used at times in post-simulation to evaluate CO<sub>2</sub> emissions from the combustion of fuels. In the original GTAP database, this corresponds to the commodities **coa**, **oil**, **gas**, **p-c**, and **gdt**.

Listing F.6: An energy bundle example

---

```

1 set e(commf) "Energy commodities" /
2     coa
3     oil
4     gas
5     p-c
6
7     ely
8 / ;

10 set NRG "Energy bundles used in model" /
11     coa "Coal"
12     oil "Oil"
13     gas "Gas"
14     ely "Electricity"
15 / ;

17 set coa(NRG) "Coal bundle used in model" /
18     coa "Coal"
19 / ;

```

---

<sup>10</sup> From the point of view of the code, these are used to determine the mapping sets **mapi1** and **mapi2**.

<sup>11</sup> The ordering in the map file is somewhat arbitrary and mostly up to user discretion, unless there are specific dependencies.

---

```

21 set oil(NRG) "Oil bundle used in model" /
22   oil      "Oil"
23 / ;

25 set gas(NRG) "Gas bundle used in model" /
26   gas      "Gas"
27 / ;

29 set ely(NRG) "Electricity bundle used in model" /
30   ely      "Electricity"
31 / ;

33 set mape(NRG,e) "Mapping of energy commodities to energy bundles" /
34   COA.(coa)
35   OIL.(oil, p-c)
36   GAS.(gas)
37   ELY.(ely)
38 / ;

```

---

The mapping set `mapzf` is critical for the rural to urban migration module (if implemented). It would normally correspond to the `mapz` mapping set from the earlier part of the mapping file—but corresponding to the model aggregation, not the data aggregation. If the migration elasticity is set to infinity, the model will assume perfect labor mobility and the `mapzf` mapping set will be ignored.

The top level demand nest for consumer demand is based on a different set of goods than the commodities defined by the set  $i$ . This allows for a 'make' or 'transition' matrix approach to consumer demand specification, see figure 7. The household commodities are specified over the set  $k$ . It is a simple mapping from the supply commodities ( $i$ ) to consumed commodities ( $k$ ). If there is a one-to-one mapping, the consumer make matrix is fully diagonal. At the moment, one of the key purposes of the make matrix is to allow for a consumer energy bundle, with a structure similar to the energy bundle in production, see figure 8. It could however be used to construct other non-diagonal elements. For example crop commodities could be combined into a single consumer crop bundle and the livestock commodities could be combined into a single consumer livestock bundle. The subsequent CES nests would then allocate the top level demand to the relevant components. The subset `fud(k)` can be used by the ENVISAGE model to focus on food demand. Listing F.7 provides an example of a consumer mapping for energy only.<sup>12</sup>

Listing F.7: An example of the consumer 'make' aggregation

---

```

1 set k "Household commodities" /
2   crp  "Cereals"
3   osd  "Oil seeds"
4   xcr  "Other crops"
5   lvs  "Meat and wool"
6   rmk  "Raw milk"
7   frs  "Forestry"
8   fsh  "Fisheries"
9   omn  "Other mining"
10  met  "Meat products"
11  vol  "Vegetable oils"
12  mil  "Dairy products"
13  ofd  "Other foods products"
14  b.t  "Beverages and tobacco products"
15  tex  "Textiles"
16  wap  "Wearing apparel"
17  lea  "Leather products"
18  wdp  "Lumber, paper and paper products"
19  crp  "Chemicals, rubber and plastics"

```

---

<sup>12</sup> Like for the other activities and commodities, the aggregation facility will automatically append the '-k' suffix to consumer labels.



---

```

20     mmn     "Basic metals and minerals"
21     ele     "Electronic equipment"
22     mvh     "Motor vehicles"
23     xma     "Other industry & manufacturing"
24     cns     "Construction"
25     ttp     "Trade and transport"
26     bsv     "Financial and business services"
27     dwe     "Dwellings"
28     xsv     "Other services"
29     nrg     "Energy bundle"
30 / ;

32 set fud(k) "Household food commodities" /
33     crp     "Cereals"
34     osd     "Oil seeds"
35     xcr     "Other crops"
36     lvs     "Meat and wool"
37     rmk     "Raw milk"
38     fsh     "Fisheries"
39     met     "Meat products"
40     vol     "Vegetable oils"
41     mil     "Dairy products"
42     ofd     "Other foods products"
43     b.t     "Beverages and tobacco products"
44 / ;

46 set mapk(commf,k) "Mapping from i to k" /
47     cer.cer
48     osd.osd
49     xcr.xcr
50     lvs.lvs
51     rmk.rmk
52     frs.frs
53     fsh.fsh
54     coa.nrg
55     oil.nrg
56     gas.nrg
57     omn.omn
58     met.met
59     vol.vol
60     mil.mil
61     ofd.ofd
62     b.t.b.t
63     tex.tex
64     wap.wap
65     lea.lea
66     wdp.wdp
67     p.c.nrg
68     crp.crp
69     mmn.mmn
70     ele.ele
71     mvh.mvh
72     xma.xma
73     ely.nrg
74     cns.cns
75     ttp.ttp
76     bsv.bsv
77     dwe.dwe
78     xsv.xsv
79 / ;

```

---

The land module depends on user-determined land bundles, see figure 11. The model allows for a fair amount of user flexibility. There are three nests, but one of the nests allows for a variable number of land bundles. The mapping of the agricultural activities to the land bundles is also determined by the user. Listing F.8 shows an example of a land bundle that corresponds to that

used in the MAGNET model.<sup>13</sup>

The MAGNET specification has only three bundles (L, FCP and COP). We have arranged these into five bundles. The top is split between the `hrt` bundle and all of the rest (that corresponds to the FCP bundle). We assume that the top level CET elasticity in MAGNET,  $\sigma^1$  holds for the top level bundle in ENVISAGE but also for the bundle composed of activities in the `hrt` bundle. This has the same impact as mapping directly the `hrt` activities to the top level bundle as in MAGNET. The `XNLB` bundle will then be composed of the `lvs` and `cer` bundles. MAGNET's CET elasticity,  $\sigma^2$  is then applied to the `XNLB` and `lvs` bundles, which has the same impact as mapping directly the `lvs` activities to the `XNLB` bundle as in MAGNET. Finally, MAGNET's  $\sigma^3$  elasticity is applied to the `cer` bundle. The set `lb0` is used to map the elasticities in the parameter file to the model elasticities—the set is not part of the the model. [This mapping is still under review and users can override these in the model parameter file].

Listing F.8: An example of land configuration

---

```

1  set lb "Land bundles" /
2    hrt  "Horticulture and other crops"
3    lvs  "Livestock and sugar"
4    cer  "Cereals and oil seeds"
5  / ;

7  set lb1(lb) "First land bundle" /
8    hrt  "Horticulture and other crops"
9  / ;

11 set maplb(lb,actf) "Mapping of activities to land bundles" /
12   hrt.v.f
13   htr.ocr
14   lvs.lvs
15   lvs.sug
16   cer.wht
17   cer.gro
18   cer.osd
19 / ;

21 set lb0 "Default land bundles" / lb1*lb3 / ;
22 set maplb0(lb, lb0) "Mapping of land bundles to original" /
23   hrt.lb1
24   lvs.lb2
25   cer.lb3
26 / ;

```

---

The water module has three basic bundle types:

1. Water bundles that have activities mapped to them—for example irrigation water in crop activities
2. Water bundles that are linked to aggregate output indices—for example industrial water use
3. Exogenous bundles not linked directly to the economy—for example ground water recharge or water for environmental services.

These fundamental water bundles are structured in a nested CET nest for which users have significant control. In the example in Listing F.9 there is one bundle of the first type (crops), three bundles of the second type (livestock, municipal and industrial) and two bundles of the third type (environmental services and ground water recharge). There are two intermediate bundles that form

---

<sup>13</sup> [edepot.wur.nl/310764](http://edepot.wur.nl/310764), page 74).

the top level CET nest—agriculture and non-agriculture, see figure 13. The set `wbnd` defines all of the possible bundles. The set `wbnd1` defines the top level bundles—in principle there could be more than two. The set `wbnd2` defines the second level bundles. The set `wbndEx` defines the exogenous bundles. These will be subtracted from total water supply. The mapping set `mapw1` maps fundamental bundles to the top-level intermediate water bundles. The mapping set `mapw2` maps all activities to one of the fundamental bundles. The subset `wbnda` indicates which water bundles are mapped directly to activity-based water demand. With the current water database, only irrigated crops use water directly. The subset `wbndi` indicates which water bundles are mapped to output indices. The elasticities for the water module need to be provided in the user-built parameter file—see below.

Listing F.9: An example of water configuration

---

```

1  set wbnd "Aggregate water markets" /
2    agr    "Agriculture"
3    nag    "Non-agriculture"
4    env    "Environmental services"
5    grd    "Ground water recharge"
6    crp    "Crops"
7    lvs    "Livestock"
8    ind    "Industrial use"
9    mun    "Municipal use"
10 / ;

12 set wbnd1(wbnd) "Top level water markets" /
13   agr    "Agriculture"
14   nag    "Non-agriculture"
15 / ;

17 set wbnd2(wbnd) "Second level water markets" /
18   crp    "Crops"
19   lvs    "Livestock"
20   ind    "Industrial use"
21   mun    "Municipal use"
22 / ;

24 set wbndEx(wbnd) "Exogenous water markets" /
25   env    "Environmental services"
26   grd    "Ground water recharge"
27 / ;

29 set mapw1(wbnd,wbnd) "Mapping of first level water bundles" /
30   agr.(crp,lvs)
31   nag.(mun,ind)
32 / ;

34 set mapw2(wbnd,actf) "Mapping of second level water bundle" ;
35 mapw2("crp",acr) = yes ;
36 mapw2("lvs",alv) = yes ;
37 mapw2("ind",man) = yes ;
38 mapw2("mun",actf)$ (not (acr(actf) or alv(actf) or man(actf))) = yes ;

40 set wbnda(wbnd) "Water bundles mapped one-to-one to activities" /
41   crp    "Crops"
42 / ;

44 set wbndi(wbnd) "Water bundles mapped to aggregate output" /
45   lvs    "Livestock"
46   ind    "Industrial use"
47   mun    "Municipal use"
48 / ;

```

---

The remaining section of the map file relates to the use of the IIASA- or GIDD-based education

projections to shape the growth of skilled and unskilled labor. This is not fundamental to the model, but is part of defining the dynamic scenario.

One subset defines skilled labor, **skl**. This is different from the **ul** and **sl** subsets that are used to allocate labor between the *LAB1* and *LAB2* bundles. The purpose here is only to drive the relative growth of unskilled and skilled labor. The mapping set, **educMap**, determines the mapping between labor skills and the education projections. There are two different education projections. The IIASA education projections have four classifications: none, primary, secondary and tertiary (labeled **ENONE**, **EPRIM**, **ESECN** and **ETERT** in the SSP database, but labeled **ELEV0\*ELEV3** in the scenario file.) The GIDD database has three classifications: 0-6 years of education, 6-9 years of education and 9 or more years (labeled **EDUC0\_6**, **EDUC6\_9**, and **EDUC9UP** in the GIDD database, but labeled **ELEV0\*ELEV2** in the scenario file.) If using the IIASA database, all four education levels should be mapped to the skills. If using the GIDD database, only three of the levels are active. The mappings can vary across regions. For example, in high-income countries, skilled workers might be only tertiary and above. In developing countries, skilled workers may also include workers with a secondary education.

## Running the aggregation

The aggregation is coded in GAMS and can be run in the GAMS IDE or from a Windows command console. For the latter, the command line is:

```
gams aggGTAP --basename=[basename] --ifAlt=[OFF|ON] --model=[GTAP|ENV]
      --ifCSV=[0|1] --ifAggTrade=[0|1]
```

The aggregation routine is named **AggGTAP.gms**. The user needs to provide the base name of the project, for example **10x10**. The aggregation facility will prepare additional output if the user wishes to use **AlterTax**. In that case, the **ifAlt** setting should be set to **ON**, otherwise it should be set to **OFF**. Aggregation is handled somewhat differently for the GTAP and ENVISAGE models. The **model** option should take the value **GTAP** for the GTAP model, and **ENV** for the ENVISAGE model. The **ifCSV** option takes either the value 0 or 1. If it is set to 1, the aggregation facility will output the SAM, energy and emissions data in a CSV formatted file that can be uploaded into Excel. The **ifAggTrade** option takes either the value 0 or 1. If it is set to 1, the output SAM will collapse bilateral trade to a single trade account, otherwise the SAM will have the full bilateral trade (on both the import and export side).

If the aggregation is intended for the ENVISAGE Model, the *aggGTAP* routine will invoke the *convertLabel* program that will convert the labels in the parameter file to the labels needed for the model.

### F.1.4 Filtering

The filtering algorithm was developed by Tom Rutherford to remove small value flows from the aggregated database ([Lanz and Rutherford \(2016\)](#)). The distributed version has been modified by Wolfgang Britz to handle some additional features. The user needs to prepare a very short include file with the algorithms's options (**[basename]Flt.gms**). Listing [F.10](#) is an example of the filter options file.

## The filter options file

The consecutively applied filtering and rebalancing approach is an extension of the method and code developed by Tom Rutherford.<sup>14</sup> It deletes components of the SAM depending on their shares on specific totals, according to the 'Relative tolerance' (**relTol**):

- Domestic and imported intermediate demand of a commodity are dropped relative to its total output
- Private/government/investment domestic respectively import demand of a commodity are dropped relative to total Private/government/investment domestic respectively import demand
- Trade flows of a product are dropped if both shares on total exports of that product and its exporter and on imports of that product and its importer are below the relative threshold
- Production is dropped if net production of a commodity, i.e. after intermediate use of that commodity in its own production is deducted, is below the relative threshold with regard to total net production

Listing F.10: An example of a filter file

---

```

1  scalars
2  ifKeepIntermediateConsumption / 1 /
3  ifKeepPrivateconsumption      / 1 /
4  ifKeepGovernmentconsumption   / 1 /
5  ifKeepInvestments              / 1 /
6  ifGDPKeep                     / 1 /
7  ifKeepFactorincomeplusbop     / 1 /
8  ifAdjDepr                     / 1 /
9  abstol                        / 1e-10 /
10 relTol                       / 0.005 /
11 relTolRed                    / 1e-6 /
12 nsteps                       / 5 /
13 minNumTransactions           / 50000 /
14 ;

16 file log / %baseName%flt.log / ;
17 put log ;

19 $ontext
20 $setglobal excRegs
21 $setglobal excSecs
22 $offtext

24 $setglobal excSecs "sol, wnd, xel"
25 $setGlobal excCombined 1

```

---

The absolute tolerance level, **absTol**, deletes any trade flow value that is below the tolerance level in absolute terms. With a value of 1.E-10, that preliminary deletion step is skipped. It is generally not recommended to use absolute deletion thresholds above 1.E-6 in combination with rebalancing as the subsequent relative thresholds will anyhow apply more refined rules.

The filtering process imposes restrictions which should maintain the regional SAMs balanced. Additional constraints ensure that production activities require added value and intermediate inputs, if not already otherwise found in the data base. As filtering systematically removes elements

---

<sup>14</sup> The following explanations are adapted from Britz's documentation. See [http://www.ilr.uni-bonn.de/em/rsrch/cgebox/cgebox\\_GUI.pdf](http://www.ilr.uni-bonn.de/em/rsrch/cgebox/cgebox_GUI.pdf).

from the SAM and the trade matrices, the process implies without further corrections shrinking the economies. During rebalancing, the algorithm can therefore add penalties for deviations from the following aggregate transactions:

- Intermediate consumption
- Private consumption
- Public consumption
- Investment
- Factor income plus BOP
- GDP

The imposition of these penalties is driven by the relevant flags in the filter options file. By adding these penalties terms, the non-deleted entries (and thus most important transactions) tend to be scaled upwards. It is generally recommended to use these penalties terms. The code will also scale all non-deleted trade flows to approximately maintain the total volume of international trade and related international transport margins.

The absolute and relative thresholds are stepwise enforced. For the first few steps, exponential increases are used, starting with minus half the number of steps. For six steps, to give an example, the first thresholds applied will be the 1.E-3 of the final one, next 1.E-2 and finally 10%. The remaining steps will use equal linear increases between 10% and the desired final ones. Once the final thresholds are active, filtering is still applied several times until no small values are found any longer. The code should ensure that the resulting transactions are still fully consistent with each other, i.e. both the resulting trade matrices and the SAMs are balanced. The changes imposed by filtering and subsequent balancing are stored in the log file. Inspecting how the stepwise enforcement of the thresholds impacts on the number of non-zero items can inform on an appropriate level for tolerances to be used.

The SAMs used during filtering are—as in the GTAP database—defined in Million dollars. An absolute threshold of 1.E-6 will hence delete any economic transactions worth a single dollar or less. In SAMs with high regional and sectoral detail, even such tiny transactions might make up to 10% of the non-zero entries. Increasing the threshold to \$1000 might remove 1/4 or more of all non-empty transactions. Similar results are found from using relative tolerances of 0.001%.

Thanks to balancing, also rather dis-aggregated versions of the model with large number of sectors and regions can be used. The biggest impact of the filtering is typically on transactions related to bilateral trade flows. Here, often 50% or more of the flows account for only 1% of the total value of these transactions. Thus, tiny changes in the relative tolerance can have a considerable impact on the number of deleted transaction, and one might need to experiment with settings in the range around 1.E-1 to 1.E-4 to find a compromise between sparsity and the implied changes on structure of the economy. For very large data sets (e.g. a 1:1 version) filtering thresholds above 1% might be needed to yield reasonable model sizes. The user can additionally define a minimum number of transactions to be kept, which reduces the need to experiment with different thresholds as the filtering process will stop once less than the desired number of transactions is reached. Tests with the model have shown that the model in full resolution of the GTAP 8.1 data base without filtering, i.e. 57 sectors and 134 regions, can be solved in partial trade liberalization scenarios, solution has failed with other shocks on models with more than 400,000 transactions, especially if the global bank mechanism active. A close look at the filtering statistics is recommended, to avoid

sharp impacts on the structure of the economy. A more detailed discussion on the relation between model dis-aggregation, filtering, solution behavior and simulated welfare impacts is provided in [Britz and van der Mensbrugghe \(2016\)](#).

**Special treatment for specific regions and sectors** When building a data base for a project, it might be desirable to apply less aggressive filtering thresholds for specific regions and/or sectors in the focus of the application. The algorithm therefore allows defining lists of regions/sectors with accompanying specific thresholds. The codes for regions/sectors needs to be inputted in the global defines called `excRegs` and `excSecs`. 'Reduced thresholds only in combination' will apply the different threshold only to the intersection of the inputted regions and sectors (`excCombined`), otherwise, all regions and sectors inputted will be receive different thresholds. Take an example where you enter for regions `xoc` and for sectors `pdr`. If 'Reduced thresholds only in combination' is NOT switched on, all transactions of the region `xoc` and all transactions for the sector `pdr` will be treated differently. If the 'Reduced thresholds only in combination' is active, only the transaction relating both to `pdr` and the region `xoc` are exemptions. However, filtering for the remaining sectors/regions has still an impact on these exemptions. For example, if production of a sector in a region is dropped, the related export flows need to be dropped as well, affecting potentially transactions in regions and sectors where tighter thresholds are used. Tests have however indicated that very few transactions are lost in regions/sector where stringent thresholds are applied as long as the overall filtering thresholds are not too aggressive.

**Diagnostics** The filter listing file (`filter.lst`) and the log file (`[basename]flt.log`) provide a number of useful diagnostics regarding the filtering process. Users can also load the resulting 'CSV' file that contains the SAM and energy/emission values both pre- and post-filtering.

### Running the filter program

The filter routine is coded in GAMS and can be run in the GAMS IDE or from a Windows command console. For the latter, the command line is:

```
gams filter --basename=[basename] --ifCSV=[0|1] --ifAggTrade=[0|1]
```

The filter routine is named `filter.gms`. The user needs to provide the base name of the project, for example 10x10. The `ifCSV` option takes either the value 0 or 1. If it is set to 1, the aggregation facility will output the SAM, energy and emissions data in a CSV formatted file that can be uploaded into Excel. The `ifAggTrade` option takes either the value 0 or 1. If it is set to 1, the output SAM will collapse bilateral trade to a single trade account, otherwise the SAM will have the full bilateral trade (on both the import and export side).

#### F.1.5 Altering a database

The third module in the data preparation routine is called `Altertax`, first developed by [Malcolm \(1998\)](#) for the GTAP model. The `Altertax` module is optional. It allows the user to make changes to the structure of the database that minimizes the distortions from the original database. It is typically used to change tax rates, for example import tariffs. The distributed version of `Altertax` relies on a version of the GTAP model written in GAMS.<sup>15</sup> The main thrust of `Altertax` is to

<sup>15</sup> The GAMS-based GTAP model has been developed by GTAP staff and documentation is available upon request. It differs from the Rutherford *GTAPinGAMS* model ([Lanz and Rutherford \(2016\)](#)) as it is intended to exactly replicate the specification of the standard GTAP model in GEMPACK ([Corong et al. \(2017\)](#) and [Hertel \(1997\)](#)).

convert most of the model elasticities to 1 thereby all CES functions are essentially converted to Cobb-Douglas functional forms. The latter have the property of conserving value shares. In addition, the Altermat version of the GTAP model assumes a Cobb-Douglas utility function for private consumption. A fixed current account is also imposed for the balance of payment closure. The overrides are provided in the `AlterTax.gms` file.

### Option file

If running Altermat, the user needs to prepare an option file that specifies the nature of the alteration—it is to be named `[basename]Alt.gms`. The file can be empty, in which case Altermat would simply re-balance the initial database—though this will already have been done in the filter routine if it was used. Listing F.11 shows a very simple change: a rise of 40% on the import tariff of motor vehicles in Argentina. One could of course overlay the entire import tariff schedule for Argentina in this file—including any bilateral dimensions. Note that the activity and commodity labels will need the appropriate suffix.

Listing F.11: An example of an Altermat shock

---

```
2  imptx.fx(s, "mvh-c", "arg", tsim) = 1.4*imptx.l(s, "mvh-c", "arg", tsim) ;
```

---

A more complicated, though artificial example is provided in Listing F.12. This example shows how to phase in the shock to ease numerical convergence. The shock is a 50% cut in all tariffs. The user provides the number of iterations to phase in the shocks, say for example 4. The iteration count is first checked to see if it is 1, in which case there is no iteration phase and the shock is imposed in one shot. If the iteration count is greater than 1, the shock will be phased in. The `solver` needs to be invoked for the first  $n - 1$  iterations, the  $n^{th}$  iteration will be solved in the standard (and final) invocation of the solver.

Listing F.12: An example of an Altermat file

---

```
1  *
2  *  Altermat shock
3  *
4  *  If new policy needs to be phased in, the basic implementation is
5  *
6  *      p(it) = pfinal*(it/n) + pinitial*(1-it/n)
7  *
8  *  See below for an example
9  *
10 *  Set the number of iterations on the command line, for example —niter=4

12 *  Cut initial tariffs by 50%
13 *
14 *
15 if(niter(tsim) eq 1,
16
17 *  No phase in of cuts
18
19   imptx.fx(r,i,rp,tsim) = 0.5*imptx.l(r,i,rp,"base") ;
20
21 else
22
23 *  Phase in the shock
24
25   for(iter=1 to niter(tsim),
26     imptx.fx(r,i,rp,tsim) = 0.5*imptx.l(r,i,rp,"base") * (iter/niter(tsim))
27     + imptx.l(r,i,rp,"Base") * (1 - iter/niter(tsim)) ;
28   if(iter < niter(tsim),
```

---



---

```

29         $$batinclude "solve.gms" gtap
30     ) ;
31 ) ;
32 ) ;

```

---

## Running Altertax

The Altertax routine is coded in GAMS and can be run in the GAMS IDE or from a Windows command console. For the latter, the command line is:

```
gams AlterTax --BaseName=[basename] --niter=1 --ifCSV=[0|1] -idir=GTAPModel
```

The filter routine is named `Altertax.gms`. The user needs to provide the base name of the project, for example `10x10`. The user needs to provide the number of iterations for a phase-in of the shock. For most modest shocks, a value of 1 should be sufficient. Note that it is up to the user to make sure the Altertax option file contains the necessary code for the phase in of the shock. The `ifCSV` option takes either the value 0 or 1. If it is set to 1, the aggregation facility will output the SAM, energy and emissions data in a CSV formatted file that can be uploaded into Excel. The `idir` option provides a pointer to the folder containing the GTAP model code. Under default configurations the model code is a sub-folder in the *Data* directory with the name *GTAPModel*.

If the `ifCSV` option is set to 1, the SAM and energy/emissions data will be output in a CSV-formatted file for input into Excel. The CSV file will have a 'time' dimension that takes three values. 'Year' equal to 1 corresponds to the base year initialization of model variables and parameters and should correspond to the input database. 'Year' 2 corresponds to a first simulation with no shock. It should re-produce the base data as well, i.e. the results in 'Year' 1 and 2 should be identical to within a very small tolerance level. 'Year' equal to 3 will correspond to the post-shock structure of the database and comparison with either 'Year' 1 or 2 will highlight deviations from the initial database.

### F.1.6 Integrated command file

The distribution is delivered with a Windows command file, `makeData.cmd`, that can be used to run all of the modules in sequence. To run the command, type the following in a Windows console:

```
makeData [baseName] [-ifFilter] [-ifAlt] [-ifEnv]
```

This will invoke the `makeData` command file. There are three options. To invoke the filter module, enter the `-ifFilter` option on the command line. To invoke the Altertax module, enter the `-ifAlt` option on the command line. To prepare an aggregation for the ENVISAGE Model, enter the `-ifEnv` option on the command line. N.B. We have made every attempt to make the `makeData` command file as robust as possible—but coding a Windows command file is not for the faint-hearted. We would be happy to have any feedback on its robustness and usefulness.

The `makeData` command file automates the copying and moving of files. In a first step, it will create a folder with the name of the aggregation code, `[basename]` if it does not exist. It will also create the four sub-folders: `Agg`, `Flt`, `Alt` and `Fn1`. The aggregation routine creates 10 GDX files (and

optionally the CSV file). The filter routine only modifies the five data files: `[basename]Dat.gdx`, `[basename]Vole.gdx`, `[basename]Emiss.gdx`, `[basename]NC02.gdx` and `[basename]Wages.gdx`. The `makeData` command file copies the other (non-modified) files from the `Agg` folder to the `Flt` folder. Similarly for `Altertax`, the command file will copy the non-modified files from the `Flt` folder to the `Alt` folder. If the respective routines are not invoked, the command file will copy ALL files from one folder to the next. In a final step, all the GDX files in the `Alt` folder will be copied to the `Fn1` folder from where the user can copy the final data and parameter files to the working folder for subsequent simulations.

## F.2 Model Simulations

### F.2.1 Introduction

We introduce in this section how to run ENVISAGE model simulations. There are three fundamental types of simulations:

- **Comparative static.** Comparative static involves introducing a shock to the benchmark database with no dynamic elements—such as factor accumulation, technology and preference changes, etc. (of course a comparative static shock could include a shock to one of these elements). The key difference between the comparative static and dynamic version of the model is the specification of the capital account as described in the model description. Capital markets close with a CET transformation function in the comparative static version, which also has no vintage capital. In the dynamic version, installed capital is assumed to be only partially mobile across activities and the model explicitly incorporates vintages.
- **Recursive dynamic with baseline calibration.** This version of the model is dynamic and the model is used to determine some model parameters subject to some targets. For example, GDP growth may be exogenous and labor productivity is calibrated to achieve a given growth target, and/or the investment to GDP ratio is targeted and the household savings rate is adjusted to meet the target, etc.
- **Recursive dynamic shock scenario with pre-calibrated dynamic parameters.** This type of dynamic scenario will use the results from the baseline calibration scenario for some of the underlying dynamic trends, for example labor productivity. In the absence of a shock, this type of simulation should re-produce the baseline scenario.

There is no single best way to run model simulations—in either comparative static or dynamic mode as GAMS provides great flexibility. We introduce herein a unified system that can work for all three types of simulations. This system is useful because it highlights the differences across the simulation types. It does have some potential drawbacks that will be highlighted below.

### F.2.2 Preliminaries

The model code is composed of a series of GAMS file, this helps with modularizing the code. Typically, all of the core model code will be available in a single folder and the user will set the `idir` option in GAMS to the location of the core model code, say for example in the *Model* directory.

Table F.6 is a list of the files that constitute the core model code. The model's declarations and equation specification are contained in the `model.gms` file.

Table F.6: Distributed model files

<i>File name</i>	<i>Description</i>
cal.gms	Calibration of model parameters
closure.gms	Default closure assumptions
compScen.gms	'Scenario' file for comparative static simulations
getData.gms	File that reads the GDX-based input files.
init.gms	Initialization of model variables
initScen.gms	Scenario file for dynamic simulations
InitVar.gms	Inter-period initialization of model variables
initVint.gms	Second period initialization of vintage volume variables
iterloop.gms	Inter-period code
miscDat.gms	Currently this file contains only an energy conversion table. It is used post-simulation.
model.gms	Core model specification
postsim.gms	Post-simulation statements—mostly creation of output CSV files
recal.gms	Inter-period core code to update vintage technology parameters and potentially other dynamically calibrated parameters such as the Armington preference parameters
recalnnn.gms	Code that updates technology parameters such as <b>and1</b>
recalnnt.gms	Code that updates technology parameters such as <b>aland</b>
recalnrg.gms	Code that updates technology parameters such as <b>aNRG</b>
recalvat.gms	Code that updates technology parameters such as <b>axp</b>
recalvnn.gms	Code that updates technology parameters such as <b>ava</b>
recalvnt.gms	Code that updates technology parameters such as <b>ak</b>
recalxanrg.gms	Code that updates technology parameters such as <b>aeio</b> with no energy nesting
recalxanrgn.gms	Code that updates technology parameters such as <b>aeio</b> with energy nesting
sam.gms	Code that writes out the simulation SAMs in CSV format
SaveParm.gms	Code that writes out key parameters
scale.gms	Code that scales model variables/equations. Currently not used.
solve.gms	Code that invokes the solver.

Figure 19 provides a schematic view of a simulation. In the current setup, the user prepares a project specific file with the name `[basename]Opt.gms`. This file contains the common options and statements across all simulations in the project. It is designed to handle both comparative static and dynamic simulations. Many of the global options can be over-ridden in subsequent simulation files (for example `runSim.gms`), however, it is normally not possible to override set definitions such as time. The first step in a simulation file, such as `runSim.gms` is to read the options file. The options file will also load a number of the core GAMS code and some user files. The 'Sets' file, which is user-based, is generated by the aggregation facility and normally requires no further input. The parameter file, described below does need user input for each project. The core code that is read includes the model specification (`model.gms`), reading of GDX-based input data files (`getData.gms`)<sup>16</sup>, initialization of model variables (`init.gms`), model calibration (`cal.gms`) and the default closure rules (`closure.gms`). After defining the default closure the model is ready to be solved.

The model is solved over time—even in the case of comparative static simulations—albeit skipping the first period, which will contain the initialized and calibrated solution using the base

<sup>16</sup> [1-Dec-2018] Reading of the base data files was separated from the model initialization routine. This allows the user to make direct changes to the database before variable initialization and parameter calibration.

data. A number of statements are executed at the beginning of each period that initializes variables, may recalibrate technology parameter (to be described), updates time-based shocks in the case of dynamic scenarios and introduces shocks for both comparative static and dynamic simulations. Once all of the period updating is finished, the solver is invoked. The simulation file will loop over all time periods—as long as the model converges. After looping over all periods, model results will be saved—optionally in a CSV file and always in GDX format.

In a typical application, the user will create a project directory with the base name of the project. The simulation files will be contained in that directory. The distribution comes with several files that provide examples of how to run both comparative static and dynamic simulations. Beyond the aggregated data files, the user needs to prepare a file with the model parameters. This file can be fairly generic if based on an aggregation of one of the existing parameter files. However, there are a number of options that users need to define to complete the model specification—such as the labor market closure.

### The user-based parameter file

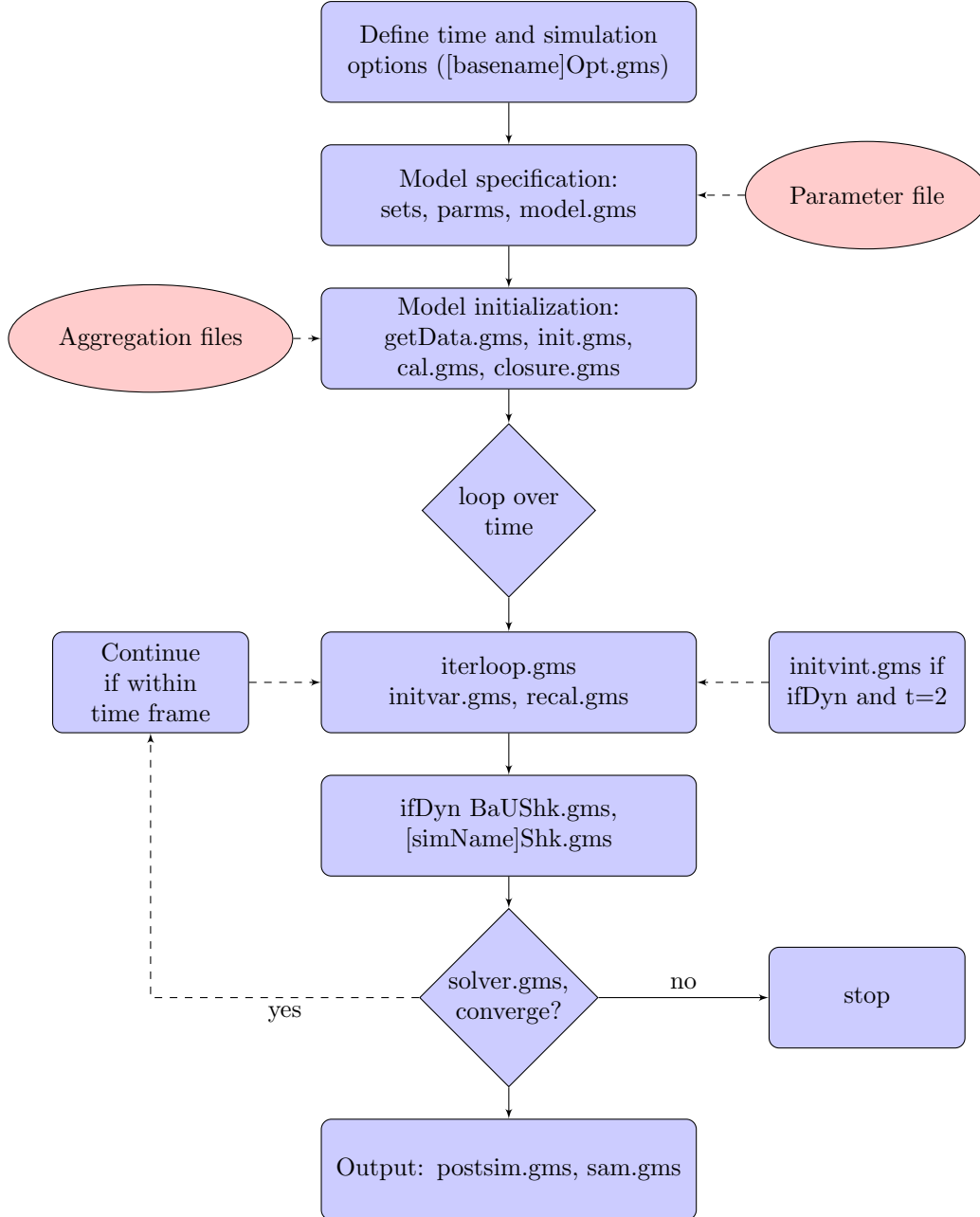
While most of the inputs to the model are prepared by the aggregation facility, it is up to the user to prepare the file with project-specific model parameters. The file has the name `[basename]Prm.gms`. There is a generic parameter file that simply initializes all model parameters to those generated by the aggregation facility. These can be overridden—after the ones from the aggregation facility are loaded, but before the initialization of the model parameters. The read-in parameters will have a suffix of '0'. The model parameters have no suffix.

The aggregated parameters reflect values used by the OECD's Environment Directorate Env-Linkages Model (Chateau et al. (2014)). The **Envisage** Model has some new features not fully reflected in the original Env-Linkages Model, such as the power and water module. The file does include some initial levels for the power elasticities, but [for the moment] does not include the elasticities for the water module. These must be entered by the user if the water module is active. Another critical set of elasticities is the natural resource supply elasticities. [NEW] It is the user's responsibility to enter these elasticities (for the original model aggregation). The user must enter a pair of elasticities for each natural resource and for all regions. The parameter to be initialized is `etanrfx0` that takes three indices: region, activity and a special one that takes the values of 'lo' and 'hi'. This latter is represented by the set `lh`. These refer to the supply elasticity to apply as a function of market conditions. The 'lo' supply elasticity is to be used for a market under contraction. The 'up' supply elasticity is to be used for an expanding market. Intuition suggests that the 'lo' elasticity will be greater than the 'hi' elasticity, i.e. it is easier to contract supply than to expand it. These should be carefully scrutinized and potentially adjusted during baseline simulation runs.

[NEW] The user must enter the investment allocation elasticities to be used with the flexible capital flow specification. The initial elasticities are to be initialized with the parameter `epsRor0`. The default value for the GTAP model is 10.

[NEW] There are a number of parameters to be entered by the user for the USAGE-inspired capital account closure. The user must enter the bounds on the growth of capital in any given year. These are entered in the parameters `'grKMin0'` and `'grKMax0'`. Normal trend growth is entered in the parameter `'grKTrend0'`. The curvature of the logistic function at the equilibrium point of the capital supply schedule is entered in the parameter `'chigrK0'`. The user must also enter the so-called normal rate of return, from which the deviations are evaluated. This is entered in the parameter `'RoRn0'`.

The comparative static version of the model uses a weighted average of the vintage-dependent

Figure 19: **Simulation flow chart**

elasticities. Users are free to change the weights, where the default values are 0.8 for *Old* and 0.2 for *New*.

There are a number of other key assumptions that are included in the parameter file. The first relates to the labor market assumptions. The user must initialize the matrix of options in the parameter `labHyp`. It has a regional and skill index and 11 attributes described below.

1. `omegam`. This determines labor market segmentation. A finite value will implement labor market segmentation—typically defined over agricultural versus non-agricultural activities. Perfect labor mobility is implemented if the value is infinity (`INF`). If it is finite, this will be the value of the labor mobility elasticity in equation (G-1),  $\omega^m$ , the elasticity of rural to urban migration with respect to the expected urban wage premium.
2. `migr0`. This determines the initial level of migration as a percent of the rural labor force. For example, if it is set to 1.0, the initial level of rural to urban migration will be set to 1% of the rural labor force. It is ignored if the migration elasticity is infinity.
3. `uezRur0`, `uezUrb0`. These parameters provide the initial unemployment rate (in percent) in respectively the rural and urban labor markets. In the case of full labor mobility, only the urban unemployment rate is active and the rural rate is ignored.
4. `uezminzRur0`, `uezminzUrb0`. These parameters provide the lower bound for the unemployment rate (in percent) in respectively the rural and urban labor markets. The rural rate is ignored in the case of full labor mobility. An error will be issued if the initial unemployment rate is less than the minimum unemployment rate.
5. `resWageRur0`, `resWageUrb0`. These parameters determine the initial level of the reservation wage with respect to the prevailing wage. A value of 1 indicates that the reservation wage is binding and the unemployment rate is greater than (or equal to) the minimum unemployment rate. A value less than one indicates the distance between the initial equilibrium wage and the reservation wage. The rural reservation wage is ignored in the case of full labor mobility. A value of `NA` indicates full employment (at all times). An error will be issued if the value is greater than 1.
6. `omegarwg` `omegarwue` `omegarwp`. The reservation wage is a function of three indicators—the growth of per capita GDP, the unemployment rate and the CPI. The elasticity is positive for income growth and the CPI. The reservation wage is negatively related to the unemployment rate, i.e. an increase in unemployment would tend to dampen the reservation wage. To avoid problems with division by zero, the relationship between the reservation wage and unemployment is converted to a relationship between the reservation wage and the employment rate, i.e.  $(1 - UE)$ . Thus if the elasticity of the reservation wage is  $\omega^{ue}$ , the elasticity of the reservation wage with respect to the employment rate is  $-\omega^{ue} (1 - UE) / UE$ . The user is expected to enter the elasticity with respect to the employment rate and it should be a positive number.

There are three additional parameters that allow for control of the dynamics of the capital market. The first, `cap_out_Ratio0`, allows for overriding the initial capital to GDP ratio. The initial capital stock is an estimate derived from the Penn World Tables.<sup>17</sup> For some countries and/or regional aggregations the estimates can lead to significant jumping off problems in the

<sup>17</sup> <http://www.rug.nl/ggdc/productivity/pwt/>

initial years. A value of **NA** will ignore the override. The parameter is entered as a multiplicative adjustment of the initial capital stock. The default value is 1.0. The second is the depreciation rate, **deprT**, which for some rapidly growing countries is too low. The default GTAP value is 4%. The third is an investment target, **invTarget0**. Users can input this for particular years, e.g. 2030 and 2050. Code in the 'Opt' file will translate these targets into linear changes for the investment target, for example between 2011 and 2030 and then between 2030 and 2050, in order to smooth the path between target years.

The final section allows the user to implement 'twists' to the Armington preference parameters. There are three sets of twist parameters and one twist-related regional mapping set. The top-level Armington twists require two different parameters depending on the value of **ArmFlag**. In the case of national sourcing, i.e. when **ArmFlag** = 0, the relevant twist parameter is **tw1**, which is region-, commodity and time-specific. With agent-based sourcing, i.e. when **ArmFlag** = 1, the relevant twist parameter is **tw1**, that is in addition agent-specific. A positive value, for example 0.02, would lead to a change in the ratio of imports to domestic demand of 2 percent—assuming constant prices. The parameter **tw2** is applied to the second level nesting. The *r* index refers to the importing country. The twist can be applied to one single region, or a group of regions. The user designates the targeted exporters in the regional mapping set **rtwtgt**. The first regional index in the mapping set refers to the targeted exporters and the second regional index refers to the importer. For example, to increase the import shares of China, Japan and Korea in the U.S. using the twist, the relevant mapping is **(chn,jpn,kor).usa**. Note that in this implementation, the same twist is applied uniformly across the targeted exporters.

### F.2.3 Common options

The first set of global options are entered as definitions through GAMS' **\$setGlobal** statement. The following enumerates the options:

1. **wDir**. By default this is set to the active directory.
2. **SSPMOD**. This option defines the choice of the economic model used for the SSP GDP projections. There are currently two choices: **OECD** and **IIASA**.<sup>18</sup>
3. **SSPSCEN**. This option defines which SSP to use for the GDP projections (and will be combined with the model choice). There are five valid options: **SSP1-SSP5**.
4. **LABSCEN**. This option influences the relative growth of skilled versus unskilled labor. If it is set to one of the SSPs, the education profiles from the SSPs—with the user mappings of skilled workers to education levels (**educMap**) together with the **skLabgrwgt** parameter (see below)—will determine the growth rate of skilled workers. The growth of unskilled workers will be determined residually so that the overall growth of workers matches the growth of the working age population. If this parameter is not equal to one of the SSPs, both skilled and unskilled workers will grow at the same rate as the working age population, i.e. no differentiation of growth by skill.
5. **POPSCEN**. This option defines which population projection to use. There are ten valid options. The IIASA SSP population projections are labeled **SSP1-SSP5**. The scenario file also includes four UN population projections: **UNMED2010**, **UNMED2012**, **UNMED2015** and **UNMED2017**.

<sup>18</sup> The PIK projections have not been processed since they reflect a fixed aggregation of 32 regions, whereas the OECD and IIASA projections were done for most countries.



Finally, the World Bank's GIDD projections are available with the label **GIDD**. The latter is intended to duplicate UNMED2015. It is only available through 2050.

6. **OVERLAYPOP**. This option allows to replace GTAP's base year population level with that from the scenario database. In any case, only the growth rates from the population projections are used from the initial base year level.<sup>19</sup>
7. **TASS**. This option determines the specification for the aggregate land supply curve. Valid options are: **KELAS** for iso-elastic supply, **LOGIST** for logistic supply curve, **HYPERB** for hyperbola supply curve, and **INFTY** for infinitely elastic (i.e. horizontal) supply curve.
8. **WASS**. This option determines the specification for the aggregate water supply curve. It has the same valid options as the aggregate land supply curve.
9. **utility**. This option determines the specification of the household utility function. Five options are available: **CD** for the Cobb-Douglas utility function, **LES** for the linear expenditure system, **ELES** for the extended linear expenditure system, **AIDADS** for the AIDADS demand system, and **CDE** for the Constant-difference-in-elasticity demand system. The **CD**, **LES** and **AIDADS** are new options. The **LES** is calibrated to the Frisch parameter, which is currently parameterized in the calibration routine (**cal.gms**), but it would be preferable to have this entered as user input. **AIDADS** is coded, but there is no calibration. It has been tested assuming the **LES** as one of the special cases of **AIDADS**. The Cobb-Douglas simply uses the base year budget shares to calibrate the marginal budget shares and the subsistence minima are set to zero.
10. **NRITER**. This option controls the number of iterations when running the model convergence in single country mode. As this has not been tested with the latest version of the model, it is best to set this option to 0.
11. **savfFlag**. This option controls the capital account closure. Three options are currently available. A value of **capFix** uses the fixed capital account closure rule. A value of **capFlexGTAP** uses the flexible capital account closure rule of the GTAP model. This closure allocates global savings in order to equalize expected returns across regions. A value of **capRFix** fixes the ratio of savings relative to nominal GDP. A value of **capFlexUSAGE** uses the flexible capital account closure rule inspired by the USAGE model. This closure allocates global savings as a function in deviations of the regional rate of return from the 'normal' rate of return.
12. **intRate**. The user sets this option to the desired value for the global real interest rate. The default value is 5% (i.e. 0.05). Note that this parameter is currently only used by the USAGE capital account closure.
13. **costCurve**. Cost curves are used to implement downward trends of new technologies (e.g. renewable electricity). These can be implemented using either a hyperbolic (**HYPERB**) or with a logistic function (**LOGIST**).
14. **NTMFlag**. This flag takes the value of 1 to implement the non-tariff measure module, or 0 otherwise.

<sup>19</sup> There is currently an inconsistency in the population levels in the SSP database and GTAP. In the former, population is in levels and in the latter it is millions. The inputs from the SSP database are currently scaled by a million to make them consistent with the GTAP units. The population scale factor can be used to improve the population scale on an as-need basis.



The remaining global options are entered as scalars. The first set provide the context for the model: comparative static and recursive dynamic with or without dynamic calibration.

1. **ifDyn**. This option takes two values. A value of 0 indicates a comparative static model. A value of 1 indicates a dynamic model. In the standard package, this will be set automatically when the simulation is invoked and the user sets the **simType** option. Valid values are **CompStat** and **RcvDyn**.
2. **ifCal**. This option takes two values and is only meaningful for dynamic scenarios. A value of 0 runs a dynamic scenario with pre-calibrated trends (generated by a baseline). A value of 1 runs a dynamic calibration scenario where specific dynamic trends are calibrated to exogenous indicators such as GDP. This value will be set in dynamic variables by setting the **ifCal** option when invoking the simulation.
3. **ifVint**. This options takes the value 0 or 1. It is set automatically when the simulation is invoked by the user. A value of 1 implements vintage capital, which is valid only for dynamic simulations. It should be set to 0 for comparative static simulations.

The remaining global options are typically invariant across simulations. The following list enumerates the options.

1. **inScale**. Scale factor for input data. The input SAM is typically in millions. A scale of  $10^{-6}$  has been found to be a useful scaling factor.
2. **outScale**. Scale factor for output data. This is typically just the inverse of **inScale**.
3. **popScale**. Scale factor for population. Often set to  $10^{-6}$ , though need to check on the population level inconsistencies.
4. **lScale**. Scale factor for labor volumes. [Need to verify scale in 'wages' database.]
5. **eScale**. Scale factor for energy. The energy volumes are in million tons of oil equivalent (MTOE). A scale factor of 0.001 is typically used.
6. **watScale**. Scale factor for water, typically  $10^{-12}$  is used. [Check units of volume database.]
7. **cScale**. Scale factor for emissions, typically 0.001 is used. [Check units for emissions.]
8. **ifCEQ**. Convert emissions to CEq. Input emissions are in CO<sub>2</sub>eq. If this flag is set to 1, the units will be converted to Ceq. The climate module [tbd] is based on emissions in Ceq.
9. **ArmFlag**. Set to 1 for agent-based Armington. If the flag is set to 0, the top level Armington sourcing is done at the aggregate level. [This is a new option.]
10. **MRIO**. Set to 1 for agent-based second-level Armington, i.e. to use the MRIO version of the model. [This is a new option.]
11. **ifNRG**. Set to 1 to use energy volumes. This will provide a volume/price split for energy commodities.
12. **ifNRGNest**. Set to 1 for energy nesting. A value of 0 will have only a single nested energy nest.

13. **ifMCP**. Set to 1 for MCP. A value of 0 will use NLP. The objective function is Walras' Law.
14. **ifLandCET**. Set to 1 to use CET for land allocation. A value of 0 will use the additive CET specification.
15. **ifSUB**. Set to 1 to substitute out equations. A value of 0 will have the full model specification without substitution. This will considerably increase the size of the model.
16. **IFPOWER**. Set to 1 for power module, which requires the power database, else set to 0.
17. **IFWATER**. Set to 1 for water module, which requires the water database, else set to 0.
18. **ifAggTrade**. Set to 1 to aggregate trade in SAM. A value of 0 will have the full bilateral trade matrices output as part of the SAM.
19. **skLabgrwgt**. Set to between 0 and 1. This determines the growth assumptions for skilled and unskilled labor. The growth of skilled labor is driven by the growth of the appropriate education categories. The growth of unskilled labor will then be calculated by residual since the growth of overall labor will be determined by the growth of the working age population (15-64). One can use this parameter to modulate this calculation. If the weight is set to 1, then the growth rate of skilled labor will match exactly the growth of the corresponding education categories, ignoring any resulting impact on the growth of unskilled labor. At the other extreme, if the weight is set to 0, the growth of both skilled and unskilled will be identical and equal to the growth rate of total labor, i.e. the education profiles are fully ignored. Users can choose an intermediate value between 0 and 1.

#### F.2.4 Comparative static simulations

In the current setup, there is very little difference between running comparative static and simulations dynamic simulations. Both the `runsim.gms` and `[basename]Opt.gms` files are coded to run both types of simulations. One of the key differences between the two is the time framework. The standard comparative static simulation has three 'time' periods labeled **base**, **check** and **shock**. The user is free to modify these in the 'Opt' file. One of the primary purposes of the comparative static version of the model is to test model initialization, calibration and homogeneity. It is also extremely useful to test new model specifications or parameterizations. With a new model version—for example a new aggregation, new specification or parameterization, it is always recommended practice to do a full diagnostic check.

A standard diagnostic check is to run the comparative static model with a single homogeneity shock in the **shock** time period. The **base** year is never run, it is meant to be able to replicate the initial database. The **check** simulation replicates **base** if the initialization and calibration are working correctly. For this reason, it is important to look at the largest residual in the listing file (or console). Open the listing file and do a search for **LOOPS**. This should take you to the diagnostics for the first simulation, i.e. the **check** simulation. Model diagnostics will appear by scrolling down—for example model size. The key diagnostic is **INITIAL POINT STATISTICS**. The maximum of **F** should be 0 or near zero. The largest error should reflect more or less the precision of the input SAM. One may also want to observe the initial Jacobian to see the range of the minimum and maximum elements. The range should be relatively narrow. A good maximum should be no larger than `1.0e003`. A larger number is an indication of a scaling issue.

A second check of the model is to look at the values for the **LHS** in the list code, which stands for the left-hand side. This is controlled by the **LIMROW** option. It may default to either 0 or 3.

In the case of 0, no equations will be listed. In the case of 3, up to 3 in any block will be listed. Setting LIMROW to a higher number will increase the number of equations printed in a block, for example: `Options limrow=100 ;`. The equations listing shows the value of the residual of an equation, i.e. it collects all endogenous variables on the left-hand side and evaluates the expression and it collects all additive exogenous variables and parameters on the right-hand side and evaluates the right-hand side expression. It then displays the value of the left-hand side that should match the displayed right-hand side. In most cases, the right-hand side will evaluate to 0, but it does not have to.<sup>20</sup> If there is a serious initialization/calibration problem, it may be necessary to increase `limrow` and visually go through the equation listings. For a quick inspection, do a search for LHS to find the first equation in the listing file. Then search for `****`. The four asterisks indicate an 'infeasible' solution. Note that in most cases, the infeasibility is likely to be a very small number—and typically reflects the accounting precision of the incoming SAM. To avoid nuisance infeasibility, one can increase the tolerance level by using the `tolinfrep` attribute of a model. In the default settings, the infeasibility tolerance has been set to `1.0e-005`, and with a well-balanced SAM, this will generally avoid any listed infeasibilities.

The homogeneity test requires a shock file. The shock file should have the same name as the simulation. Thus if the simulation is assigned the name `COMP`, the user should create the a file named `COMPShk.gms`. Listing F.13 illustrates one way of implementing the homogeneity shock. It tests the value of the time period, and then increases the value of the exogenous numéraire by 50%.

Listing F.13: Homogeneity shock

---

```

2  if(sameas(tsim,"shock"),
3     pnum.fx(tsim) = 1.5*pnum.l(tsim) ;
4  ) ;

```

---

It is useful to do a quick check of the diagnostics by loading the output CSV file into Excel—and most convenient in Excel's pivot tables. The first thing to look at is the resulting SAM's. All accounts should be perfectly balanced for each individual region and for each time period. The second diagnostic is to compare the `base` values with the `check` values. They should be identical to within very narrow margins. A third check is to compare the `check` values with the `shock` values. All volume variables should be identical and all value variables and prices should have increased by the same percentage amount as the numéraire, for example 50%.

These are minimal diagnostics, but the user may want to do others—particularly if there is a change in specification. These would include shocks to taxes, for example tariffs, shock to factor stocks, productivity shocks, etc.

The distribution comes with an option file for the '10x10' aggregation, as well as the generic `runsim.gms` file. The latter is set to handle many different situations but users may find it useful and/or necessary to make a copy of it and insert modifications. The distribution also comes with a command file that is needed to run the file in a Windows console. To run the homogeneity test, the command line could be the following:

```
runsim Homog Comp CompStat 0
```

The first argument will be the simulation name. In comparative static simulations the second argument is ignored, but must be present. In dynamic simulations it is the name of the baseline file.

---

<sup>20</sup> For example the labor equilibrium condition will have the exogenous supply of labor on the right-hand side.

The third argument must be `CompStat` for comparative static simulations. The fourth argument is ignored for comparative static simulations, it is used for dynamic simulations. It is best to set it to zero.

Users will need to modify the `runsim` command file for their installations. An example is depicted below.

```
gams runSim --simName=%1 --BauName=%2 --simType=%3 --ifCal=%4 --baseName=10x10
--odir=v:\Output\EnvLink\10x10 -idir=..\model5n
-scrdir=v:\Output\EnvLink\10x10 -ps=9999 -pw=150
```

The `runSim` command file expects at least the four arguments as described above. It is setup for a specific project. In this case it is setup for a project named '10x10'. The user needs to specify a folder for the output directory. Use `.` to specify the current directory. It is also required to specify the model directory. In the example above, the model directory is at the same level as the simulation directory and named `model5n`. The remaining parameters are optional and the user is free to add others.<sup>21</sup>

[NEW 01-Dec-2018] The code in the 'Opt' file has been modified somewhat to improve the handling of starting point for the model simulation. The old code allowed to start a simulation by reading in a previous baseline using the `execute_loadpoint` option of GAMS. This is handy if you are running shock simulations starting in a future year relative to the reference year. The new code allows the user to specify any simulation as a starting point. The user can identify the starting simulation on the command line with the `startName` option. The Windows command file `runsimR.cmd` shows an example of how to use an existing comparative static simulation as a starting point. If there is no additional shock, the model should solve immediately.

## F.2.5 Dynamic simulations

Dynamic simulations are in principle only somewhat more complicated than comparative static simulations—though in practice raise considerably more problems. There are two key differences. The first is the time dimension. This is under the discretion of the user. For GTAP V9, the starting year should be set at 2011. The terminal year and the intermediate years are at full discretion. The model is setup to handle year step sizes of more than 1 [though this probably needs additional testing since it hasn't been used in a while]. One potential problem with multi-year time steps could be convergence, though this has proven to be less of a problem further out when the model has reached some sort of steady state. The second key difference is the implementation of the vintage capital formulation. This in principle should be transparent to the user.

In principle there is nothing to change in the 'Opt' file nor in the `runsim.gms` file. There are default statements that are implemented in the file `iterloop.gms` that generate the necessary closures for dynamic simulations—either the baseline, or pre-calibrated scenarios. Users have some control in shaping the baseline by modifying or adding statements in the `BaUShk.gms` file. In the distributed file, there are two adjustments to the standard baseline. The first can phase out net capital flows between two years. The second targets the investment to GDP ratio for some given future year—both the target and the year are provided by the user. It endogenizes the savings rate in order to achieve the desired investment ratio.

<sup>21</sup> Simulation output can be voluminous depending on the aggregation and the time span of a simulation. With increasing use of backups and the cloud, it is sometimes useful to store output on non-critical storage in order to minimize bandwidth problems and exceeding allowed storage capacity.

Other key dynamic assumptions are embedded in the `initScen.gms` file that is in the *Model* directory. In particular the default assumptions on labor productivity wedges (`glAddShift` and `glMltShift`), autonomous yield growth (`yexo`), autonomous energy efficiency improvement (`aeei` and `aeec`) and improvements in trade margins (`tteff`) are included in this file.[We may revisit this looking ahead as it is not satisfactory to have these assumptions included in the core model code.]

The mechanics of the baseline are fairly straightforward—shaping the baseline is not always straightforward. One of the key issues to arise is the growth of the capital stock. Base year conditions may lead to extreme jumping off conditions. For example over-investment may lead to sharply declining rates of return, and under-investment the reverse. The initial growth of the capital stock will depend on base year investment (and savings) rates, net capital flows, the initial stock of capital and the depreciation rate. On balance, one would like to have as a starting point for the baseline relatively steady returns to capital, unless there is a good reason to expect rates of return to rise or drop. It may involve adjusting the initial capital stock and rates of depreciation. It is not possible to change the initial savings rate including the capital account, but these can be adjusted over time to yield a desired path for capital accumulation and rates of return. This is one of the purposes for using a long-run target for the investment to GDP.

To run the baseline simulation, the command line could be the following:

```
runsim BaU BaU RcvDyn 1
```

The first argument and second arguments are the name of the baseline. In the case of running the baseline, the second argument is redundant, but necessary. The third argument is to invoke a dynamic-type scenario. The fourth argument invokes the code for dynamic calibration. On occasion, the model will fail to converge to the terminal year and typically this is due to factor prices going to some extreme—notably zero for capital in the case of investment exceeding demand for capital. One may have to restrict the simulation to an intermediate year and assess intermediate results to understand the nature of the problem. This can be done by restricting the time loop, for example `loop(tsim$(years(tsim) 1e 2020)`.

Once the user has a satisfactory baseline, dynamic shock simulations can be implemented. They are associated with a specific shock file with a base name linked to the simulation name, [`simName`]`Shk.gms`, for example [`SSP2ccShk.gms`] to run the SSP2 baseline with climate change impacts. To run a pre-calibrated dynamic scenario, the command line could be the following:

```
runsim SSP2CC BaU RcvDyn 0
```

The first argument is the name of the simulation, and `runSim.gms` will try and open and implement a shock file with a filename that incorporates the simulation name. The second is the name of the baseline file. At a minimum this is necessary because the shock simulation needs to extract from the baseline calibrated information from the baseline—notably the labor productivity parameter and real government expenditures. Both are assumed to be exogenous in the shock simulations.<sup>22</sup>

[NEW 01-Dec-2018] The code in the 'Opt' file has been modified to allow for more flexibility in terms of starting from an existing simulation. The old code only allowed for starting from the

<sup>22</sup> Users can extract other information from the baseline on an as-needed basis.

baseline simulation. The new code allows for starting from any simulation (including the baseline). An existing simulation is not always the best starting point for solving for a given year, particularly for a future year. The code allows the user to provide a year until which the simulation starts from an existing simulation or the solution from a previous period. To test the baseline simulation, for example, one could re-run the baseline with no shock but using the calibrated variables from the baseline, for example:

```
runsimrd noShk BaU RcvDyn 0 BaU 2030
```

where the command file `runsimrd.cmd` is used to invoke GAMS. This says to run a calibrated recursive dynamic scenario using the calibrated variables from the baseline scenario called BaU. In addition, use as the starting point the BaU scenario through the year 2030. If the dynamics is set up correctly, the model should solve within 1 iteration in each period. In effect, the *noShock* scenario is similar in concept to the *Check* scenario for comparative statics. The rule of thumb is that you can use the baseline till the first year of a shock. In most cases, initializing with the previous period's result is better once the shock has been implemented as the shock tends to lead to sharp deviations from the baseline. For example, if a shock starts in the year 2020, use the baseline values through 2020 as a starting point, then for subsequent years, the model with initialize values with the solution from the previous period—not the baseline simulation. The Windows command file `runAll.cmd` contains the batch commands to run the baseline and the no-shock scenario in sequence. Note that it will make a copy of the 'runSim.gms' file for each of the simulations and thus the list files will have separate names.

## F.3 Post-simulation processing

There is some rudimentary post-simulation processing that is still under development. It involves two steps. The first step is a GAMS file that reads in a solution and produces a number of model-based indicators and stores them in different CSV cubes. The CSV cubes can hold results from one or more simulations. These can be loaded into Excel files, preferably into pivot tables, and/or alternatively they can be loaded into statistical or graphical packages such as R for automated generation of figures.

### F.3.1 User inputs

The user typically prepares two files—a Windows command file, which facilitates automating the processing of the GDX file(s), and a user-based option file that is specific to the simulations. The user-based option file is assumed to have the name 'baseNameTab.gms', where 'baseNameTab' is replaced with the base name for the set of simulations, for example '10x10'. Listing F.14 provides an example of a 'Tab' file. There are a number of global options that first need to be initialized.

1. **xclDir**. This sets the directory for holding the Excel files. N.B. For the moment it is best to use the full path name as testing with relative paths has created errors. For example do not use ".\Doc\" or "Doc\". This seems to be an issue with the use of the VB scripts under Windows.
2. **inDir**. This sets the directory for the CSV files. Note that this will be initialized from the command line parameter '%oDir%'.

3. **wDir**. This sets the working directory. It normally defaults to the directory that was used to invoke the 'makCSV' program.
4. **simTgt**. The following set of options initialize the fields for the pivot tables. **simTgt** should be set to the code for a valid simulation or scenario, such as BaU.
5. **regTgt**. **regTgt** should be set to the code for a valid region in the model. Refer to the 'Sets' file to see the valid regions for a specific aggregation.
6. **timeTgt**. **timeTgt** should be set to a valid reporting year. Typically this will be the base, or reference, year.
7. **actTgt**. **actTgt** should be set to a valid activity, which can also be referenced in the 'Sets' file.

The user next provides the reporting years. The set **tr** is a subset of the full simulation time framework, which can include all solution years. The time framework also differs between comparative static and dynamic scenarios. The next option also permits sub-setting the number of activities reported. It defaults to all activities and activity aggregations. The user also needs to specify a conversion factor that is used to convert output from 'primary' electricity activities. 'Primary' electric activities are those that do not rely on energy inputs such as coal or gas. The latter have a conversion loss typically assumed to be a factor of 3, i.e. it takes 3 MTOE of coal or gas to produce 1 MTOE of electricity.<sup>23</sup> This factor is only valid for the power-enabled version of the model.

The final set of options is for the cubes (i.e. tables) that will be extracted from the GDX files. The set **tables** contains the full list of tables that are available in the **makTab.gms** file. The user determines which of the tables to produce by defining the table subset **ifTab**.

Listing F.14: 'Tab' file example for creating CSV cubes

```

2  * -----
3  *
4  *      Options for creating CSV cubes
5  *
6  *      Users typically select years to output and CSV cubes to create
7  *
8  * -----
9
10 * Set folder for Excel files
11
12 $setGlobal xclDir "v:\env10\10x10\doc\"
13
14 * Options for CreatePivot file
15
16 $setGlobal indir      %oDir%
17 $setGlobal wdir       %system.fp%
18 $setGlobal modDir     "..\Model"
19 $setGlobal simTgt     BaU
20 $setGlobal regTgt     EastAsia
21 $setGlobal timeTgt    2011
22 $setGlobal actTgt     Agriculture-a
23
24 * Select report years
25
26 $iftheni "%simType%" == "compStat"

```

<sup>23</sup> See <https://unstats.un.org/unsd/envaccounting/londongroup/meeting13/LG13-12a.pdf>.

---

```

28     set tr(t) "Reporting years" / base, check, shock / ;

30 $elseifi "%simType%" == "RcvDyn"

32     set tr(t) "Reporting years" / 2011, 2014, 2017, 2020, 2025, 2030 / ;

34 $endif

36 * Select reporting activities (a subset of aga — activities + aggregate activities

38 set aggaga(aga) "Activities to report" ;

40 * Report all

42 aggaga(aga) = yes ;

44 scalar elyPrmNrgConv "Primary electric conversion factor" / 3 / ;

46 * Pivot tables to create

48 * List of tables

50 set tables /
51     gdppop      "Macro data"
52     factp       "Factor prices"
53     kappah      "Household direct tax rate"
54     rgovshr     "Government expenditures"
55     savinv      "Savings investment balance"
56     xp          "Output by activity"
57     va          "Value added by activity and factor"
58     inv         "Investment"
59     emi         "Emissions"
60     cost        "Production costs"
61     ydecomp     "Growth decomposition"
62     trade       "Trade by sector"
63     bilat       "Bilateral trade"
64     lab         "Labor demand"
65     pow         "Power sector variables"
66     tot         "Terms of trade"
67     nrg         "Energy indicators"
68     depl        "Depletion variables"
69     shock       "For future use"
70 / ;

72 * Selected tables

74 set ifTab(tables) /
75     gdppop      "Macro data"
76     factp       "Factor prices"
77     kappah      "Household direct tax rate"
78     rgovshr     "Government expenditures"
79     savinv      "Savings investment balance"
80     xp          "Output by activity"
81     va          "Value added by activity and factor"
82     inv         "Investment"
83     emi         "Emissions"
84     cost        "Production costs"
85     ydecomp     "Growth decomposition"
86     tot         "Terms of trade"
87     trade       "Trade by sector"
88     bilat       "Bilateral trade"
89     lab         "Labor demand"
90 * pow         "Power sector variables"
91 * nrg         "Energy indicators"
92 * depl        "Depletion variables"
93 * shock       "For future use"
94 / ;

```

---



### F.3.2 Extracting the indicators

The user can invoke the **makCSV** GAMS program by entering the following command (or equivalent in the GAMS IDE):

```
gams makCSV --simname=BaU --BaUName=BaU --simType=RcvDyn --ifCal=1
--ifAppend=0 --BaseName=10x10 --odir=z:\Output\Env10\10x10 -idir=..\Model
```

Many of the command line options are identical to the **runsim** command. The user specifies the name of the simulation (**simname**), the name of the baseline simulation, potentially the same (**BaUName**)<sup>24</sup>, the simulation type (**simType**), and the dynamic calibration flag (**ifCal**). One of the key command line arguments is **ifAppend**. If **ifAppend** is set to 0, all CSV cubes will be created anew—erasing any existing CSV cubes. If **ifAppend** is set to 1, the list of indicators being extracted will be appended to the end of an existing CSV file—thus enabling the concatenation of indicators from multiple scenarios in a single cube. If running from a Windows command file—the first invocation of **makCSV** should set the append parameter to 0, and all subsequent invocations should set the append parameter to 1. The output directory should be where the GDX files are located and will also be the location of the output CSV files.

Once the CSV cubes have been created, a new command file will also have been created the contains the commands to convert the CSV cubes into Excel files, or, if the Excel files already exist, will 'refresh' the pivot tables contained in the Excel file. This command file will only work in a Windows-based terminal window as the commands rely on VB scripts. To invoke the command, simply type 'baseNamePivot', for example '10x10Pivot'. The command file will have been created by the 'makCSV' program. It will only invoke the appropriate commands for the requested tables. Sample 'runtab.cmd' files will automatically invoke the command.

The whole system relies on four GAMS files and one 'VBS' file that are typically located in the 'Model' directory. The five files are:

1. **makCSV**. **makCSV.gms** is the master GAMS file that reads the options, invokes the indicator extraction file (**makTab**) and creates the 'baseNamePivotPivot.cmd' file.
2. **makTab**. **makTAB.gms** is the core GAMS file that will extract the data from GDX files, create indicators, and save them to the CSV files.
3. **setupPivot**. **setupPivot.gms** is invoked by the **makCSV.gms** to create the command lines for the 'baseNamePivot.cmd' file. It is a small file invoked using 'batinclude' with arguments that simplifies the creation of the command file.
4. **CreatePivot**. **createPivot.gms** writes the VB script that creates the Excel files with the built-in pivot tables. It is only invoked for Excel files that are being created for the first time. Otherwise, the 'refresh' VB script is invoked.
5. **refresh**. **refresh.vbs** is a VB script that refreshes a pivot table in an Excel file (without opening Excel).

<sup>24</sup> The baseline is on occasion used to calculate indicators relative to the baseline.

### F.3.3 Standard indicators

The GAMS file 'makCSV.gms' is the file that contains code to create indicators from simulation results. It extracts selected model results from a simulation (stored as a GDX container), creates an indicator (potentially without transformation) and in most cases will also aggregate across user-defined regions and sectors. Classes of indicators are stored together in CSV cubes. There are several advantages to having indicators in separate cubes. First, it keeps the sizes of the CSV files to a reasonable size; making them easier and quicker to load. Second, the geometry of the cube can be tailored to specific indicators. For example, macroeconomic indicators typically do not need sector or other qualifiers. Table F.7 describes the current list of CSV cubes.

The following sections describe each of the indicators. The left-hand side should be thought of as being indexed by aggregate region ( $R$ ), time and simulation. Normally each modeled region is mapped to its own aggregate region, but could also be mapped to a true aggregate region that would be composed of 1 or more modeled regions. The right-hand side is normally composed of a model indicator. There is often a one-to-one mapping, i.e. the aggregation is just a simple sum of a model indicator.

#### The GDPPop cube

The formulas below calculate the indicators contained in the GDPPop cube. The formulas need little explanation. GDP at 2005 PPP exchange rates uses the base year PPP exchange rate derived from the SSP scenario file. The first two population indicators, P1564 and PopT, refer to the exogenous information that is used as input for a model simulation. The indicator Pop refers to the population that emerges from the model. This could differ from PopT depending on the initial level of the population—though the growth rates should be identical.

$$RGDPMP_R = \sum_{r \in R} RGDPMP_r$$

$$RGDPMPPP_R = \sum_{r \in R} RGDPMP_r PPP = \sum_{r \in R} RGDPMP_r \frac{GDPScen_{Mod, SSP, GDPPPP05, r, t0}}{GDPScen_{Mod, SSP, GDP, r, t0}}$$

$$GDPMP_R = \sum_{r \in R} GDPMP_r$$

$$PGDPMP_R = \sum_{r \in R} GDPMP_r \Big/ \sum_{r \in R} RGDPMP_r$$

$$P1564_R = \sum_{r \in R} PopT_{r, P1564}$$

$$PopT_R = \sum_{r \in R} PopT_{r, PTOTL}$$

$$Pop_R = \sum_{r \in R} Pop_r$$

$$RGDPPC_R = \sum_{r \in R} RGDPMP_r \Big/ \sum_{r \in R} Pop_r$$

The GDP cube also contains some indicators of welfare. The indicators  $EV$ ,  $EVG$  and  $EVI$  represent the equivalent variation for households, government and investment expenditures, respectively. The household welfare indicator is derived from the indirect utility function for the user-chosen utility function. Government and investment expenditure functions have a simple formula for equivalent variation based on the CES preference function. Note that the model allows for multiple households, though the standard GTAP database only includes a single representative household. The remaining three indicators measure the volume of private, public and investment expenditures.

$$\begin{aligned}
 EV_R &= \sum_{r \in R} \sum_h EV_{r,h} \\
 EVG_R &= \sum_{r \in R} YFD_{r,gov} \frac{YFD_{r,gov,t0}}{YFD_{r,gov}} \\
 EVI_R &= \sum_{r \in R} YFD_{r,inv} \frac{YFD_{r,inv,t0}}{YFD_{r,inv}} \\
 EVT_R &= EV_R + EVG_R + EVI_R
 \end{aligned}$$

$$\begin{aligned}
 XFD_R &= \sum_{r \in R} \sum_h XFD_{r,h} \\
 XFDG_R &= \sum_{r \in R} XFD_{r,gov} \\
 XFDI_R &= \sum_{r \in R} XFD_{r,inv}
 \end{aligned}$$

### The FactP cube

The **FactP** cube provides aggregate (or economy-wide) factor prices. The labels for capital, land and water return are hard-coded. The labels for the aggregate wages will be the same as the labels for the labor types (e.g. `nsk` and `skl`).

$$\begin{aligned}
 AW_{R,l} &= \sum_{r \in R} \sum_a PF_{r,l,a} XF_{r,l,a} \Big/ \sum_{r \in R} \sum_a XF_{r,l,a} \\
 TRENT_R &= \sum_{r \in R} \sum_a PF_{r,cap,a} XF_{r,cap,a} \Big/ \sum_{r \in R} \sum_a XF_{r,cap,a} \\
 PTLand_R &= \sum_{r \in R} PTLand_r TLand_r \Big/ \sum_{r \in R} TLand_r \\
 PTNRS_R &= \sum_{r \in R} \sum_a PF_{r,nrs,a} XF_{r,nrs,a} \Big/ \sum_{r \in R} \sum_a XF_{r,nrs,a} \\
 PTH2O_R &= \sum_{r \in R} \sum_a PF_{r,wat,a} XF_{r,wat,a} \Big/ \sum_{r \in R} \sum_a XF_{r,wat,a}
 \end{aligned}$$

**The Kappah cube**

This cube contains only the average rate of transfer between households and government.<sup>25</sup>

$$Kappah_R = \sum_{r \in R} \kappa_r^h YH_r \bigg/ \sum_{r \in R} YH_r$$

**The RGovShr cube**

This cube contains only the average ratio of real government expenditure to real GDP.<sup>26</sup>

$$RGovShr_R = 100 \cdot \sum_{r \in R} XFD_{r,gov} \bigg/ \sum_{r \in R} RGDPMP_r$$

---

<sup>25</sup> The main purpose of this is to assess the impacts of the standard closure rule that has the government deficit fixed and net household transfers adjust to meet the fiscal target.

<sup>26</sup> In the standard baseline, this ratio is fixed and drives the growth of real government expenditures. In shock simulations, the ratio is endogenous and real government expenditures are fixed to baseline levels.

Table F.7: CSV cubes

<i>File name</i>	<i>Description</i>
GDPPop	Contains GDP and population including: real GDP at market price ( <b>rgdmp</b> ), nominal GDP at market price ( <b>gdpmp</b> ), GDP at market price deflator ( <b>pgdmp</b> ), working age population ( <b>P1564</b> ), input population ( <b>PopT</b> ), simulated population ( <b>Pop</b> ), and real GDP per capital ( <b>rgdppc</b> ).
Factp	Contains aggregate factor prices including: wages identified by skill, capital ( <b>trrent</b> ), land ( <b>ptland</b> ), natural resources ( <b>ptnrs</b> ) and water, if available ( <b>pth2o</b> ).
kappah	Contains net direct transfers to households as a share of income ( <b>kappah</b> ).
RGovShr	Contains real government expenditures as a share of real GDP ( <b>rgovshr</b> ).
savin	Contains various indicators relating to savings and investment including: nominal GDP at market price ( <b>gdpmp</b> ), real GDP at market price ( <b>rgdmp</b> ), real investment as a share of real GDP ( <b>rinvshr</b> ), aggregate real investment ( <b>rinv</b> ), aggregate nominal investment ( <b>inv</b> ), aggregate nominal household saving ( <b>savh</b> ), household savings rate ( <b>aps</b> ), nominal government saving ( <b>savg</b> ), nominal foreign saving ( <b>savf</b> ) and nominal depreciation ( <b>deprY</b> ).
Output	Contains various indicators relating to output: unweighted output ( <b>xp</b> ), output weighted by base year prices ( <b>xpw</b> ), nominal output ( <b>xpd</b> ), output price using unweighted volumes ( <b>px</b> ), output price using base year price-weighted volumes ( <b>pxn</b> ), GAPS unweighted volume ( <b>xpg_GAPS</b> ), and Envisage output using GAPS base year volume ( <b>xpg_Env</b> ).
VA	Contains various indicators relating to value added by activity and factor: nominal value added ( <b>va_d</b> ), value added weighted by base year prices ( <b>va_n</b> ), price of value added using base-year price-weighted volumes ( <b>pva_n</b> ), unweighted volumes ( <b>va</b> ), and price of value added using unweighted volumes ( <b>pva</b> ).
INV	Contains various indicators relating to investment by sector: volume of sectoral investment ( <b>inv_sec</b> ), nominal investment by sector ( <b>invd_sec</b> ), sectoral share of aggregate nominal investment ( <b>invd_shr</b> ), sectoral share of nominal value added ( <b>va_shr</b> ), and ratio of sectoral investment share and value added share ( <b>invRatio</b> ).
EMI	Contains various indicators relating to greenhouse gas emissions: total emissions ( <b>emi</b> ), emissions linked to intermediate and final demand ( <b>emi_io</b> ), emissions linked to factor use ( <b>emi_fp</b> ), and emissions linked to output ( <b>emi_xp</b> ).
Cost	Contains various indicators relating to decomposing production cost: change in production cost ( <b>delpx</b> ), change in output ( <b>delxp</b> ), contribution of change in cost arising from change in input price ( <b>pcshr</b> ), contribution of change in cost arising from change in technology mix ( <b>xcshr</b> ), and cost share of input ( <b>cshr</b> ).
YDecomp	Contains various indicators relating to decomposing GDP growth: nominal GDP at factor cost ( <b>gdpcf</b> ), real GDP at factor cost ( <b>rgdpcf</b> ), real GDP at factor cost deflator ( <b>pgdpcf</b> ), contribution of GDP growth from factor growth ( <b>qdel</b> ), and contribution of GDP growth from factor productivity ( <b>ldel</b> ).
Trade	Contains various indicators relating to trade: nominal exports ( <b>exp_d</b> ), real exports ( <b>exp</b> ), export tax rate ( <b>etax</b> ), nominal imports at FOB prices ( <b>imp_fob_d</b> ), real imports at FOB prices ( <b>imp_fob</b> ), nominal imports at CIF prices ( <b>imp_cif_d</b> ), real imports at CIF prices ( <b>imp_cif</b> ) import tariff ( <b>mtax</b> ) and domestic absorption ( <b>Absorb</b> ).
Lab	Contains various indicators relating to labor markets: labor supply ( <b>ls</b> ), wage ( <b>twage</b> ), real wage ( <b>trwage</b> ), rural to urban migration ( <b>migr</b> ), labor supply by zone ( <b>lsz</b> ), average wage by zone ( <b>awage</b> ), equilibrium wage by zone ( <b>ewage</b> ) average real wage ( <b>arwage</b> ) and average equilibrium wage ( <b>aewage</b> ).
Power	Contains various indicators relating to power: output ( <b>XP</b> ), and average cost ( <b>PX</b> ).

### The SavInv cube

The **SavInv** cube contains the indicators below. Note that investment should equal aggregate savings (household, government and foreign) plus depreciation for any regional aggregation. Moreover, the sum of foreign saving at the world level should equal 0.

$$\begin{aligned}
 RGDPMP_R &= \sum_{r \in R} RGDPMP_r \\
 GDPMP_R &= \sum_{r \in R} GDPMP_r \\
 RInvShr_R &= 100 \cdot \sum_{r \in R} XFD_{r,inv} \Big/ \sum_{r \in R} RGDPMP_r \\
 RInv_R &= \sum_{r \in R} XFD_{r,inv} \\
 Inv_R &= \sum_{r \in R} YFD_{r,inv} \\
 SavH_R &= \sum_{r \in R} S_r^h \\
 SavG_R &= \sum_{r \in R} S_r^g \\
 SavF_R &= \sum_{r \in R} PW^{sav} S_r^f \\
 DeprY_R &= \sum_{r \in R} DeprY_r \\
 APS_R &= 100 \cdot \sum_{r \in R} S_r^h \Big/ \sum_{r \in R} YD_r
 \end{aligned}$$

### The Output cube

The output cube contains various indicators related to output. Three reflect different measures of the level of output—nominal (**XP**D), unweighted real (**XP**) and weighted real (**XP**W). For most purposes the appropriate index is the weighted real—particularly when aggregating across activities using different units. Potential exceptions are when aggregating energy indicators as these may be evaluated in the same units, for example MTOE or GWhr. Similarly there are two price indices: unweighted (**PX**) and weighted (**PX**N). The cube also contains the total factor productivity parameter (for installed or *Old* capital).

$$\begin{aligned}
 XP_{R,A} &= \sum_{r \in R} \sum_{a \in A} XP_{r,a} \\
 XPW_{R,A} &= \sum_{r \in R} \sum_{a \in A} PX_{r,a,0} XP_{r,a}
 \end{aligned}$$

$$\begin{aligned}
XPD_{R,A} &= \sum_{r \in R} \sum_{a \in A} PX_{r,a} XP_{r,a} \\
PX_{R,A} &= \sum_{r \in R} \sum_{a \in A} PX_{r,a} XP_{r,a} \bigg/ \sum_{r \in R} \sum_{a \in A} XP_{r,a} \\
PXN_{R,A} &= \sum_{r \in R} \sum_{a \in A} PX_{r,a} XP_{r,a,0} \bigg/ \sum_{r \in R} \sum_{a \in A} PX_{r,a,0} XP_{r,a,0} \\
\lambda_{xpOld}_{R,A} &= \sum_{r \in R} \sum_{a \in A} \lambda_{r,a,Old}^{xp} PX_{r,a,Old,0} XP_{r,a,Old} \bigg/ \sum_{r \in R} PX_{r,a,Old,0} XP_{r,a,Old}
\end{aligned}$$

[Optional] There are two additional volume indicators linked to an external estimate of base year volumes.<sup>27</sup> The first indicator is the external indicator itself—in its own units and with its own growth trends (e.g. from an FAO or IEA projection). The second is the growth trend using the Envisage results but re-based to that of the external data. This allows for easier comparison of trends between institutions. In the case of the comparison with the FAO GAPS projections, the FAO projections are labeled **xpg\_GAPS** and the Envisage projections are labeled **xpg\_Env**. Note that the base years for GAPS and Envisage do not line up (2012 versus 2011), nor do the time steps line up.

### The VA cube

The value added cube contains various indicators related to value added. In addition to the various endowments that are part of the model definition, the indicators also include aggregations for labor (**tlab**), all non-labor factors (**tcap**) and all factors (**tot**). In most cases, the **VA\_N** indicator is the most appropriate for assessing aggregate value added. The **VA** indicator can be used in some cases to evaluate labor in person years depending on the calibration choices of the user.

$$\begin{aligned}
VA\_D_{R,A,V} &= \sum_{r \in R} \sum_{a \in A} \sum_{f \in V} PF_{r,f,a} XF_{r,f,a} \\
VA\_N_{R,A,V} &= \sum_{r \in R} \sum_{a \in A} \sum_{f \in V} PF_{r,f,a,0} XF_{r,f,a} \\
PVA\_N_{R,A,V} &= \sum_{r \in R} \sum_{a \in A} \sum_{f \in V} PF_{r,f,a} XF_{r,f,a} \bigg/ \sum_{r \in R} \sum_{a \in A} \sum_{f \in V} PF_{r,f,a,0} XF_{r,f,a} \\
VA_{R,A,V} &= \sum_{r \in R} \sum_{a \in A} \sum_{f \in V} XF_{r,f,a} \\
PVA_{R,A,V} &= \sum_{r \in R} \sum_{a \in A} \sum_{f \in V} PF_{r,f,a} XF_{r,f,a} \bigg/ \sum_{r \in R} \sum_{a \in A} \sum_{f \in V} XF_{r,f,a}
\end{aligned}$$

<sup>27</sup> This is particularly useful where the external estimate of base year volumes is in recognizable units—such as tons, joules, etc.

### The INV cube

The investment cube contains investment by sector—both in volume and value terms. The formulas take into account multi-year step sizes. The `InvRatio` indicator measures an activity's share of investment relative to its share of value added (in nominal terms).

$$\begin{aligned}
INV\_SEC_{R,A} &= \sum_{r \in R} \sum_{a \in A} \frac{\delta_{r,t-n}}{1 - (1 - \delta_{r,t-n})^n} \left[ \sum_v K_{r,a,v,t}^v - (1 - \delta_{r,t-n})^n \sum_v K_{r,a,v,t-n}^v \right] \frac{1}{\chi_r^k} \\
INVD\_SEC_{R,A} &= \sum_{r \in R} \sum_{a \in A} \frac{\delta_{r,t-n}}{1 - (1 - \delta_{r,t-n})^n} \left[ \sum_v K_{r,a,v,t}^v - (1 - \delta_{r,t-n})^n \sum_v K_{r,a,v,t-n}^v \right] \frac{PFD_{r,inv,t-n}}{\chi_r^k} \\
INVD\_SHR_{R,A} &= \frac{\sum_{r \in R} \sum_{a \in A} \frac{\delta_{r,t-n}}{1 - (1 - \delta_{r,t-n})^n} \left[ \sum_v K_{r,a,v,t}^v - (1 - \delta_{r,t-n})^n \sum_v K_{r,a,v,t-n}^v \right] \frac{PFD_{r,inv,t-n}}{\chi_r^k}}{\sum_{r \in R} \sum_a \frac{\delta_{r,t-n}}{1 - (1 - \delta_{r,t-n})^n} \left[ \sum_v K_{r,a,v,t}^v - (1 - \delta_{r,t-n})^n \sum_v K_{r,a,v,t-n}^v \right] \frac{PFD_{r,inv,t-n}}{\chi_r^k}} \\
VA\_SHR_{R,A} &= \frac{\sum_{r \in R} \sum_{a \in A} \sum_f PF_{r,f,a} XF_{r,f,a}}{\sum_{r \in R} \sum_a \sum_f PF_{r,f,a} XF_{r,f,a}} \\
InvRatio_{R,A} &= \frac{INVD\_SHR_{R,A}}{VA\_SHR_{R,A}}
\end{aligned}$$

### The Emissions cube

The emissions cube contains 5 indicators: total emissions, emissions linked to consumption of intermediate and final demand, emissions linked to factor-use and emissions linked to output. Emission indicators are output for all types of emissions. Physical emissions are all denominated in unit of gigatons.<sup>28</sup> Greenhouse gases are also output in units of CEq and CO2eq. If the model uses units of CEq, CO2eq is calculated by multiplying the emission level by a factor of 44/12. In the contrary case, emissions are converted to CEq using the factor 12/44.

$$\begin{aligned}
Emi_{R,Tot,em} &= \sum_{r \in R} Emi_{Tot,r,em} \\
Emi\_io_{R,A,em} &= \sum_{r \in R} \sum_{a \in A} \sum_i Emi_{r,em,i,a} \\
Emi\_io_{R,fd,em} &= \sum_{r \in R} \sum_i Emi_{r,em,i,fd} \\
Emi\_fp_{R,A,em} &= \sum_{r \in R} \sum_{a \in A} \sum_{fp} Emi_{r,em,fp,a} \\
Emi\_xp_{R,A,em} &= \sum_{r \in R} \sum_{a \in A} Emi_{r,em,Tot,a}
\end{aligned}$$

<sup>28</sup> Base non-greenhouse gas emissions are in units of gigagrams. They are divided by 1000 to convert to gigatons.



### The Trade cube

The first set of equations relate to exports. These are aggregated over regions and commodities, where the index  $I$  is an aggregate commodity, which may be an individually modeled commodity. The first two measure aggregate exports by commodity  $I$  at current FOB prices and at constant FOB prices. The third formula measures the export tax.

$$\begin{aligned}
 EXP\_D_{R,I} &= \sum_{s \in R} \sum_{i \in I} \sum_d PWE_{s,i,d} XW_{s,i,d} \\
 EXP_{R,I} &= \sum_{s \in R} \sum_{i \in I} \sum_d PWE_{s,i,d,0} XW_{s,i,d} \\
 ETAX_{R,I} &= 100 \cdot \frac{\sum_{s \in R} \sum_{i \in I} \sum_d PWE_{s,i,d} XW_{s,i,d}}{\sum_{s \in R} \sum_{i \in I} \sum_d PE_{s,i,d} XW_{s,i,d}} - 100
 \end{aligned}$$

The next set of five equations relate to imports. The first two equations evaluate current and constant imports at FOB prices, i.e. these indicators ignore the trade and transport margins. The next two equations use CIF import prices. The final equation evaluates the import tariff in percentage terms. Note that the variable  $XW$  represents the supply of exports and is adjusted by the iceberg parameter to convert to demand for imports.

$$\begin{aligned}
 IMP\_FOB\_D_{R,I} &= \sum_{d \in R} \sum_{i \in I} \sum_s PWE_{s,i,d} \lambda_{s,i,d}^w XW_{s,i,d} \\
 IMP\_FOB_{R,I} &= \sum_{d \in R} \sum_{i \in I} \sum_s PWE_{s,i,d,0} \lambda_{s,i,d}^w XW_{s,i,d} \\
 IMP\_CIF\_D_{R,I} &= \sum_{d \in R} \sum_{i \in I} \sum_s PWM_{s,i,d} \lambda_{s,i,d}^w XW_{s,i,d} \\
 IMP\_CIF_{R,I} &= \sum_{d \in R} \sum_{i \in I} \sum_s PWM_{s,i,d,0} \lambda_{s,i,d}^w XW_{s,i,d} \\
 MTAX_{R,I} &= 100 \cdot \frac{\sum_{d \in R} \sum_{i \in I} \sum_s PDM_{s,i,d} \lambda_{s,i,d}^w XW_{s,i,d}}{\sum_{s \in R} \sum_{i \in I} \sum_s PWM_{s,i,d} \lambda_{s,i,d}^w XW_{s,i,d}} - 100
 \end{aligned}$$

Domestic absorption is defined as the sum of demand by source, i.e. domestic and imported. Its evaluation depends on the Armington trade specification. The first formula evaluates domestic absorption in the case of the agent-specific Armington specification. The second formula measures domestic absorption in the case of the national-agent Armington specification.

$$\begin{aligned}
 Absorb_{R,I} &= \sum_{d \in R} \sum_{i \in I} \sum_{aa} \gamma_{r,i,aa}^{edd} PDT_{r,i} XD_{r,i,aa} + \gamma_{r,i,aa}^{edm} PMT_{r,i} XM_{r,i,aa} \\
 Absorb_{R,I} &= \sum_{d \in R} \sum_{i \in I} \sum_{aa} \gamma_{r,i,aa}^{eda} PAT_{r,i} XA_{r,i,aa}
 \end{aligned}$$

### F.3.4 Calculating the indicators

Calculating the indicators involves running the `makCSV` file for each simulation. The command line arguments are similar to those used to run the simulations themselves. One key argument is the `ifAppend` argument. Typically it is set to '0' for the first simulation and '1' for subsequent simulations. This forces the results to be concatenated in the CSV cubes. The sequence is easiest to run by using an existing command file that contains the full suite of commands to create the CSV cubes. In addition, the command file contains the code to create or refresh the Excel files and the pivot tables in the Excel file. If the Excel file already exists, the command file will request a refresh of the pivot table(s). The advantage of this is that it will preserve any worksheets, charts, tables that a user may have created and saved using a previous CSV cube. The `refresh.vbs` script is invoked to refresh an existing Excel file. If the Excel file does not exist, a new one will be created with a single pivot table.

### F.3.5 Creating the Excel files

The user can request a new Excel file by modifying the code in `CreatePivot.gms`. The basic purpose of the GAMS program is to create a Visual Basic script (vbs) that will automatically create a pivot table in an Excel file based on the relevant CSV cube. Each cube has its own geometry and there is a set of built-in geometries in the GAMS file—for example `VarByCost` or `VarByActivity`. When using `CreatePivot.gms`, a specific geometry must be chosen for each CSV cube that corresponds to the geometry of the cube. Users are free to add new geometries, using the existing ones as a guideline. Some of the geometries allow for labels to be sorted—see for example `VarByActivity`. The pivot tables that are created all have two dimensions—a row and a column. The rows will have the position number 1, columns position number 2, and the values position number 5. All other dimensions normally default to position number 3. Users can modify the pivot table geometry once the file has been created (and this will be preserved when the files are refreshed). It can be convenient to have some of the labels sorted automatically—though VB script can sometimes be fussy about this. Regions are sorted in all cases.<sup>29</sup> Some of the geometries allow for sorting by either activity or commodity.

### F.3.6 Creating graphs in R

A number of R scripts are available to create and save figures automatically. Most of these will generate a single graph for each region (both model regions and the aggregate regions) as well as for each simulation. The scripts are designed to save all the graphs for a simulation in a simulation folder. The scripts have the following outline:

- The scripts are called from a command line and the only argument is the name of CSV cube, for example `GDPPop`, without the `.csv` extension.
- Each script will read a script common to all that sets some global parameters, such as input and output directories. This common script is called `PlotOpt.R`. It normally only needs to be reviewed for a new model configuration.
- Each script also relies on two user-based files that contain the list of regions (`regions.csv`) and the time periods (`years.csv`). These files are aggregation specific.

<sup>29</sup> Only the model core regions are sorted, it would be nice to have sorting for the aggregate regions as well.

- The script relies on two libraries—`sqldf` and `ggplot2`. The first allows for reading CSV files as if they are cubes and users can then use the standard SQL commands to select and filter the incoming data.<sup>30</sup> The graphs are created using the `ggplot2` package.
- The function `read.csv.sql` is used to read in the CSV cube. This allows to skip some fields altogether and/or to select specific rows of the cube. For example, the `plotRGDPoc.R` script reads only two variables in the cube—`Pop` and `rgdmp`.<sup>31</sup>
- The body of the script loops over all possible simulations and regions and creates a figure for each combination based on the indicators read in from the CSV cube.
- The figures are stored by simulation. If the parent folder is called `PROJNAME\DOC` then the figures will be stored in `PROJNAME\DOC\SIM1`, `PROJNAME\DOC\SIM2`, ..., `PROJNAME\DOC\SIMn`. There could be other ways to organize the folder structure depending on the specific needs of the user.

Running of the scripts can be automated. See for example the file `plotAll.cmd`.

---

<sup>30</sup> The alternative is to use the standard `read.csv` function. The downside to this is that it will read the entire cube, even if only a small slice is desired.

<sup>31</sup> Unfortunately, `read.csv.sql` does not strip double quotes from the incoming cube so the R script has been designed to handle the string labels with double quotes.

# Bibliography

- Aguiar, A., B. Narayanan, and R. McDougall. 2016. “An Overview of the GTAP 9 Data Base.” *Journal of Global Economic Analysis*, 1(1): 181–208. doi:[10.21642/JGEA.010103AF](https://doi.org/10.21642/JGEA.010103AF).
- Britz, W., I. Pérez, A. Zimmermann, and T. Heckeles. 2007. “Definition of the CAPRI Core Modelling System and Interfaces with other Components of SEAMLESS-IF.” SEAMLESS Integrated Project, SEAMLESS Report No. 26. [http://www.seamless-ip.org/Reports/Report\\_26\\_PD3.5.1.1.pdf](http://www.seamless-ip.org/Reports/Report_26_PD3.5.1.1.pdf).
- Britz, W., and D. van der Mensbrugghe. 2016. “Reducing unwanted consequences of aggregation in large-scale economic models - A systematic empirical evaluation with the GTAP model.” *Economic Modelling*, 59: 463–472. doi:[10.1016/j.econmod.2016.07.021](https://doi.org/10.1016/j.econmod.2016.07.021).
- Burniaux, J. 1987. *Le Radeau de la Méduse: Analyse des dilemmes alimentaires*. Paris, France: Economica.
- Burniaux, J., and D. van der Mensbrugghe. 1991. “Trade Policies in a Global Context: Technical Specification of the Rural/Urban-North/South (RUNS) Applied General Equilibrium Model.” OECD, OECD Development Centre Technical Papers No. 48, November. doi:[10.1787/18151949](https://doi.org/10.1787/18151949).
- Burniaux, J.M., G. Nicoletti, and J. Oliveira-Martins. 1992. “GREEN: A Global Model for Quantifying the Costs of Policies to Curb CO<sub>2</sub> Emissions.” *OECD Economic Studies*, 19: 49–92. <http://www.oecd.org/eco/growth/35044127.pdf>.
- Burniaux, J.M., and D. van der Mensbrugghe. 1994. “The RUNS Global Trade Model.” *Economic & Financial Modelling*, pp. 161–282.
- Cardenete, M.A., A.I. Guerra, and F. Sancho. 2012. *Applied General Equilibrium: An Introduction*. Heidelberg: Springer. doi:[10.1007/978-3-642-24746-0](https://doi.org/10.1007/978-3-642-24746-0).
- Chateau, J., R. Dellink, and E. Lanzi. 2014. “An Overview of the OECD ENV-Linkages Model: Version 3.” OECD, OECD Environment Working Papers No. 65. doi:[10.1787/5jz2qck2b2vd-en](https://doi.org/10.1787/5jz2qck2b2vd-en).
- Chepeliev, M. 2018. “Development of the GTAP 9.2 Air Pollution Data Base.” Global Trade Analysis Project (GTAP), Department of Agricultural Economics, Purdue University, West Lafayette, IN, GTAP Research Memorandum No. Forthcoming.
- Chepeliev, M., and D. van der Mensbrugghe. 2017. “Global Energy Subsidies Reform: Inclusive Approaches to Welfare Assessment.” Global Trade Analysis Project (GTAP), Department of Agricultural Economics, Purdue University, West Lafayette, IN, Presented at the 20th Annual Conference on Global Economic Analysis, West Lafayette, Indiana, United States. [https://www.gtap.agecon.purdue.edu/resources/res\\_display.asp?RecordID=5299](https://www.gtap.agecon.purdue.edu/resources/res_display.asp?RecordID=5299).

- Corong, E. 2017. "Incorporating Domestic Margins into GTAP Data Base and Model." Global Trade Analysis Project (GTAP), Department of Agricultural Economics, Purdue University, West Lafayette, IN, Paper prepared for the 20th Annual Conference on Global Economic Analysis, West Lafayette, IN, USA.
- Corong, E., T. Hertel, R. McDougall, M. Tsigas, and D. van der Mensbrugghe. 2017. "The Standard GTAP Model, Version 7." *Journal of Global Economic Analysis*, 2(1): 1–119. doi:[10.21642/JGEA.020101AF](https://doi.org/10.21642/JGEA.020101AF).
- Dellink, R., J. Chateau, E. Lanzi, and B. Magné. 2017. "Long-term economic growth projections in the Shared Socioeconomic Pathways." *Global Environmental Change*, 42: 200–214. doi:<http://dx.doi.org/10.1016/j.gloenvcha.2015.06.004>.
- Dixon, P.B., and M.T. Rimmer. 2006. "The Displacement Effect of Labour-Market Programs: MONASH Analysis." *Economic Record*, 82: S26–S40. doi:[10.1111/j.1475-4932.2006.00330.x](https://doi.org/10.1111/j.1475-4932.2006.00330.x).
- Dixon, P.B., and M.T. Rimmer. 2002. *Dynamic general equilibrium modelling for forecasting and policy: a practical guide and documentation of MONASH*, 1st ed., vol. 256. Amsterdam: Elsevier.
- Dixon, P.B., and M.T. Rimmer. 2005. "Mini-Usage: reducing barriers to entry in dynamic CGE Modeling." Report. <https://www.gtap.agecon.purdue.edu/resources/download/2251.pdf>.
- Giesecke, J.A., N.H. Tran, E.L. Corong, and S. Jaffee. 2013. "Rice Land Designation Policy in Vietnam and the Implications of Policy Reform for Food Security and Economic Welfare." *Journal of Development Studies*, 49(9): 1202–1218. doi:[10.1080/00220388.2013.777705](https://doi.org/10.1080/00220388.2013.777705).
- Hanoch, G. 1975. "Production and Demand Models with Direct or Indirect Implicit Additivity." *Econometrica*, 43(3): 395–419. <http://www.jstor.org/stable/1914273>.
- Harris, J.R., and M.P. Todaro. 1970. "Migration, Unemployment and Development: A Two-Sector Analysis." *The American Economic Review*, 60(1): 126–142. <http://www.jstor.org/stable/1807860>.
- Hertel, T.W., ed. 1997. *Global Trade Analysis: Modeling and Applications*. New York, NY: Cambridge University Press.
- Hertel, T.W., E. Peterson, Y. Surry, P.V. Preckel, and M.E. Tsigas. 1991. "Implicit Additivity as a Strategy for Restricting the Parameter Space in Computable General Equilibrium Models." *Economic and Financial Computing*, 1: 265–289.
- Howe, H. 1975. "Development of the extended linear expenditure system from simple saving assumptions." *European Economic Review*, 6(3): 305–310. doi:[10.1016/0014-2921\(75\)90014-8](https://doi.org/10.1016/0014-2921(75)90014-8).
- Jorgenson, D.W., H. Jin, D.T. Slesnick, and P.J. Wilcoxon. 2013. "Chapter 17 - An Econometric Approach to General Equilibrium Modeling." In *Handbook of Computable General Equilibrium Modeling SET, Vols. 1A and 1B*, edited by P. B. Dixon and D. W. Jorgenson. Elsevier, vol. 1 of *Handbook of Computable General Equilibrium Modeling*, pp. 1133–1212. doi:[10.1016/B978-0-444-59568-3.00017-1](https://doi.org/10.1016/B978-0-444-59568-3.00017-1).
- Kee, H.L., A. Nicita, and M. Olarreaga. 2009. "Estimating Trade Restrictiveness Indices." *The Economic Journal*, 119(534): 172–199. doi:[10.1111/j.1468-0297.2008.02209.x](https://doi.org/10.1111/j.1468-0297.2008.02209.x).

- Lanz, B., and T. Rutherford. 2016. "GTAPinGAMS: Multiregional and Small Open Economy Models." *Journal of Global Economic Analysis*, 1(2): 1–77. doi:[10.21642/JGEA.010201AF](https://doi.org/10.21642/JGEA.010201AF).
- Lluch, C. 1973. "The extended linear expenditure system." *European Economic Review*, 4(1): 21–32. doi:[10.1016/0014-2921\(73\)90028-7](https://doi.org/10.1016/0014-2921(73)90028-7).
- Malcolm, G. 1998. "Adjusting Tax Rates in the GTAP Data Base." Global Trade Analysis Project (GTAP), Purdue University, West Lafayette, IN, GTAP Technical Paper No. 12. <https://www.gtap.agecon.purdue.edu/resources/download/580.pdf>.
- McDougall, R. 2003. "A New Regional Household Demand System for GTAP (Revision 1)." Global Trade Analysis Project (GTAP), Purdue University, West Lafayette, IN, GTAP Technical Paper No. 20. <https://www.gtap.agecon.purdue.edu/resources/download/1593.pdf>.
- Meinshausen, M., T.M.L. Wigley, and S.C.B. Raper. 2011. "Emulating atmosphere-ocean and carbon cycle models with a simpler model, MAGICC6—Part 2: Applications." *Atmospheric Chemistry and Physics*, 11(4): 1457–1471. doi:[10.5194/acp-11-1457-2011](https://doi.org/10.5194/acp-11-1457-2011).
- OECD. 2016. *The Economic Consequences of Outdoor Air Pollution*. doi:<https://doi.org/http://dx.doi.org/10.1787/9789264257474-en>.
- OECD. 1989/1990. "Special Issue on Modelling the Effects of Agricultural Policies." *OECD Economic Studies*, , (13)pp. 49–92.
- Rimmer, M.T., and A.A. Powell. 1992a. "Demand Patterns across the Development Spectrum: Estimates of the AIDADS System." Centre of Policy Studies and the Impact Project, Monash University, Preliminary Working Paper No. OP-75. <http://www.copsmodels.com/ftp/workpapr/op-75.pdf>.
- Rimmer, M.T., and A.A. Powell. 1996. "An implicitly additive demand system." *Applied Economics*, 28(12): 1613–1622. doi:[10.1080/000368496327589](https://doi.org/10.1080/000368496327589).
- Rimmer, M.T., and A.A. Powell. 1992b. "An Implicitly Directly Additive Demand System: Estimates for Australia." Centre of Policy Studies and the Impact Project, Monash University, Preliminary Working Paper No. OP-73. <http://www.copsmodels.com/ftp/workpapr/op-73.pdf>.
- Smeets Křístková, Z., M. van Dijk, and H. van Meijl. 2016. "Projections of long-term food security with R&D driven technical change—A CGE analysis." *NJAS - Wageningen Journal of Life Sciences*, pp. –. doi:[10.1016/j.njas.2016.03.001](https://doi.org/10.1016/j.njas.2016.03.001).
- Surry, Y. 1993. "The Constant Difference of Elasticities Function with Applications to the EC Animal Feed Sector." *Journal of Agricultural Economics*, 44(1): 110–125. doi:[10.1111/j.1477-9552.1993.tb00254.x](https://doi.org/10.1111/j.1477-9552.1993.tb00254.x).
- van der Mensbrugghe, D. 1994. "GREEN: The Reference Manual." OECD, OECD Economics Department Working Papers No. 143, January. doi:[10.1787/054442860076](https://doi.org/10.1787/054442860076).
- van der Mensbrugghe, D. 2011. "LINKAGE Technical Reference Document: Version 7.1." The World Bank, Washington, DC, Processed. <http://go.worldbank.org/7NP2KK1OH0>.
- van der Mensbrugghe, D. 2018. "The Standard GTAP Model in GAMS, Version 7." *Journal of Global Economic Analysis*, 3(1): 1–83. doi:[10.21642/JGEA.030101AF](https://doi.org/10.21642/JGEA.030101AF).

- Varian, H.R. 1992. *Microeconomic Analysis: Third Edition*. W. W. Norton & Company, Inc. <http://books.wwnorton.com/books/detail.aspx?ID=11247>.
- Yu, W., T. Hertel, P. Preckel, and J. Eales. 2003. "Projecting World Food Demand Using Alternative Demand Systems." Global Trade Analysis Project (GTAP), Purdue University, West Lafayette, IN, GTAP Working Paper No. 21. <https://www.gtap.agecon.purdue.edu/resources/download/2873.pdf>.