Energy Technology Systems Analysis Programme

http://www.etsap.org/tools.htm

Documentation for the TIMES Model

PART I

April 2005

Authors:

Richard Loulou
Uwe Remne
Amit Kanudia
Antti Lehtila
Gary Goldstein

General Introduction

This documentation is composed of thee Parts.

Part I comprises eight chapters constituting a general description of the TIMES paradigm, with emphasis on the model's general structure and its economic significance. Part I also includes a simplified mathematical formulation of TIMES, a chapter comparing it to the MARKAL model, pointing to similarities and differences, and chapters describing new model options.

Part II is a comprehensive reference manual intended for the technically minded modeler or programmer looking for an in-depth understanding of the complete model details, in particular the relationship between the input data and the model mathematics, or contemplating making changes to the model's equations. Part II includes a full description of the sets, attributes, variables, and equations of the TIMES model.

Part III describes the GAMS control statements required to run the TIMES model. GAMS is a modeling language that translates a TIMES database into the Linear Programming matrix, and then submits this LP to an optimizer and generates the result files. In addition to the GAMS program, two model interfaces (VEDA-FE and VEDA-BE) are used to create, browse, and modify the input data, and to explore and further process the model's results. The two VEDA interfaces are described in detail in their own user's guides.

PART I: TIMES CONCEPTS AND THEORY

TABLE OF CONTENTS FOR PART I

1 Introduction to the TIMES model

1.1 A brief summary

TIMES (an acronym for The Integrated MARKAL-EFOM[1] System) is an economic model generator for local, national or multi-regional energy systems, which provides a technology-rich basis for estimating energy dynamics over a long-term, multi-period time horizon. It is usually applied to the analysis of the entire energy sector, but may also applied to study in detail single sectors (e.g. the electricity and district heat sector).

Reference case estimates of end-use energy service demands (e.g., car road travel; residential lighting; steam heat requirements in the paper industry; etc.) are provided by the user for each region. In addition, the user provides estimates of the existing stock of energy related equipment in all sectors, and the characteristics of available future technologies, as well as present and future sources of primary energy supply and their potentials.

Using these as inputs, the TIMES model aims to supply energy services at minimum global cost (more accurately at minimum loss of surplus) by simultaneously making equipment investment and operating, primary energy supply, and energy trade decisions, by region. For example, if there is an increase in residential lighting energy service relative to the reference scenario (perhaps due to a decline in the cost of residential lighting, or due to a different assumption on GDP growth), either existing generation equipment must be used more intensively or new – possibly more efficient – equipment must be installed. The choice by the model of the generation equipment (type and fuel) is based on the analysis of the characteristics of alternative generation technologies, on the economics of the energy supply, and on environmental criteria. TIMES is thus a vertically integrated model of the entire extended energy system.

The scope of the model extends beyond purely energy oriented issues, to the representation of environmental emissions, and perhaps materials, related to the energy system. In addition, the model is admirably suited to the analysis of energy-environmental policies, which may be represented with accuracy thanks to the explicitness of the representation of technologies and fuels in all sectors.

In TIMES – like in its MARKAL forebear – the quantities and prices of the various commodities are in equilibrium, i.e. their prices and quantities in each time period are such that the suppliers produce exactly the quantities demanded by the consumers. This equilibrium has the property that the total surplus is maximized.

1.2 Using the TIMES model

The TIMES model is particularly suited to the *exploration* of possible energy futures based on contrasted *scenarios*. Given the long horizons simulated with TIMES, the

[1] MARKAL (MARket ALlocation model, Fishbone et al, 1981, 1983, Berger et al. 1992) and EFOM (Van Voort et al, 1984) are two bottom-up energy models which inspired the structure of TIMES.

scenario approach is really the only choice (whereas for the shorter term, econometric methods may provide useful projections). Scenarios, unlike forecasts, do not pre-suppose knowledge of the main drivers of the energy system. Instead, a scenario consists of a set of *coherent assumptions* about the future trajectories of these drivers, leading to a coherent organization of the system under study. A scenario builder must therefore carefully test the assumptions made for internal coherence, via a credible *storyline*.

In TIMES, a complete scenario consists of four types of inputs: energy service demands, primary resource potentials, a policy setting, and the descriptions of a set of technologies. We now present a few comments on each of these four components.

1.2.1 The Demand Component of a TIMES scenario

In the case of the TIMES model demand drivers (population, GDP, family units, etc.) are obtained externally, via other models or from accepted other sources. As one example, the TIMES global model constructed for the EFDA[2] used the GEM-E3[3] general equilibrium model to generate a set of *coherent* (total and sectoral) GDP growth rates in the various regions. Note that GEM-E3 itself uses other drivers as inputs in order to derive GDP trajectories. These GEM-E3 drivers consist of measures of technological progress, population, degree of market competitiveness, and a few other perhaps qualitative assumptions. For population and household projections, both GEM-E3 and TIMES used the same exogenous sources (IPCC, Nakicenovic 2000, Moomaw and Moreira, 2001). Other approaches may be used to derive TIMES drivers, whether via models or other means.

For the EFDA model, the main drivers were: Population, GDP, GDP per capita, number of households, and sector GDP. For sectoral TIMES models, the demand drivers may be different depending on the system boundaries.

Once the drivers for a TIMES model are determined and quantified the construction of the reference demand scenario requires computing a set of energy service demands over the horizon. This is done by choosing elasticities of demands to their respective drivers, in each region, using the following general formula:

$$Demand = Driver^{Elasticity}$$

As mentioned above, the demands are provided for the reference scenario. However, when the model is run for alternate scenarios (for instance for an emission constrained case, or for a set of alternate technological assumptions), it is likely that the demands will be affected. TIMES has the capability of estimating the response of the demands to the changing conditions of an alternate scenario. To do this, the model requires still another set of inputs, namely the assumed elasticities of the demands to their own prices. TIMES

[2] EFDA: European Fusion Development Agreement
[3] GEM-E3 General Equilibrium Model for Economy, Energy and Environment

is then able to endogenously adjust the demands to the alternate cases without exogenous intervention. In fact, the TIMES model is driven not by demands but by *demand curves*.

To summarize: the TIMES demand scenario components consist in a set of assumptions on the drivers (GDP, population, households) and on the elasticities of the demands to the drivers and to their own prices.

1.2.2 The Supply Component of a TIMES Scenario

The second constituent of a scenario is a set of *supply curves* for primary energy and material resources. Multi-stepped supply curves can be easily modeled in TIMES, each step representing a certain potential of the resource available at a particular cost. In some cases, the potential may be expressed as a cumulative potential over the model horizon (e.g. reserves of gas, crude oil, etc), as a cumulative potential over the ressource base (e.g. available areas for wind converters differentiated by velocities, available farmland for biocrops, roof areas for PV installations) and in others as an annual potential (e.g. maximum extraction rates, or for renewable resources the available wind, biomass, or hydro potentials). Note that the supply component also includes the identification of trading possibilities, where the amounts and prices of the traded commodities are determined endogenously (within any imposed limits).

1.2.3 The Policy Component of a TIMES Scenario

Insofar as some policies impact on the energy system, they may become an integral part of the scenario definition. For instance, a No-Policy scenario may perfectly ignore emissions of various pollutants, while alternate policy scenarios may enforce emission restrictions, or emission taxes, etc. The detailed technological nature of TIMES allows the simulation of a wide variety of both micro measures (e.g. technology portfolios, or targeted subsidies to groups of technologies), and broader policy targets (such as general carbon tax, or permit trading system on air contaminants). A simpler example might be a nuclear policy that limits the future capacity nuclear plants. Another example might be the imposition of fuel taxes, or of industrial subsidies, etc.

1.2.4 The Techno-economic component of a TIMES Scenario

The fourth and last constituent of a scenario is the set of technical and economic parameters assumed for the transformation of primary resources into energy services. In TIMES, these techno-economic parameters are described in the form of technologies (or processes) that transform some commodities into others (fuels, materials, energy services, emissions). In TIMES, some technologies may be imposed and others may simply be available for the model to choose. The quality of a TIMES model rests on a rich, well developed set of technologies, both current and future, for the model to choose from. The emphasis put on the technological database is one of the main distinguishing factors of the class of Bottom-up models, to which TIMES belongs. Other classes of models will tend to emphasize other aspects of the system (e.g. interactions with the rest of the economy) and treat the technical system in a more succinct manner via aggregate production functions.

Remark: two scenarios may differ in all or in only some of their components. For instance, the same demand scenario may very well lead to multiple scenarios by varying the primary resource potentials and/or technologies and/or policies, insofar as the alternative scenario assumptions do not alter the basic demand inputs (Drivers and Elasticities). The scenario builder must always be careful about the overall coherence of the various assumptions made on the four components of a scenario.

Organization of PART I

Chapter 2 provides a general overview of the representation in TIMES of the Reference Energy System (RES) of a typical region or country, focusing on its basic elements, namely technologies and commodities. Chapter 3 discusses the economic rationale of the model, and Chapter 4 presents a streamlined representation of the Linear Programming problem used by TIMES to compute the equilibrium. Chapter 5 contains a comparison of the respective features of TIMES and MARKAL, intended primarily for users already familiar with MARKAL, while Chapter 6 describes in detail the elastic demand feature and other economic and mathematical properties of the TIMES equilibrium. Chapters 7 and 8, respectively describe two model options: Lumpy Investments (LI), and Endogenous Technological Learning (ETL).

2 The basic structure of the TIMES model

It is useful to distinguish between a model's *structure* and a particular instance of its implementation. A model's structure exemplifies its fundamental approach for representing and analyzing a problem—it does not change from one implementation to the next. All TIMES models exploit an identical mathematical structure. However, because TIMES is data[4] driven, each (regional) model will vary according to the data inputs. For example, in a multi-region model one region may, as a matter of user data input, have undiscovered domestic oil reserves. Accordingly, TIMES generates technologies and processes that account for the cost of discovery and field development. If, alternatively, user supplied data indicate that a region does not have undiscovered oil reserves no such technologies and processes would be included in the representation of that region's Reference Energy System (RES, see sections 2.3 and 2.4). Due to this property TIMES can also be called a model generator that, based on the input information provided by the modeler, generates an instance of a model. In the following, if not stated otherwise, the expression model is used with two meanings: the instance of a TIMES model or more generally the model generator TIMES.

The structure of TIMES is ultimately defined by variables and equations determined from the data input provided by the user. This information collectively defines each TIMES regional model database, and therefore the resulting mathematical representation of the RES for each region. The database itself contains both qualitative and quantitative data. The *qualitative data* includes, for example, lists of energy carriers, the technologies that the modeler feels are applicable (to each region) over a specified time horizon, as well as the environmental emissions that are to be tracked. This information may be further classified into subgroups, for example energy carriers may be split by type (e.g., fossil, nuclear, renewable, etc). *Quantitative data*, in contrast, contains the technological and economic parameter assumptions specific to each technology, region, and time period. When constructing multi-region models it is often the case that a technology may be available for use in two distinct regions; however, cost and performance assumptions may be quite different (i.e., consider a residential heat pump in Canada versus the same piece of equipment in China). This chapter discusses both qualitative and quantitative assumptions in the TIMES modeling system.

The TIMES energy economy is made up of producers and consumers of commodities such as energy carriers, materials, energy services, and emissions. TIMES, like most equilibrium models, assumes competitive markets for all commodities. The result is a supply-demand equilibrium that maximizes the net total surplus (i.e. the sum of producers' and consumers' surpluses) as will be fully discussed in chapters 3 and 6. TIMES may, however, depart from perfectly competitive market assumptions by the introduction of user-defined explicit constraints, such as limits to technological penetration, constraints on emissions, exogenous oil price, etc. Market imperfections can also be introduced in the form of taxes, subsidies and hurdle rates.

---

[4] Data in this context refers to parameter assumptions, technology characteristics, projections of energy service demands, etc. It does not refer to historical data series.

Operationally, a TIMES run configures the *energy system* (of a *set of regions*) over a certain *time horizon* in such a way as to *minimize the net total cost* (or equivalently *maximize the net total surplus*) of the system, while satisfying a number of *constraints*. TIMES is run in a dynamic manner, which is to say that all investment decisions are made in each period with full knowledge of future events. The model is said to have *perfect foresight* (or to be *clairvoyant*). In addition to time-periods (which may be of variable length), there are time divisions within a year, also called *time-slices,* which may be defined at will by the user (see Figure 2.1). For instance, the user may want to define seasons, day/night, and/or weekdays/weekends. Time-slices are especially important whenever the mode and cost of production of an energy carrier at different times of the year are significantly different. This is the case for instance when the demand for an energy form fluctuates across the year and a variety of technologies may be chosen for its production. The production technologies may themselves have different characteristics depending on the time of year (e.g. wind turbines or run-of-the-river hydro plants). In such cases, the matching of supply and demand requires that the activities of the technologies producing and consuming the commodity be tracked at each time slice. Examples of commodities requiring time-slicing may include electricity, district heat, natural gas, industrial steam, and hydrogen. Two additional reasons for defining sub yearly time slices are a) the fact that the commodity is expensive (or even impossible) to store (thus requiring that production technologies be suitably activated in each time slice to match the demand), and b) the existence of an expensive infrastructure whose capacity should be sufficient to bear the peak demand for the commodity. The net result of these chracteristics is that the deployment in time of the various production technologies may be very different in different time-slices, and furthermore that specific investment decisions be taken to insure adequate reserve capacity at peak.

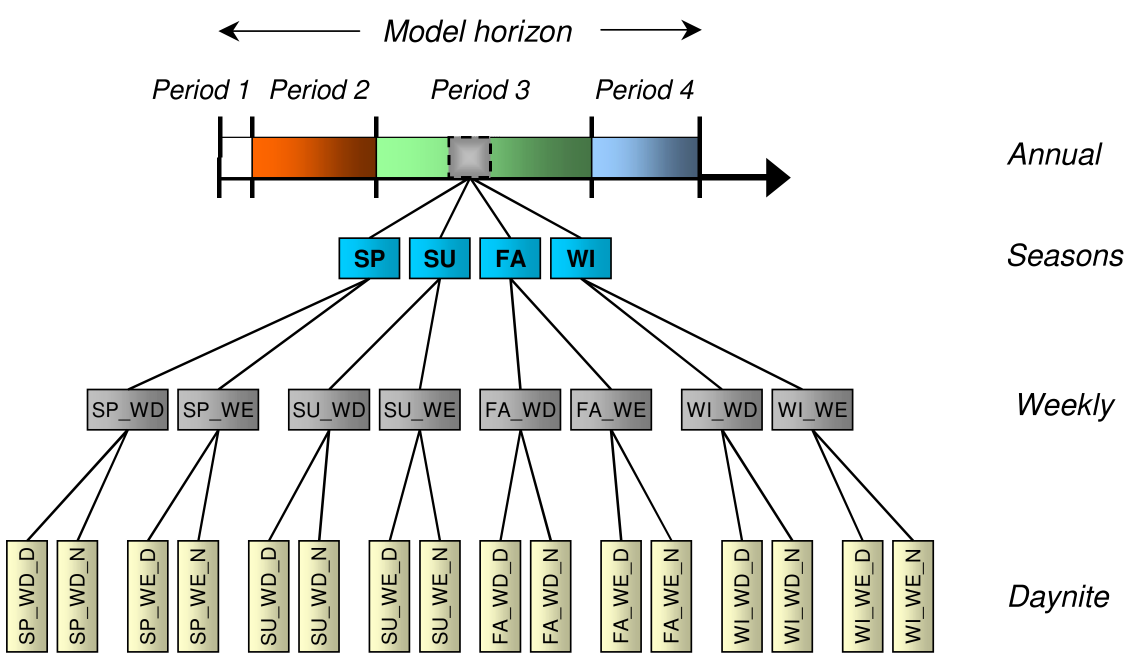

Figure 2.1: Example of a timeslice tree

2.1 Time horizon

The time horizon is divided into a (user-chosen) number of time-periods, each model period containing a (possibly different) number of years. For TIMES each year in a given period is considered identical, except for the cost objective function which differentiates between payments in each year of a period. For all other quantities (capacities, commodity flows, operating levels, etc) any model input or output related to period ***t*** applies to each of the years in that period, with the exception of investment variables, which are usually made only once in a period[5],. In this respect, TIMES is similar to MARKAL but differs from the approach used in EFOM, where capacities and flows were assumed to evolve linearly between so-called milestone years.

The initial period is usually considered a past period, over which the model has no freedom, and for which the quantities of interest are all fixed by the user at their historical values. It is often advised to choose an initial period consisting of a single year, in order to facilitate calibration to standard energy statistics. Calibration to the initial period is one of the more important tasks required when setting up a TIMES model. The main variables to be calibrated are: the capacities and operating levels of all technologies, as well as the extracted, exported, imported, produced, and consumed quantities for all energy carriers, and the emissions if modeled.

In TIMES years preceding the first period also play a role. Although no explicit variables are defined for these years, data may be provided by the modeler on past investments. Note carefully that the specification of past investments influences not only the initial period's calibration, but also the model's behavior over several future periods, since the past investments provide residual capacity in several years within the modeling horizon proper.

2.2 Decoupling of data and model horizon

In TIMES, special efforts have been made to de-couple the specification of data from the definition of the time periods for which a model is run. Two TIMES features facilitate this decoupling.

First, the fact that investments made in past years are recognized by TIMES makes it much easier to modify the choice of the initial and subsequent periods without major revisions of the database.

Second, the specification of process and demand input data in TIMES is made by specifying the *years* when the data apply, and the model takes care of interpolating and extrapolating the data to represent the particular *periods* chosen by the modeler for a particular model run.

---

[5] There are exceptional cases when an investment must be repeated more than once in a period, namely when the period is so long that it exceeds the technical life of the investment. These cases are described in detail in section 5.2 of PART II.

These two features combine to make a change in the definition of periods quite easy and error-free. For instance, if a modeler decides to change the initial year from 1995 to 2005, and perhaps change the number and durations of all other periods as well, only one type of data change is needed, namely to define the investments made from 1995 to 2004 as past investments. All other data specifications need not be altered[6]. This feature represents a great simplification of the modeler's work. In particular, it enables the user to define time periods that have varying lengths, without changing the input data.

2.3 The RES concept

The TIMES energy economy consists of three types of entities:

- *Technologies* (also called *processes*) are representations of physical devices that transform commodities into other commodities. Processes may be primary sources of commodities (e.g. mining processes, import processes), or transformation activities such as conversion plants that produce electricity, energy-processing plants such as refineries, end-use demand devices such as cars and heating systems, etc,
- *Commodities* consisting of energy carriers, energy services, materials, monetary flows, and emissions. A commodity is generally produced by some process(es) and/or consumed by other process(es), and
- *Commodity flows,* that are the links between processes and commodities. A flow is of the same nature as a commodity but is attached to a particular process, and represents one input or one output of that process.

It is helpful to picture the relationships among these various entities using a network diagram, referred to as a Reference Energy System (RES). In TIMES, the RES processes are represented as boxes and commodities as vertical lines. Commodity flows are represented as links between process boxes and commodity lines. Using graph theory terminology, a RES is an oriented graph, where both the processes and the commodities are the *nodes* of the graph. They are interconnected by the flows, which are the *arcs* of the graph. Each arc (flow) is oriented and links exactly one process node with one commodity node. Such a graph is called bi-partite, since its set of nodes may be partitioned into two subsets and there are no arcs directly linking two nodes in the same subset.

Figure 2.2 depicts a small portion of a hypothetical RES containing a single energy service demand, namely residential space heating. There are three end-use space heating technologies using the gas, electricity, and heating oil energy carriers (commodities), respectively. These energy carriers in turn are produced by other technologies, represented in the diagram by one gas plant, three electricity-generating plants (gas fired, coal fired, oil fired), and one oil refinery. To complete the production chain on the primary energy side, the diagram also represents an extraction source for natural gas, an

[6] However, if the horizon has been lengthened, additional data for the new years at the end of the horizon must of course be provided, unless the original database horizon already covers the new model horizon.

extraction source for coal, and two sources of crude oil (one extracted domestically and then transported by pipeline, and the other one imported). This simple RES has a total of 13 commodities and 13 processes. These elements form Note that in the RES every time a commodity enters/leaves a process (via a particular flow) its name is changed (e.g., wet gas becomes dry gas, crude becomes pipeline crude). This simple rule enables the inter-connections between the processes to be properly maintained throughout the network.

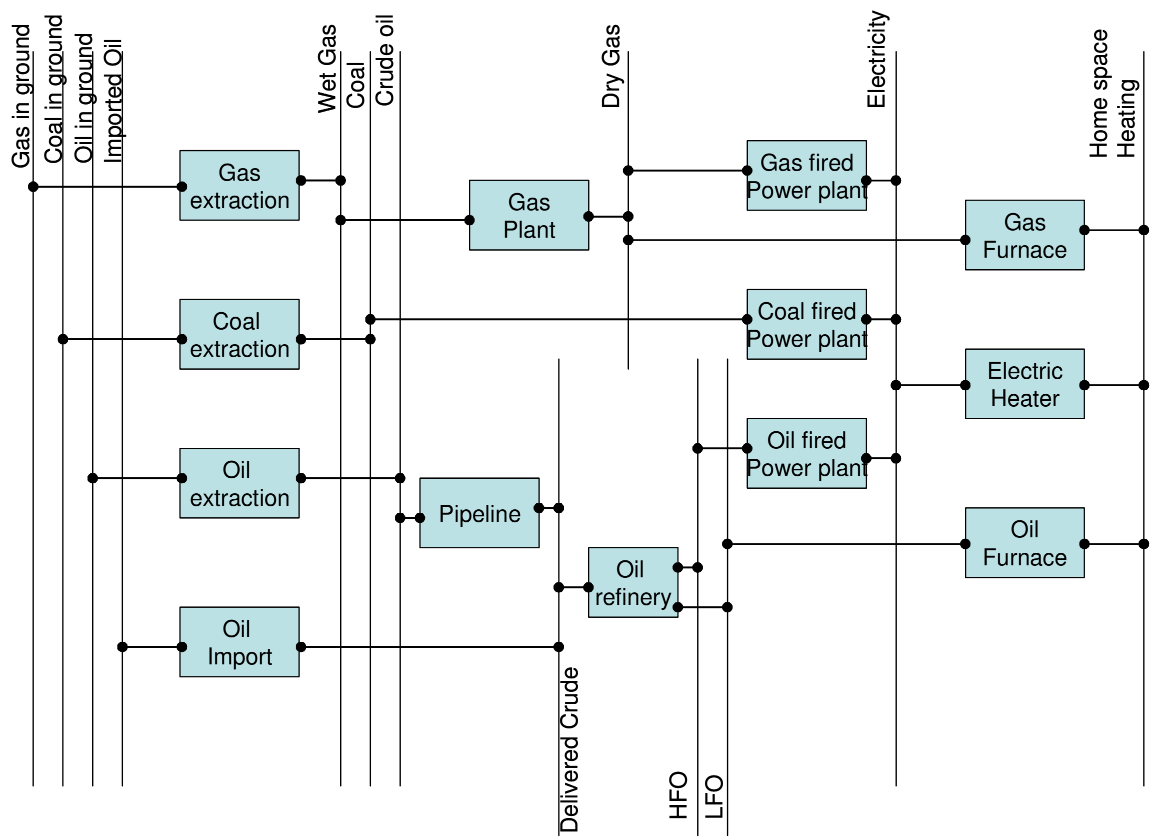

Figure 2.2. Partial view of a simple Reference Energy System (all arcs are oriented left to right)

To organize the RES, and inform the modeling system of the nature of its components, the various technologies, commodities, and flows may be classified into *sets*. Each TIMES set regroups components of a similar nature. The entities belonging to a set are referred to as *members, items* or *elements* of that set. The same item may appear in multiple technology or commodity sets. While the topology of the RES can be represented by a multi-dimensional network, which maps the flow of the commodities to the various technologies, the set membership conveys the nature of the individual components and is often most relevant to post-processing (reporting) than influencing the model structure itself.

Contrary to MARKAL, TIMES has relatively few sets for forming process or commodity groups. In MARKAL the processes are differentiated depending on whether they are sources, conversion processes, end-use devices, etc., and processes in each set have their own specialized attributes. In TIMES most processes are endowed with essentially the same attributes (with the exceptions of storage and inter-regional exchange processes), and unless the user decides otherwise (e.g. by providing values for some attributes and ignoring others), they have the same variables attached to them, and must obey similar constraints. Therefore, the differentiation between the various species of processes or commodities is made through data specification only, thus eliminating the need to define specialized membership sets (unless desired for processing results). Most of the TIMES features (e.g. sub-annual time-slice resolution, vintaging) are available for all processes and the modeler chooses the features being assigned to a particular process by specifying a corresponding indicator set (e.g. PRC_TSL, PRC_VINT).

However, the TIMES commodities are still classified into several *Major Groups*. There are five such groups: energy carriers, materials, energy services, emissions, and monetary flows. The use of these groups is essential in the definition of some TIMES constraints, as discussed in chapter 4.

2.4 Overview of the TIMES attributes

TIMES has some attributes that were not available in MARKAL. More importantly, some attributes correspond to powerful new features that confer to TIMES additional flexibility. The complete list of attributes is shown in PART II, and we provide below only succinct comments on the types of attribute attached to each entity of the RES or to the RES as a whole.

Attributes may be cardinal (e.g. numbers) or ordinal (e.g. sets). For example, some of ordinal attributes are defined for process to describe subsets of flows that are then used to construct specific flow constraints. Subsection 4.4 describes such flow constraints,a nd Chapter 2 of PART II gives the complete list of TIMES sets.

The cardinal attributes are usually called *parameters*. We give below a brief idea of the types of parameters available in the TIMES model generator.

2.4.1 Parameters associated with processes

TIMES process-oriented parameters fall into three general categories. First are *technical parameters* that include efficiency, availability factor(s), commodity consumptions per unit of activity, shares of fuels per unit activity, technical life of the process, construction lead time, dismantling lead-time and duration, amounts of the commodities consumed (respectively released) by the construction (respectively dismantling) of one unit of the process, and contribution to the peak equations. The efficiency, availability factors, and commodity inputs and outputs of a process may be defined in several flexible ways depending on the desired process flexibility, on the time-slice resolution chosen for the process and on the time-slice resolution of the commodities involved. Certain parameters are only relevant to special processes, such as storage processes or processes that implement trade between regions.

The other class of process parameters is *economic and policy parameters* that include a variety of costs attached to the investment, dismantling, maintenance, and operation of a process. In addition, taxes and subsidies may be defined in a very flexible manner. Other economic parameters are the economic life of a process (which is the time during which the investment cost of a process is amortized, which may differ from the operational lifetime) and the process specific discount rate, also called *hurdle rate*, both of which serve to calculate the annualized payments on the process investment cost.

Finally, the modeler may impose a variety of *bounds* (upper, lower, equality) on the investment, capacity, and activity of a process.

Note that many process parameters may be *vintaged* (i.e. dependent upon the date of installation of new capacity), and furthermore may be defined as being dependent on the *age* of the technology. The latter feature is implemented by means of special data grouped under the SHAPE parameter, which introduces user-defined shaping indexes that can be applied to age-dependent parameters. For instance, the annual maintenance cost of an automobile could be defined to remain constant for say 3 years and then increase in a linear manner each year after the third year.

2.4.2 Parameters associated with commodities

Commodity-oriented parameters fall into three categories.

Technical parameters associated with commodities include overall efficiency (for instance grid efficiency), and the time-slices over which that commodity is to be tracked. For demand commodities, in addition the annual projected demand and load curves (if the commodity has a subannual time-slice resolution) can be specified.

Economic parameters include additional costs, taxes, and subsidies on the overall or net production of a commodity.. These cost elements are then added to all other (implicit) costs of that commodity. In the case of a demand service, additional parameters define the demand curve (i.e. the relationship between the quantity of demand and its price). These parameters are: the demand’s own-price elasticity, the total allowed range of variation of the demand value, and the number of steps to use for the discrete approximation of the curve.

Policy based parameters include bounds (at each period or cumulative) on the overall or net production of a commodity, or on the imports or exports of a commodity by a region.

In TIMES the net or the total production of each commodity may be explicitly represented by a variable, if needed for imposing a bound or a tax. No such direct possibility was available in MARKAL, although the same result could be achieved via clever modeling.

2.4.3 Parameters attached to commodity flows into and out of processes

A commodity flow (more simply, a *flow*) is an amount of a given commodity produced or consumed by a given process. Some processes have several flows entering or leaving it, perhaps of different types (fuels, materials, demands, or emissions). In TIMES, unlike in MARKAL, each flow has a variable attached to it, as well as several attributes (parameters or sets)

Technical parameters (along with some set attributes), permit full control over the maximum and/or minimum share a given input or output flow may take within the same commodity group. For instance, a flexible turbine may accept oil or gas as input, and the modeler may use a parameter to limit the share of oil to at most 40% of the total fuel input. Other parameters and sets define the amount of certain outflows in relation to certain inflows (e.g., efficiency, emission rate by fuel). For instance, in an oil refinery a parameter may be used to set the total amount of refined products equal to 92% of the total amount of crude oils (s) entering the refinery, or to calculate certain emissions as a fixed proportion of the amount of oil consumed. If a flow has a sub-annual time-slice resolution, a load curve can be specified for the flow[7].

Economic parameters include delivery and other variable costs, taxes and subsidies attached to an individual process flow.

2.4.4 Parameters attached to the entire RES

These parameters include currency conversion factors (in a multi-regional model), region-specific time-slice definitions, a region-specific general discount rate, and reference year for calculating the discounted total cost (objective function). In addition, certain switches control the activation of the data interpolation procedure as well as special model features to be employed (e.g., run with ETL, see chapter 8).

2.5 Process and Commodity classification

Although TIMES does not explicitly differentiate processes or commodities that belong to different portions of the RES (with the exception of storage and trading processes), there are three ways in which some differentiation does occur.

[7] It is possible to define not only load curves for a flow, but also bounds on the share of a flow in a specific time-slice relative to the annual flow, e.g. the flow in the time-slice "WinterDay" has to be at least 10 % of the total annual flow.

First, TIMES does require the definition of Primary Commodity Groups (*pcg*), i.e. subsets of commodities *of the same nature* entering or leaving a process. For each given process, the modeler defines a pcg as a subset of commodities of the same nature, eitehr entering or leaving the process as flows. TIMES uses the pcg to define the activity of the process, and also its capacity. Besides establishing the process activity and capacity, these groups are convenient aids for defining certain complex quantities related to process flows, as discussed in section 4.4 and in PART II.

As noted previously TIMES does not require that the user provide many set memberships. However, the TIMES report step does pass some set declarations to the VEDA-BE result-processing system to facilitate construction of results analysis tables. These include process subsets to distinguish demand devices, energy processes, material processes (by weight or volume), refineries, electric production plants, coupled heat and power plants, heating plants, storage technologies and distribution (link) technologies; and commodity subsets for energy, useful energy demands (split into six aggregate sub-sectors), environmental indicators, and materials.

The third instance of commodity or process differentiation is not embedded in TIMES, but rests on the modeler. A modeler may well want to choose process and commodity names in a judicious manner so as to more easily identify them when browsing through the input database or when examining results. As an example, the World Multi-regional TIMES model developed within ETSAP adopts a naming convention whereby the first three characters denote the sector and the next three the fuel (e.g., light fuel oil used in the residential sector is denoted RESLFO). Similarly, process names are chosen so as to identify the sub-sector or end-use (first three characters), the main fuel used (next three), and the specific technology (last four). For instance, a standard (001) residential water heater (RHW) using electricity (ELC) is named RWHELC001. Naming conventions may thus play a critical role in allowing the easy identification of an element's position in the RES.

Similarly, energy services may be labeled so that they are more easily recognized. For instance, the first letter may indicate the broad sector (e.g. 'T' for transports) and the second letter designate any homogenous sub-sectors (e.g. 'R' for road transport), the third character being free.

In the same fashion, fuels, materials, and emissions are identified so as to immediately designate the sector and sub-sector where they are produced or consumed. To achieve this some fuels have to change names when they change sectors, which is accomplished via processes whose primary role is to change the name of a fuel. In addition, such a process may serve as a bearer of sector wide parameters such as distribution cost, price markup, tax, that are specific to that sector and fuel. For instance, a tax may be levied on industrial distillate use but not on agricultural distillate use, even though the two commodities are physically identical.

3 Economic rationale of the TIMES modeling approach

This chapter provides a detailed economic interpretation of the TIMES and other partial equilibrium models based on maximizing total surplus. Partial equilibrium models have one common feature – they simultaneously configure the production and consumption of commodities (i.e. fuels, materials, and energy services) and their prices. The price of producing a commodity affects the demand for that commodity, while at the same time the demand affects the commodity's price. A market is said to have reached an equilibrium at prices p^* and quantities q^* when no consumer wishes to purchase less than q^* and no producer wishes to produce more than q^* at price p^*. Both p^* and q^* are vectors whose dimension is equal to the number of different commodities being modeled. As will be explained below, when all markets are in equilibrium the total economic surplus is maximized.

The concept of total surplus maximization extends the direct cost minimization approach upon which earlier bottom-up energy system models were based. These simpler models had fixed energy service demands, and thus were limited to minimizing the cost of supplying these demands. In contrast, the TIMES demands for energy services are themselves elastic to their own prices, thus allowing the model to compute a *bona fide* supply-demand equilibrium. This feature is a fundamental step toward capturing the main feedback from the economy to the energy system.

Section 3.1 provides a brief review of different types of energy models. Section 3.2 discusses the economic rationale of the TIMES model with emphasis on the features that distinguish TIMES from other bottom-up models (such as the early incarnations of MARKAL, see Fishbone and Abilock, 1981, Berger et al., 1992, though MARKAL has since been extended beyond these early versions). Section 3.3 describes the details of how price elastic demands are modeled in TIMES, and section 3.4 provides additional discussion of the economic properties of the model.

3.1. A brief classification of energy models

Many energy models are in current use around the world, each designed to emphasize a particular facet of interest. Differences include: economic rationale, level of disaggregation of the variables, time horizon over which decisions are made (and which is closely related to the type of decisions,i.e. only operational planning or also investment decisions), and geographic scope. One of the most significant differentiating features among energy models is the degree of detail with which commodities and technologies are represented, which will guide our classification of models in two major classes.

3.1.1 'Top-Down' Models

At one end of the spectrum are aggregated *General Equilibrium* (GE) models. In these each sector is represented by a production function designed to simulate the potential substitutions between the main factors of production (also highly aggregated into a few variables such as: energy, capital, and labor) in the production of each sector's output. In this model category are found a number of models of national or global energy systems.

These models are usually called “Top-Down”, because they represent an entire economy via a relatively small number of aggregate variables and equations. In these models, production function parameters are calculated for each sector such that inputs and outputs reproduce a single base historical year.[8] In policy runs, the mix of inputs[9] required to produce one unit of a sector’s output is allowed to vary according to user-selected elasticities of substitution. Sectoral production functions most typically have the following general form:

$$\boldsymbol{X}_s = \boldsymbol{A}_0\left(\boldsymbol{B}_K \cdot \boldsymbol{K}_s^{\rho} + \boldsymbol{B}_L \cdot \boldsymbol{L}_S^{\rho} + \boldsymbol{B}_E \cdot \boldsymbol{E}_S^{\rho}\right)^{1/\rho} \tag{3-1}$$

where $\boldsymbol{X}_S$ is the output of sector *S,*
$\boldsymbol{K}_S$, $\boldsymbol{L}_S$, *and* $\boldsymbol{E}_S$ are the inputs of capital, labor and energy needed to produce one unit of output in sector *S,*
$\boldsymbol{\rho}$ is the elasticity of substitution parameter,
$\boldsymbol{A}_0$ and the $\boldsymbol{B}$’s are scaling coefficients.

The choice of ρ determines the ease or difficulty with which one production factor may be substituted for another: the smaller ρ is (but still greater than or equal to 1), the easier it is to substitute the factors to produce the same amount of output from sector *S*. Also note that the degree of factor substitutability does not vary among the factors of production — the ease with which capital can be substituted for labor is equal to the ease with which capital can be substituted for energy, while maintaining the same level of output. GE models may also use alternate forms of production function (3-1), but retain the basic idea of an explicit substitutability of production factors.

3.1.2 ‘Bottom-Up’ Models

At the other end of the spectrum are the very detailed, *technology explicit* models that focus primarily on the energy sector of an economy. In these models, each important energy-using technology is identified by a detailed description of its inputs, outputs, unit costs, and several other technical and economic characteristics. In these so-called ‘Bottom-Up’ models, a sector is constituted by a (usually large) number of logically arranged technologies, linked together by their inputs and outputs (*commodities*, which may be energy forms or carriers, materials, emissions and/or demand services). Some bottom-up models compute a partial equilibrium via maximization of the total net (consumer and producer) surplus, while others simulate other types of behavior by economic agents, as will be discussed below. In bottom-up models, one unit of sectoral output (e.g., a billion vehicle kilometers, one billion tonnes transported by heavy trucks or one Petajoule of residential cooling service) is produced using a mix of individual technologies’ outputs. Thus the production function of a sector is *implicitly* constructed,

[8] These models assume that the relationships (as defined by the form of the production functions as well as the calculated parameters) between sector level inputs and outputs are in equilibrium in the base year.
[9] Most models use inputs such as labor, energy, and capital, but other input factors may conceivably be added, such as arable land, water, or even technical know-how. Similarly, labor may be further subdivided into several categories.

rather than explicitly specified as in more aggregated models. Such implicit production functions may be quite complex, depending on the complexity of the reference energy system of each sector (sub-RES).

3.1.3 Recent Modeling Advances

While the above dichotomy applied fairly well to earlier models, these distinctions now tend to be somewhat blurred by recent advances in both categories of model. In the case of aggregate top-down models, several general equilibrium models now include a fair amount of fuel and technology disaggregation in the key energy producing sectors (for instance: electricity production, oil and gas supply). This is the case with MERGE[10] and SGM[11], for instance. In the other direction, the more advanced bottom-up models are 'reaching up' to capture some of the effects of the entire economy on the energy system. For instance, the TIMES model has end-use demands (including demands for industrial output) that are sensitive to their own prices, and thus capture the impact of rising energy prices on economic output and *vice versa*. Recent incarnations of technology-rich models are multi-regional, and thus are able to consider the impacts of energy-related decisions on trade. It is worth noting that while the multi-regional top-down models have always represented trade, they have done so with a very limited set of traded commodities – typically one or two, whereas there may be quite a number of traded energy forms and materials in multi-regional bottom-up models.

MARKAL-MACRO (see [9]) is a hybrid model combining the technological detail of MARKAL with a succinct representation of the macro-economy consisting of a single producing sector. Because of its succinct single-sector production function, MARKAL-MACRO is able to compute a general equilibrium in a single optimization step. The NEMS[12] model is another example of a full linkage between several technology rich modules of the various energy subsectors and a set of macro-economic equations, although the linkage here is not as tight as in MARKAL-MACRO, and thus requires iterative resolution methods.

The TIMES model introduces further enhancements over and above those of MARKAL. In TIMES, the horizon may be divided into periods of unequal lengths, thus permitting a more flexible modeling of long horizons: typically, one may adopt short periods in the near-term (the initial period often consists of a single base year), and longer ones in the out years; TIMES includes both technology related variables (as in MARKAL) and flow related variables (as in the EFOM model,van der Voort et. al., 1984), thus allowing the easy creation of more flexible processes and constraints; the expression of the TIMES objective function (total system cost) tracks the payments of investments and other costs much more precisely that in other bottom-up models; and, several other new features of TIMES that are fully discussed in chapters 4 and 5, and in PART II of this documentation.

[10] Model for Evaluating Regional and Global Effects (Manne et al., 1995)
[11] Second Generation Model (Edmonds et al., 1991)
[12] National Energy Modeling System, a detailed integrated equilibrium model of an energy system linked to the economy at large, US Dept of Energy, Energy Information Administration (2000)

In spite of these advances in both classes of models, there remain important differences. Specifically:

- Top-down models encompass macroeconomic variables beyond the energy sector proper, such as wages, consumption, and interest rates, and
- Bottom-up models have a rich representation of the variety of technologies (existing and/or future) available to meet energy needs, and, they often have the capability to track a wide variety of traded commodities.

The Top-down vs. Bottom-up approach is not the only relevant difference among energy models. Among Top-down models, the so-called Computable General Equilibrium models (CGE) described above differ markedly from the *macro econometric models*. The latter do not compute equilibrium solutions, but rather simulate the flows of capital and other monetized quantities between sectors (see e.g. Meade, 1996 on the LIFT model). They use econometrically derived input-output coefficients to compute the impacts of these flows on the main sectoral indicators, including economic output (GDP) and other variables (labor, investments). The sectoral variables are then aggregated into national indicators of consumption, interest rate, GDP, labor, and wages.

Among technology explicit models also, two main classes are usually distinguished: the first class is that of the partial equilibrium models such as MARKAL and TIMES, that use optimization techniques to compute a least cost (or maximum surplus) path for the energy system. The second class is that of *simulation* models, where the emphasis is on representing a system not governed purely by financial costs and profits. In these simulation models (e.g. CIMS, Jaccard et al. 2003), investment decisions taken by a representative agent (firm or consumer) are only partially based on profit maximization, and technologies may capture a share of the market even though their life-cycle cost may be higher than that of other technologies. Simulation models use market-sharing formulas that preclude the easy computation of equilibrium – at least not in a single pass. The SAGE incarnation of the MARKAL model possesses a market sharing mechanism that allows it to reproduce certain behavioral characteristics of observed markets.

The next section focuses on the TIMES model, the most recent advanced partial equilibrium model.

3.2 The TIMES paradigm

Since certain portions of this and the next sections require an understanding of the concepts and terminology of Linear Programming, the reader requiring a brush-up on this topic may first read section 4.5, and then, if needed, some standard textbook on LP, such as Hillier and Lieberman (1990) or Chvàtal (1983). The application of Linear Programming to microeconomic theory is covered in Gale (1960), and in Dorfman, Samuelson, and Solow (1958, and subsequent editions).

A brief description of TIMES would express that it is a:

- *Technology explicit*;
- *Multi-regional*;

- *Partial equilibrium* model;

that assumes:

- *Price elastic demands;* and
- *Competitive markets:* with
- *Perfect foresight (*resulting in *Marginal value Pricing)*

We now proceed to flesh out each of these properties.

3.2.1 A technology explicit model

As already presented in chapter 2 (and described in much more detail in Part II), each technology is described in TIMES by a number of technical and economic parameters. Thus each technology is explicitly identified (given a unique name) and distinguished from all others in the model. A mature TIMES model may include several thousand technologies in all sectors of the energy system (energy procurement, conversion, processing, transmission, and end-uses) in each region. Thus TIMES is not only technology explicit, it is technology rich as well. Furthermore, the number of technologies and their relative topology may be changed at will, purely via data input specification, without the user ever having to modify the model's equations. The model is thus to a large extent *data driven*.

3.2.2 Multi-regional feature

Some existing TIMES models covering the entire energy system include up to 15 regional modules, while some existing sectoral models consist of up to 27 regions. The number of regions in a model is limited only by the difficulty of solving LP's of very large size. The individual regional modules are linked by energy and material trading variables, and by emission permit trading variables, if desired. The trade variables transform the set of regional modules into a single multi-regional (possibly global) energy model, where actions taken in one region may affect all other regions. This feature is of course essential when global as well as regional energy and emission policies are being simulated. Thus a multi-regional TIMES modeled is geographically integrated.

3.2.3 Partial equilibrium properties

TIMES computes a partial equilibrium on energy markets. This means that the model computes both the *flows* of energy forms and materials as well as their *prices*, in such a way that, at the prices computed by the model, the suppliers of energy produce exactly the amounts that the consumers are willing to buy. This equilibrium feature is present at every stage of the energy system: primary energy forms, secondary energy forms, and energy services. A supply-demand equilibrium model has as economic rationale the maximization of the total surplus, defined as the sum of suppliers and consumers surpluses. The mathematical method used to maximize the surplus must be adapted to the particular mathematical properties of the model. In TIMES, these properties are as follows:

- Outputs of a technology are linear functions of its inputs (subsection 3.2.3.1);
- Total economic surplus is maximized over the entire horizon (3.2.3.2), and
- Energy markets are competitive, with perfect foresight (3.2.3.3).

As a result of these assumptions the following additional properties hold:

- The market price of each commodity is equal to its marginal value in the overall system (3.2.3.4), and
- Each economic agent maximizes its own profit or utility (3.2.3.5).

.

3.2.3.1 Linearity

A linear input-to-output relationship first means that each technology represented may be implemented at any capacity, from zero to some upper limit, without economies or diseconomies of scale. In a real economy, a given technology is usually available in discrete sizes, rather than on a continuum. In particular, for some real life technologies, there may be a minimum size below which the technology cannot be implemented (or else at a prohibitive cost), as for instance a nuclear power plant, or a hydroelectric project. In such cases, because TIMES assumes that all technologies may be implemented in any size, it may happen that the model's solution shows some technology's capacity at an unrealistically small size. However, in most applications, such a situation is relatively infrequent and often innocuous, since the scope of application is at the country or region's level, and thus large enough so that small capacities are unlikely to occur.

On the other hand, there may be situations where plant size matters, for instance when the region being modeled is very small. In such cases, it is possible to enforce a rule by which certain capacities are allowed only in multiples of a given size (e.g., build or not a gas pipeline), by introducing integer variables. This option, referred to as Lumpy Investment (LI) is available in TIMES and is discussed in chapter 7. This approach should, however, be used sparingly because it greatly increases solution time. Alternatively and more simply, a user may add user-defined constraints to force to zero any capacities that are clearly too small.

It is the linearity property that allows the TIMES equilibrium to be computed using Linear Programming techniques. In the case where economies of scale or some other non-convex relationship is important to the problem being investigated, the optimization program would no longer be linear or even convex. We shall examine such a case in chapter 8 when discussing Endogenous Technology Learning.

The fact that TIMES's equations are linear, however, *does not mean that production functions behave in a linear fashion*. Indeed, the TIMES production functions are usually highly non-linear (although convex), representing non-linear functions as a stepped sequence of linear functions. As a simple example, a supply of some resource may be represented as a sequence of segments, each with rising (but constant within its interval) unit cost. The modeler defines the 'width' of each interval so that the resulting supply curve may simulate any non-linear convex function. In brief, diseconomies of scale are usually present at the sectoral level.

3.2.3.2 Maximization of total surplus: Price equals Marginal value

The *total surplus* of an economy is the sum of the suppliers' and the consumers' surpluses. The term *supplier* designates any economic agent that produces (and sells) one or more commodities (i.e., in TIMES, an energy form, a material, an emission permit, and/or an energy service). A *consumer* is a buyer of one or more commodities. In TIMES, the suppliers of a commodity are technologies that procure a given commodity, and the consumers of a commodity are technologies or demands that consume a given commodity. Some technologies may be both suppliers and consumers, but not of the same commodity (since a technology never has the same commodity as input and output, with the exception of storage technologies[13]). Therefore, for each commodity the RES defines a set of suppliers and a set of consumers.

It is customary in microeconomics to represent the set of suppliers of a commodity by their *inverse production function*, that plots the marginal production cost of the commodity (vertical axis) as a function of the quantity supplied (horizontal axis). In TIMES, as in other linear optimization models, the supply curve of a commodity is not explicitly expressed as a function of aggregate factor inputs such as capital, labor and energy (as they would in typical production functions used in the economic literature). However, it is a standard result of Linear Programming theory that the inverse supply function is step-wise constant and increasing in each factor (see Figures 2 and 3. for the case of a single commodity[14]). Each horizontal step of the inverse supply function indicates that the commodity is produced by a certain technology or set of technologies in a strictly linear fashion. As the quantity produced increases, one or more resources in the mix (either a technological potential or some resource's availability) is exhausted, and therefore the system must start using a different (more expensive) technology or set of technologies in order to produce additional units of the commodity, albeit at higher unit cost. Thus, each change in production mix generates one step of the staircase production function with a value higher than the preceding step. The width of any particular step depends upon the technological potential and/or resource availability associated with the set of technologies represented by that step.

[13] Even here, a storage process consumes a given commodity at a certain time-slice or period, and restitutes it at a later time-slice or period. Therefore, the output commodity is not identical to the input commodity.
[14] This is so because in Linear Programming the shadow price of a constraint remains constant over a certain interval, and then changes abruptly, giving rise to a stepwise constant functional shape.

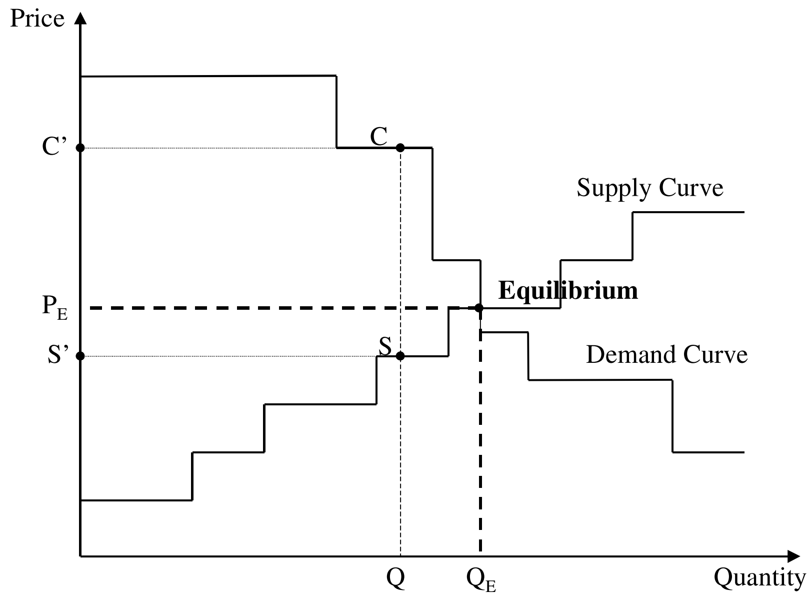

Figure 3.1. Equilibrium in the case of an energy form: the model implicitly constructs both the supply and the demand curves

In a similar manner, each TIMES instance defines a series of inverse demand functions. In the case of demands, two cases are distinguished. First, if the commodity in question is an energy carrier whose production and consumption are endogenous to the model, then its demand function is implicitly constructed within TIMES, and is a step-wise constant, decreasing function of the quantity demanded, as illustrated in Figure 3.1 for a single commodity. If on the other hand the commodity is a demand for an energy service, then its demand curve is defined by the user via the spcification of the own-price elasticity of that demand, and the curve is in this instance a smoothly decreasing curve as illustrated in Figure 3.2[15]. In both cases, the supply-demand equilibrium is at the intersection of the supply function and the demand function, and corresponds to an equilibrium quantity Q_E and an equilibrium price P_E[16]. At price P_E, suppliers are willing to supply the quantity Q_E and consumers are willing to buy exactly that same quantity Q_E. Of course, the TIMES equilibrium concerns many commodities, and the equilibrium is a multi-dimensional analog of the above, where Q_E and P_E are now vectors rather than scalars.

The above description of the TIMES equilibrium is valid for any energy form that is entirely endogenous to TIMES, i.e. an energy carrier, material, or emission permit. In the case of an energy service, TIMES does not implicitly construct the demand function. Rather, the user *explicitly* defines the demand function by specifying its own price

[15] This smooth curve will be discretized later for computational purposes, as described in chapter 6

[16] As may be seen in figures 2 and 3, the equilibrium is not necessarily unique. In the case shown in Figure 2, any point on the vertical segment containing the equilibrium is also an equilibrium, with the same Q_E but a different P_E. In other cases, the multiple equilibria may have the same price and different quantities.

elasticity. Each energy service demand is assumed to have a constant own price elasticity function of the form (see Figure 3.2):

$$D/D_0 = (P/P_0)^E \qquad (3.3\text{-}1)$$

Where $\{D_0, P_0\}$ is a reference pair of demand and price values for that energy service over the forecast horizon, and E is the (negative) own price elasticity of that energy service demand, as chosen by the user (note that though not shown by the notation, this price elasticity may vary over time). The pair $\{D_0, P_0\}$ is obtained by solving TIMES for a reference scenario. More precisely, D_0 is the demand projection estimated by the user in the reference case based upon explicitly defined relationships to economic and demographic drivers, and P_0 is the shadow price of that energy service demand obtained by running the reference case scenario of TIMES.

Using Figure 3.1 as an example, the definition of the suppliers' surplus corresponding to a certain point S on the inverse supply curve is the difference between the total revenue and the total cost of supplying a commodity, i.e. the gross profit. In Figure 3.1, the surplus is thus the area between the horizontal segment SS' and the inverse supply curve. Similarly, the consumers' surplus for a point C on the inverse demand curve, is defined as the area between the segment CC' and the inverse demand curve. This area is a consumer's analog to a producer's profit; more precisely it is the cumulative opportunity gain of all consumers who purchase the commodity at a price lower than the price they would have been willing to pay. For a given quantity Q, the total surplus (suppliers' plus consumers') is simply the area between the two inverse curves situated at the left of Q. It should be clear from Figure 3.1 that the total surplus is maximized exactly when Q is equal to the equilibrium quantity Q_E. Therefore, we may state (in the single commodity case) the following Equivalence Principle:

"The supply-demand equilibrium is reached when the total surplus is maximized"

This is a remarkably useful result, as it leads to a method for computing the equilibrium, as will be see in much detail in Chapter 6. In the multi-dimensional case, the proof of the above statement is less obvious, and requires a certain qualifying property (called the integrability property) to hold (Samuelson, 1952, Takayama and Judge, 1972). One sufficient condition for the integrability property to be satisfied is realized when the cross-price elasticities of any two energy forms are equal, viz.

$$\partial \boldsymbol{P_j} / \partial \boldsymbol{Q_i} = \partial \boldsymbol{P_i} / \partial \boldsymbol{Q_j} \quad \text{for all } \boldsymbol{i}, \boldsymbol{j}$$

In the case of commodities that are energy services, these conditions are trivially satisfied in TIMES because we have assumed zero cross price elasticities. In the case of an energy carrier, where the demand curve is implicitly derived, it is also easy to show that the integrability property is always satisfied[17]. Thus the equivalence principle is valid in all cases.

[17] This results from the fact that in TIMES each price $\boldsymbol{P_i}$ is the shadow price of a balance constraint (see section 4.5), and may thus be (loosely) expressed as the derivative of the objective function $\boldsymbol{F}$ with respect to the right-hand-side of a balance constraint, i.e. $\partial \boldsymbol{F} / \partial \boldsymbol{Q_i}$. When that price is further differentiated with

In summary, the equivalence principle guarantees that the TIMES supply-demand equilibrium maximizes total surplus.The total surplus concept has long been a mainstay of social welfare economics because it takes into account both the surpluses of consumers and of producers.[18]

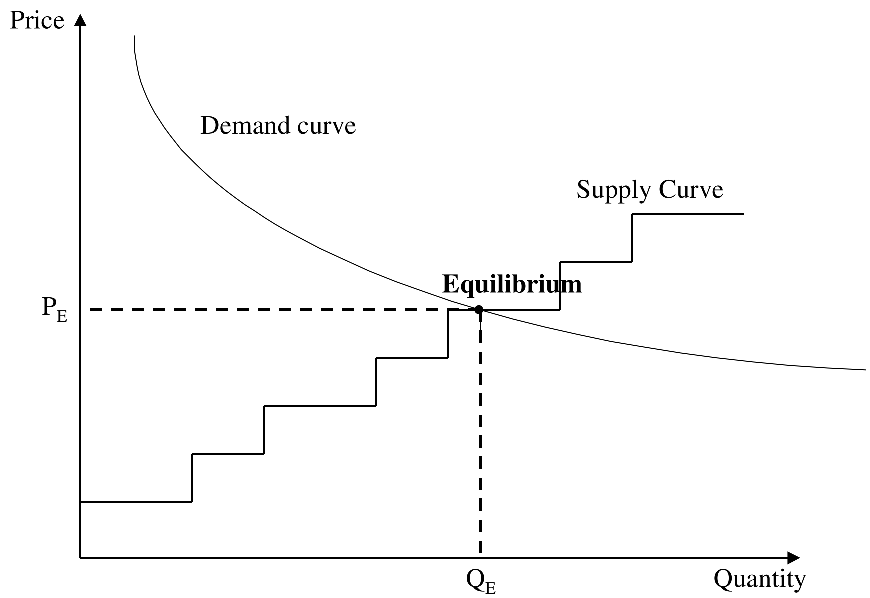

Figure 3.2. Equilibrium in the case of an energy service: the user, explicitly provides the demand curve, usually using a simple functional form

Remark: In older versions of MARKAL, and in several other least-cost bottom-up models, energy service demands are exogenously specified by the modeler, and only the cost of supplying these energy services is minimized. Such a case is illustrated in Figure 3.3 where the "inverse demand curve" is a vertical line. The objective of such models was simply the minimization of the total cost of meeting exogenously specified levels of energy service.

---

respect to another quantity Q_j, one gets $\partial^2 \boldsymbol{F} / \partial \boldsymbol{Q}_i \bullet \partial \boldsymbol{Q}_j$, which, under mild conditions is always equal to $\partial^2 \boldsymbol{F} / \partial \boldsymbol{Q}_j \bullet \partial \boldsymbol{Q}_i$, as desired.

[18] See e.g. Samuelson, P, and W. Nordhaus (1977)

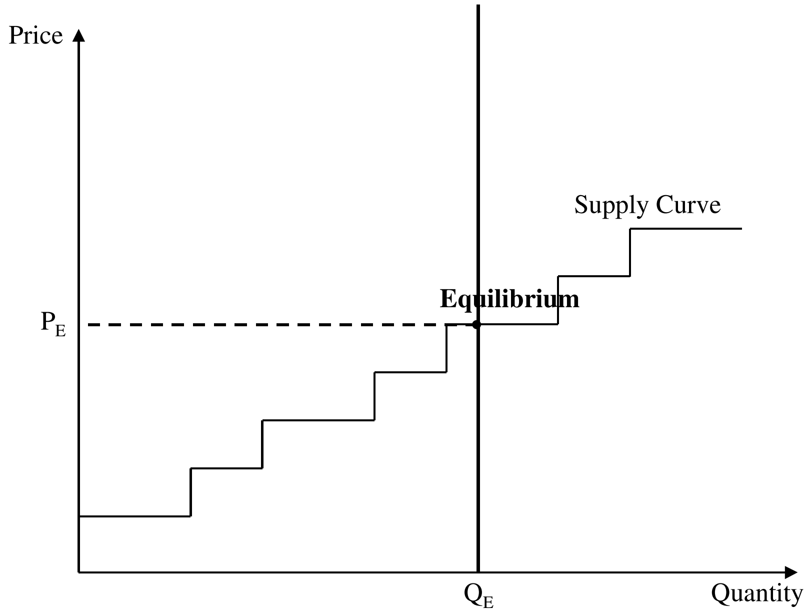

Figure 3.3. Equilibrium when an energy service demand is fixed

3.2.3.3 Competitive energy markets with perfect foresight

Competitive energy markets are characterized by perfect information and atomic economic agents, which together preclude any of them from exercising market power. That is, neither the level any individual producer supplies, nor the level any individual consumer demands, affects the equilibrium market price (because there are many other buyers and sellers to replace them). It is a standard result of microeconomic theory that the assumption of competitive markets entails that the market price of a commodity is equal to its marginal value in the economy. This is of course also verified in the TIMES economy, as discussed in the next subsection.

In TIMES, the perfect information assumption extends to the entire planning horizon, so that each agent has perfect foresight, i.e. complete knowledge of the market's parameters, present and future. Hence, the equilibrium is computed by maximizing total surplus in one pass for the entire set of periods. Such a farsighted equilibrium is also called an *inter-temporal, dynamic* or *clairvoyant* equilibrium.

Note that there are at least two ways in which the perfect foresight assumption may be relaxed: in one variant, agents are assumed to have foresight over a limited portion of the horizon, say one period. Such an assumption of limited foresight is embodied in the SAGE variant of MARKAL. There is for the time being no such variant of the TIMES model. In another variant, foresight is assumed to be imperfect, meaning that agents may only assume probabilities for certain key future events. This assumption is at the basis of the Stochastic Programming option in Standard MARKAL. A Stochastic version of TIMES is at the planning stage.

3.2.3.4 Marginal value pricing

We have seen in the preceding subsections that the TIMES equilibrium occurs at the intersection of the inverse supply and inverse demand curves. It follows that the equilibrium prices are equal to the marginal system values of the various commodities. From a different angle, the duality theory of Linear Programming (section 4.5) indicates that for each constraint of the TIMES linear program there is a dual variable. This dual variable (when an optimal solution is reached) is also called the constraint's *shadow price*[19], and is equal to the marginal change of the objective function per unit increase of the constraint's right-hand-side. For instance (section 4.5), the shadow price of the balance constraint of a commodity (whether it be an energy form, material, a service demand, or an emission) represents the competitive market price of the commodity.

The fact that the price of a commodity is equal to its marginal value is an important feature of competitive markets. Duality theory does not necessarily indicate that the marginal value of a commodity is equal to the marginal cost of *producing* that commodity. For instance, in the equilibrium shown in Figure 3.4 the price does not correspond to *any* marginal supply cost, since it is situated at a discontinuity of the inverse supply curve. In this case, the price is precisely determined by demand rather than by supply, and the term *marginal cost pricing* (so often used in the context of optimizing models) is incorrect. The term *marginal value pricing* is a more appropriate term to use.

It is important to note that marginal value pricing *does not imply that suppliers have zero profit*. Profit is exactly equal to the suppliers' surplus, and Figures 2 through 5 show that it is generally positive. Only the last few units produced may have zero profit, if, and when, their production cost equals the equilibrium price, and even in this case zero profit is not automatic as exemplified in Figure 3.3.

In TIMES the shadow prices of commodities play a very important diagnostic role. If some shadow price is clearly out of line (i.e. if it seems much too small or too large compared to the anticipated market prices), this indicates that the model's database may contain some errors. The examination of shadow prices is just as important as the analysis of the quantities produced and consumed of each commodity and of the technological investments.

[19] The term Shadow Price is often used in the mathematical economics literature, whenever the price is derived from the marginal value of a commodity. The qualifier 'shadow' is used to distinguish the competitive market price from the price observed in the real world, which may be different, as is the case in regulated industries or in sectors where either consumers or producers exercise market power. When the equilibrium is computed using LP optimization, as is the case for MARKAL, the shadow price of each commodity is computed as the dual variable of that commodity's balance constraint, as will be further developed in section 3.4.

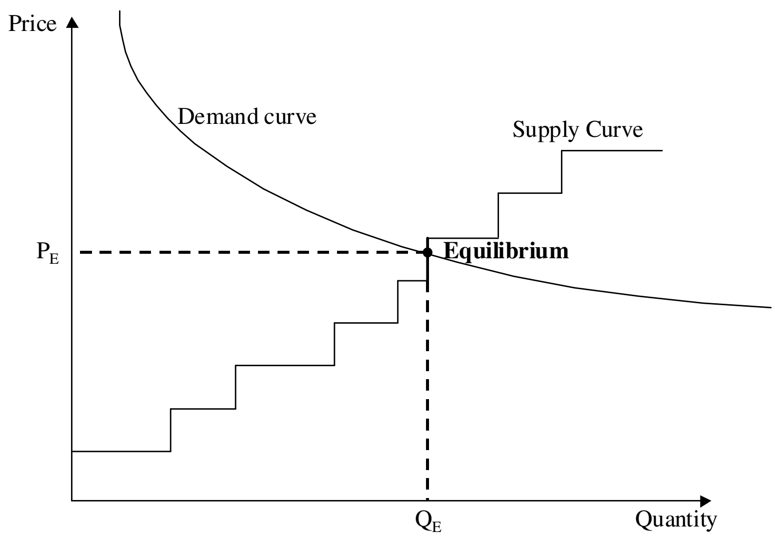

Figure 3.4. Case where the equilibrium price is not equal to any marginal supply cost.

3.2.3.5 Profit maximization: the Invisible Hand

An interesting property may be derived from the assumptions of competitiveness. While the avowed objective of the TIMES model is to maximize the overall surplus, it is also true that each economic agent in TIMES maximizes its own 'profit'. This property is akin to the famous 'invisible hand' property of competitive markets, and may be established rigorously by the following theorem that we state in an informal manner:

Theorem: Let (p,q*) be the pair of equilibrium vectors. If we now replace the original TIMES linear program by one where the commodity prices are fixed at value p*, and we let each agent maximize its own profit, there exists a vector of optimal quantities produced or purchased by the agents that is equal to q*[20].*

This property is important inasmuch as it provides an alternative justification for the class of equilibria based on the maximization of total surplus. It is now possible to shift the model's rationale from a global, societal one (surplus maximization), to a local, decentralized one (individual utility maximization). Of course, the equivalence suggested by the theorem is valid only insofar as the marginal value pricing mechanism is strictly enforced—that is, neither individual producers' nor individual consumers' behaviors affect market prices—both groups are price takers. Clearly, many markets are not

[20] However, the resulting Linear Program has multiple optimal solutions. Therefore, although q* is an optimal solution, it is not necessarily the one found when the modified LP is solved.

competitive in the sense the term has been used here. For example, the behavior of a few, state-owned oil producers has a dramatic impact on world oil prices, that then depart from their marginal system value. Market power[21] may also exist in cases where a few consumers dominate a market. The entire annual crop of a given region's supply of coffee beans may, for example, be purchased by a handful of purchasers who may then greatly influence prices.

[21] An agent has market power if its decisions, all other things being equal, have an impact on the market price. Monopolies and Oligopolies are example of markets where one or several agents have market power.

4 A simplified description of the TIMES optimization program

This chapter contains a simplified formulation of the TIMES Linear Program. A Linear Programming problem consists in the minimization (or maximization) of an *objective function* defined as a mathematical expression of *decision variables*, subject to *constraints* (also called *equations*[22]) also expressed mathematically. Very large instances of Linear Programs involving hundreds of thousands of constraints and variables may be formulated in the GAMS language, and solved via powerful Linear Programming *optimizers*[23]. The Linear Program described in this chapter is much simplified, since it ignores many exceptions and complexities that are not essential to a basic understanding of the principles of the model. Section 4.5 gives additional details on general Linear Programming concepts. The full details of the parameters, variables, objective function, and constraints of TIMES are given in Part II of this documentation.

An optimization problem formulation consists of three types of entities:

- *the decision variables:* i.e. the unknowns, or endogenous quantities, to be determined by the optimization,
- *the objective function*: expressing the criterion to be minimized or maximized, and
- *constraints*: equations or inequalities involving the decision variables that must be satisfied by the optimal solution.

4.1 Indices

The model data structures (sets and parameters), variables and equations use the following indexes:

r: indicates the region
t or v: time period; ***t*** corresponds to the current period, and ***v*** is used to indicate the vintage year of an investment When a process is not vintaged then ***v = t***.
p: process (technology);
s: time-slice; this index is relevant only for user-designated commodities and processes that are tracked at finer than annual level (e.g. electricity, low-temperature heat, and perhaps run-of-river hydro or natural gas, etc.). Time-slice defaults to “ANNUAL”, indicating that a commodity is tracked only annually.
c: commodity (energy, material, emission, demand);

4.2 Decision Variables

[22] This rather improper term includes equality as well as inequality relationships between mathematical expressions.
[23] For more information on optimizers see The Solver Manual of the GAMS A USER'S GUIDE, Anthony Brooke, David Kendrick, Alexander Meeraus, and Ramesh Raman, December 1998.

The decision variables represent the *choices* to be made by the model, i.e. the *unknowns*. The various kinds of decision variables in a TIMES model are:

NCAP(r,v,p)**:** new capacity addition (investment) for technology ***p***, in period ***v*** and region ***r***. For all technologies the ***v*** value corresponds to the vintage of the process*,* i.e. year in which it is invested in. For vintaged technologies (declared as such by the user) the vintage (***v***) information is reflected in other process variables, discussed below. Typical units are PJ/year for most energy technologies, Million tonnes per year (for steel, aluminum, and paper industries), Billion vehicle-kilometers per year (B-vkms/year) or million cars for road vehicles and GW for electricity equipment (1GW=31.536 PJ/year), etc.

CAP(r,v,t,p): installed capacity of process ***p***, in region r and period ***t*** (optionally with vintage ***v***)**.** It represents the total capacity in place in period *t*, considering the residual capacity at the beginning of the modeling horizon and new investments made prior to and including period *t* that have not reached their technical lifetime. Typical units: same as investments. The ***CAP*** variables are actually not explicitly defined in the model, but are derived from the NCAP variables and data on past investments and plant lifetimes.

CAPT(r,t,p): total installed capacity of technology ***p***, in region ***r*** and period ***t***, all vintages together. The ***CAPT*** variables are only defined when some bound or user-constraint are specified for them. They do not enter any other equation.

ACT(r,v,t,p,s): activity level of technology ***p***, in region ***r*** and period ***t*** (optionally vintage ***v*** and time-slice ***s***)**.** Typical units: PJ for all energy technologies. The ***s*** index is relevant only for processes that produce or consume commodities specifically declared as time-sliced. Moreover, it is the process that determines which time slices prevail, via a special attribute. By default, only annual activityis tracked.

FLOW(r,v,t,p,c,s): the quantity of commodity ***c*** consumed or produced by process ***p***, in region ***r*** and period ***t*** (optionally with vintage ***v*** and time-slice ***s***). Typical units: PJ for all energy technologies. The *FLOW* variables confer considerable flexibility to the processes modeled in TIMES, as they allow the user to define flexible processes for which input and/or output flows are not rigidly linked to the process' activity.

SIN(r,v,t,p,c,s)/SOUT(r,v,t,p,c,s): the quantity of commodity ***c*** stored or discharged by storage process ***p,*** in time-slice ***s***, period ***t*** (optionally with vintage ***v)***, and region ***r***.

TRADE(r,t,p,c,s,imp) and TRADE(r,t,p,c,s,exp):) quantity of commodity ***c*** (PJ per year) sold (***exp***) or purchased (***imp***) by region ***r*** through export (resp. import) process p in period ***t*** (optionally in time-slice ***s***). Note that the topology defined for the exchange process ***p*** specifies the traded commodity ***c***, the region ***r***, and the regions ***r'*** with which region ***r*** is trading commodity ***c***. In the case of bi-lateral trading, if it is desired that region ***r*** trade with several other regions, then each such trade requires the definition of a separate bi-lateral exchange process. Note that it is also possible to define multi-lateral trading relationships between region ***r*** and several other regions ***r'*** by defining one of the regions as the common market for trade in commodity ***c***. In this case, the commodity is 'put on the market' and may be bought by any other region participating in the market. This case is convenient for global commodities such as emission permits or crude oil. Finally, exogenous trading may also be modeled by specifying the r' region as an external region. Exogenous trading is required for models that are not global, since exchanges with non modeled regions cannot be considered endogenous.

D(r,t,d): demand for end-use energy service ***d*** in region ***r*** and period ***t.*** In the reference scenario, this variable is fixed by the user. In non-reference scenarios ***D(r,t,d)*** may differ from the reference case demand due to the responsiveness of demands to prices (based on each service demand's own-price elasticity). Note that in this simplified formulation, we do not show the variables used to decompose ***D(r,t,d)*** into a sum of step-wise quantities (see chapter 6 and Part II for details).

Other variables: TIMES has a number of commodity related variables that are not strictly needed but are convenient for reporting purposes and/or for applying certain bounds to them. Examples of such variables are: the total amount produced of a commodity (***COMPRD***), or the total amount consumed of a commodity (***COMCON***).[24]

4.3 TIMES objective function: discounted total system cost

The Surplus Maximization objective is first transformed into an equivalent Cost Minimization objective by taking the negative of the surplus, and calling this the *total system cost*. This practice is in part inspired from historical custom from the days of the fixed demand MARKAL model. The TIMES objective is therefore to minimize the total cost of the system, properly augmented by the 'cost' of lost demand. All cost elements are appropriately discounted to a selected year.

While the TIMES constraints and variables are linked to a *period*, the components of the system cost are expressed for each *year* of the horizon (and even for some years outside the horizon). This choice is meant to provide a smoother, more realistic rendition of the stream of payments in the energy system, as will be discussed below. Each year, the total cost includes the following elements:

- *Capital Costs* incurred for *investing* into and/or *dismantling* processes*;*
- Fixed and variable annual *Operation and Maintenance (O&M) Costs,* and other annual costs occurring during the dismantling of technologies;
- Costs incurred for *exogenous imports* and for domestic resource *production*;
- Revenues from exogenous *exports;*
- *Delivery* costs for required commodities consumed by processes;

[24] These extra variables, as well as the flow variables, add only a moderate computational burden to the optimization process thanks to the use of a *reduction algorithm* to detect and eliminate redundant variables and constraints before solving the LP. These variables and constraints are later reinstated in the solution file for reporting purposes.

- *Taxes* and *subsidies* associated with commodity flows and process activities or investments;
- *Revenues from recuperation of embedded commodities,* accrued when a process's dismantling releases some valuable commodities;
- *Salvage value* of processes and embedded commodities at the end of the planning horizon;
- *Welfare loss* resulting from reduced end-use demands. Chapter 6 presents the mathematical derivation of this quantity.

As already mentioned, in TIMES, special care is taken to precisely track the monetary flows related to process investments and dismantling in each year of the horizon. Such tracking is made complex by several factors:

- First, TIMES recognizes that there may be a lead-time between the beginning and the end of the construction of some large processes, thus spreading the investment installments over several years;
- Second, TIMES also recognizes that for some other processes (e.g. new cars), the investments in new capacity occur progressively over several years of a time period, rather than in one lump amount (as in MARKAL);
- Third, there is the possibility that a certain investment decision made at period ***t*** will have to be repeated more than once during that same period (this will occur if the $\boldsymbol{t^{th}}$ period is long compared to the process life);
- Fourth, TIMES recognizes that there may be dismantling capital costs at the end-of-life of some processes (e.g. a nuclear plant), and these costs, while attached to the investment variable indexed by period ***t***, are actually incurred much later, and
- Finally, TIMES assumes that the payment of any capital cost is spread over an *economic life* that may be different from the *technical life* of the process, and annualized at a different rate than the overall discount rate.

These various TIMES features, while adding precision and realism to the cost profile, also introduce complex mathematical expressions in the objective function. In this simplified formulation, we do not provide much detail on these complex expressions, which are fully described in section 5.1 of Part II. We limit our description to giving general indications on the cost elements composing the objective function, as follows:

- The investment and dismantling costs are transformed into streams of annual payments, computed for each year of the horizon (and beyond, in the case of dismantling costs and recycling revenues), along the lines suggested above;
- A *salvage value* of all investments still active at the end of the horizon (EOH) is calculated and its value is assigned to the (single) year following the EOH;
- The other costs listed above, which are all annual costs, are added to the annualized capital cost payments, minus salvage value, to form the ***ANNCOST*** quantity (below), and
- TIMES then computes for each region a total net present value of the stream of annual costs, discounted to a user selected reference year. These regional discounted costs are then aggregated into a single total cost, which constitutes the objective function to be minimized by the model in its equilibrium computation.

$$NPV = \sum_{r=1}^{R} \sum_{y \in YEARS} (1+d_{r,y})^{REFYR-y} \bullet ANNCOST(r,y)$$

where:

NPV is the net present value of the total cost for all regions (the TIMES objective function);
ANNCOST(r,y) is the total annual cost in region *r* and year *y*;
$d_{r,y}$ is the general discount rate;
REFYR is the reference year for discounting;
YEARS is the set of years for which there are costs, including all years in the horizon, plus past years (before the initial period) if costs have been defined for past investments, plus a number of years after EOH where some investment and dismantling costs are still being incurred, as well as the Salvage Value;
R is the set of regions in the area of study, and

As already mentioned, the exact computation of ***ANNCOST*** is quite complex and is postponed until PART II, section 5.1.

4.4 Constraints

While minimizing total discounted cost the TIMES model must satisfy a large number of constraints (the so-called *equations* of the model) which express the physical and logical relationships that must be satisfied in order to properly depict the associated energy system. TIMES constraints are of several kinds. We list and briefly discuss the main types of constraints. A full description is given in Part II. If any constraint is not satisfied, the model is said to be *infeasible*, a condition caused by a data error or an over-specification of some requirement.

In the descriptions of the equations that follow, the equation and variable names (and their indexes) are in ***bold italic*** type, and the parameters (and their indexes), corresponding to the input data, are in regular *italic* typeset. Furthermore, some parameter indexes have been omitted in order to provide a streamlined presentation.

4.4.1 Capacity Transfer (conservation of investments)

Investing in a particular technology increases its installed capacity for the duration of the physical life of the technology. At the end of that life, the total capacity for this technology is decreased by the same amount. When computing the available capacity in some time period, the model takes into account the capacity resulting from all investments up to that period, some of which may have been made prior to the initial period but are still in operating condition (embodied by the residual capacity of the technology), and others that have been decided by the model at, or after, the initial period, up to and including the period in question.

The total available capacity for each technology *p,* in region *r*, in period *t* (all vintages), is equal to the sum of investments made by the model at past and current periods, and whose physical life has not yet ended, plus capacity in place prior to the modeling horizon that is still available. The exact formulation of this constraint is made quite complex by the fact that TIMES accepts variable time periods, and therefore the end of life of an investment may well fall in the middle of a future time period. We ignore here these complexities and provide a streamlined version of this constraint. Full details are shown in Part II.

EQ_CPT(r,t,p) - Capacity transfer

> ***CAPT(r,t,p) = Sum*** *{over all periods t' preceding or equal to t such that t-t'<LIFE(r,t',p) of* ***NCAP(r,t',p)****}* *+ RESID(r,t,p)* **(4-1)**
>
> where *RESID(r,t,p)* is the (exogenously provided) capacity of technology *p* due to investments that were made prior to the initial model period and still exist in region *r* at time *t*.

4.4.2 Definition of process activity variables

Since TIMES recognizes activity variables as well as flow variables, it is necessary to relate these two types of variables. This is done by introducing a constraint that equates an overall activity variable, ***ACT(r,v,t,p,s)***, with the appropriate set of flow variables, ***FLOW(r,v,t,p,c,s)***, properly weighted This is accomplished by first identifying the group of commodities that defines the activity (and thereby the capacity as well) of the process. In a simple process, one consuming a single commodity and producing a single commodity, the modeler simply chooses one of these two flows to define the activity, and thereby the process normalization (input or output). In more complex processes, with several commodities (perhaps of different types) as inputs and/or outputs, the definition of the activity variable requires first to choose the *primary commodity group (pcg)* that will serve as the activity-defining group. For instance, the *pcg* may be the group of energy carriers, or the group of materials of a given type, or the group of GHG emissions, etc. The modeler then identifies whether the activity is defined via inputs or via outputs that beong to the selected *pcg*. Conceptually, this leads to the following relationship:

EQ_ACTFLO(r,v,t,p,s) – Activity definition

> ***ACT(r,v,t,p,s)*** = **SUM**{*c in pcg of* ***FLOW(r,v,t,p,c,s)*** */ ACTFLO(r,v, p,c)*} **(4-2)**

where *ACTFLO(r,v, p,c)* is a conversion factor (often equal to 1) from the activity of the process to the flow of a particular commodity.

4.4.3 Use of capacity

In each time period the model may use some or all of the installed capacity according to the Availability Factor (AF) of that technology. Note that the model may decide to use *less* than the available capacity during certain time-slices, or even throughout one or more whole periods, if such a decision contributes to minimizing the overall cost. Optionally, there is a provision for the modeler to force specific technologies to use their capacity to their full potential.

For each technology *p*, period *t*, vintage *v*, region *r*, and time-slice *s*, the activity of the technology may not exceed its available capacity, as specified by a user defined availability factor.

EQ_CAPACT (r,v,t,p,s) - Use of capacity

ACT (r,v,t,p,s) ≤ or = *AF(r,v,t,p,s)* CAPUNIT(r,p))*FR(r,s)** ***CAP(r,v,t,p)*** **(4-3)**

Here *CAPUNIT(r,p)* is the conversion factor between units of capacity and activity (often equal to 1, except for power plants). The *FR(r,s)* parameter is equal to the duration of timeslice s. The availability factor *AF* also serves to indicate the nature of the constraint as an inequality or an equality. In the latter case the capacity is forced to be fully utilized,. Note that the ***CAP(r,v,t,p)*** variable is not explicitly defined in TIMES. Instead it is replaced in (4-3) by a fraction (less than or equal to 1) of the investment variable ***NCAP(r,v,p)***[25]a sum of past investments that are still operating, as in equation (4-1).

Example: *a coal fired power plant's activity in any time-slice is bounded above by 80% of its capacity, i.e.* ***ACT (r,v,t,p,s)*** *≤ 0 .8*31.536* **** CAP(r,v,t,p),*** *where CAPUNIT(r,p) = 31.536 is the conversion factor between the units of the capacity variable (GW) and the activity-based capacity unit (PJ/a) The activity-based capacity unit is obtained from the activity unit(PJ) by division by a denominator of one year.*

The *s* index of the *AF* coefficient in equation (4-3) indicates that the user may specify time-sliced dependency on the availabilityof the installed capacity of some technologies, if desirable. This is especially needed when the operation of the equipment depends on the availability of a resource that cannot be stored, such as wind and sun, or that can be only partially stored, such as water in a reservoir. In other cases, the user may provide an *AF* factor that does not depend on *s*, which is then applied to the entire year. The operation profile of a technology within a year, if the technology has a sub-annual

[25] That fraction is equal to 1 if the technical life of the investment made in period ***v*** fully covers period ***t***. It is less than 1 (perhaps 0) otherwise. Chapter 5 of PART II provides details of the computation of that fraction.

process resolution, is determined by the optimization routine. The number of ***EQ_CAPACT*** constraints is at least equal to the number of time-slices at which the equipment operates. For technologies with only an annual characterization the number of constraints is reduced to one per period (where *s="ANNUAL"*).

4.4.4 Commodity Balance Equation:

In each time period, the production by a region plus imports from other regions of each commodity must balance the amount consumed in the region or exported to other regions. In TIMES, the sense of each balance constraint (≥ or =) is user controlled, via a special parameter attached to each commodity. However, the constraint defaults to an equality in the case of materials (i.e. the quantity produced and imported is *exactly* equal to that consumed and exported), and to an inequality in the case of energy carriers, emissions and demands (thus allowing some surplus production). For those commodities for which time-slices have been defined, the balance constraint must be satisfied in each time-slice.

The balance constraint is very complex, due to the many terms involving production or consumption of a commodity. We present a much simplified version below, to simply indicate the basic meaning of this equation.

For each commodity ***c,*** time period ***t (***vintage ***v)***, region ***r***, and time-slice ***s*** (if necessary or "ANNUAL" if not), this constraint requires that the disposition of each commodity balances its procurement. The disposition includes consumption in the region plus exports; the procurement includes production in the region plus imports.

EQ_COMBAL(r,t,c,s) - Commodity Balance

[Sum *{over all* ***p,c*** *∈ TOP(r,p,c,"out") of:* **[*FLOW(r,v,t,p,c,s)* + *SOUT(r,v,t,p,c,s)**STG_EFF(r,v,p)*]** *}* **+**

Sum *{over all* ***p,c*** *∈ RPC_IRE(r,p,c,"imp") of:* ***TRADE(r,t,p,c,s,"imp")****}* +

Sum *{over all* ***p*** *of: Release(r,t,p,c)**NCAP(r,t,p,c)****}]*** *** COM_IE(r,t,c,s)*

≥ or = **(4-4)**

Sum *{over all* ***p,c*** *∈* ***TOP(r,p,c,"in")*** *of:* ***FLOW(r,v,t,p,c,s)* + *SIN(r,v,t,p,c,s)****}* **+**

Sum {*over all* ***p,c*** *∈* ***RPC_IRE(r,p,c,"exp")****} of:* ***TRADE(r,t,p,c,s,'exp")*** **+**

Sum *{over all* ***p*** *of: Sink(r,t,p,c)**NCAP(r,t,p,c)****}* +

FR(c,s) * DM(c,t)

where:

The constraint is $\geq$ for energy forms and = for materials and emissions (unless these defaults are overridden by the user).

TOP(r,p,c,"in/out") identifies that there is an input/output flow of commodity c into/from process *p* in region *r;*

RPC_IRE(r,p,c,"imp/exp"") identifies that there is an import/export flow into/from region *r* of commodity *c* via process *p;*

STG_EFF(r,v,p) is the efficiency of storage process p;

COM_IE(r,t,c) is the infrastructure efficiency of commodity c;

Release(r,t,p,c) is the amount of commodity ***c*** recuperated per unit of capacity of process ***p*** dismantled (useful to represent some materials or fuels that are recuperated while dismantling a facility);

Sink(r,t,p,c) is the quantity of commodity ***c*** required per unit of new capacity of process ***p*** (useful to represent some materials or fuels consumed for the construction of a facility);

FR(s) is the fraction of the year covered by time-slice ***s*** (equal to 1 for non- time-sliced commodities).

***Example**: Gasoline consumed by vehicles plus gasoline exported to other regions must not exceed gasoline produced from refineries plus gasoline imported from other regions.*

4.4.5 Defining flow relationships in a process

A process with one or more (perhaps heterogeneous) commodity flows is essentially defined by one or more independent input and output flow variables. In the absence of relationships between these flows, the process would be completely undetermined, i.e. its outputs would be independent form its inputs.We therefore need one or more constraints stating that the ratio of the sum of some of its output flows to the sum of some of its input flows is equal to a constant (which is akin to an efficiency). In the case of a single commodity in, and a single commodity out of a process, this equation defines the traditional efficiency of the process. With several commodities, this constraint may leave some freedom to individual output (or input) flows, as long as their sum is in fixed proportion to the sum of input (or output) flows. An important rule for this constraint is that each sum must be taken over commodities of the same type (i.e. in the same group, say: energy carriers, or emissions, etc.). In TIMES, for each process the modeler identifies the input commodity group *cg*1*,* and the output commodity group *cg*2, and chooses a value for the efficiency ratio, named *FLOFUNC(p,cg1,cg2).* The following equation embodies this:

EQ_PTRANS(r,v,t,p,cg1,cg2,s) *–Efficiency definition*

$$\textbf{\textit{SUM}}\left\{ c \textit{ in } cg2 \textit{ of : } \textbf{\textit{FLOW(r,v,t,p,c,s)}} \right\} =$$

$$FLOFUNC(r,v,cg1,cg2,s) * \textbf{\textit{SUM}} \left\{ c \textit{ within } cg1 \textit{ of: } COEFF(r,v,p,cg1,c,cg2,s) * \textbf{\textit{FLOW(r,v,t,p,c,s)}} \right\} \qquad \textbf{(4-5)}$$

Where *COEFF(r,v,p,cg1,c,cg2,s)* takes into account the harmonization of different time-slice resolution of the flow variables, which have been omitted here for simplicity, as well as commodity-dependent transformation efficiencies.

4.4.6 Limiting flow shares in flexible processes

When either of the commodity groups *cg*1 or *cg*2 contains more than one element, the previous constraint allows a lot of freedom on the values of flows. The process is therefore quite flexible. The flow share constraint is intended to limit the flexibility, by constraining the share of each flow within its own group. For instance, a refinery output might consist of three refined products: *c1*=light, *c2*=medium, and *c3*=heavy distillate. If losses are 9% of the input, then the user must specify FLOFUNC = 0.91 to define the overall efficiency. The user may then want to limit the flexibility of the slate of outputs by means of three *FLOSHAR(ci) coefficients, say 0.4, 0.5, 0.6,* resulting in three flow share constraints as follows:

FLOW(c1) ≤ 0.4[FLOW(c1) + FLOW(c2) + FLOW(c3)],* so that *c*1 is at most 40% of the total output,
FLOW(c2) ≤ 0.5[FLOW(c1) + FLOW(c2) + FLOW(c3)],* so that *c*2 is at most 50% of the total output,
FLOW(c3) ≤ 0.6[FLOW(c1) + FLOW(c2) + FLOW(c3)],* so that *c*3 is at most 60% of the total output,

The general form of this constraint is:

EQ_INSHR(c,cg,p,r,t,s) and *EQ_OUTSHR(c,cg,p,r,t,s)*

$$\textbf{\textit{FLOW(c)}} \leq \geq = FLOSHAR(c) * Sum \{over\ all\ c' \ in\ cg\ of: \textbf{\textit{FLOW(c')}} \} \qquad \textbf{(4-6)}$$

The commodity group *cg* may be on the input or output side of the process.

4.4.7 Peaking Reserve Constraint (time-sliced commodities only)

This constraint imposes that the total capacity of all processes producing a commodity at each time period and in each region must exceed the average demand in the time-slice where peaking occurs by a certain percentage. This percentage is the Peak Reserve

Factor, *RESERV(r,t,c)*, and is chosen to insure against several contingencies, such as: possible commodity shortfall due to uncertainty regarding its supply (e.g. water availability in a reservoir); unplanned equipment down time; and random peak demand that exceeds the average demand during the time-slice when the peak occurs. This constraint is therefore akin to a safety margin to protect against random events not explicitly represented in the model. In a typical cold country the peaking time-slice for electricity (or natural gas) will be Winter-Day, and the total electric plant generating capacity (or gas supply plant) must exceed the Winter-Day demand load by a certain percentage. In a warm country the peaking time-slice may be Summer-Day for electricity (due to heavy air conditioning demand). The user controls for which time-slices a peaking equation is maintained.

For each time period ***t*** and for region ***r***, there must be enough installed capacity to exceed the required capacity in the season with largest demand for commodity ***c*** by a safety factor *E* called the *peak reserve factor*.

EQ_PEAK(r,t,c,s) *- commodity peak requirements*

Sum *{over all* ***p producing c with c=pcg*** *of CAPUNIT(r,p) * Peak(r,v,p,c,s)* ******* *FR(s)* ****CAP(r,v,t,p) * ACTFLO(r,v,p,c)*** *}* **+**

Sum *{over all* ***p*** *producing* ***c*** *with* ***c#pcg*** *of Peak(r,v,p,c,s) *FLOW(r,v,t,p,c,s) +* ***TRADE(r,t,p,c,s,i)***

≥ **(4-7)**

*[1+RESERVE(r,t,c,s)] * [* ***Sum*** *{over all* ***p consuming c*** *of* ***FLOW(r,v,t,p,c,s) + TRADE(r,t,p,c,s,e)*** *}]*

where:

RESERV(r,t,c,s) is the region-specific reserve coefficient for commodity *c* in time-slice *s*, which allows for unexpected down time of equipment, for demand at peak, and for uncertain resource availability, and

Peak(r,v,p,c,s) (never larger than 1) specifies the fraction of technology *p*'s capacity in a region *r* for a period *t* and commodity *c* (electricity or heat only) that is allowed to contribute to the peak load in slice *s*; many types of supply processes are predictably available during the peak and thus have a peak coefficient equal to 1, whereas others (such as wind turbines or solar plants in the case of electricity) are attributed a peak coefficient less than 1, since they are on average only fractionally available at peak (e.g., a wind turbine typically has a peak coefficient of .25 or .3, whereas a hydroelectric plant, a gas plant, or a nuclear plant typically has a peak coefficient equal to 1).

For simplicity it has been assumed in (4-7) that the time-slice resolution of the peaking commodity and the time-slice resolution of the commodity flows (FLOW, TRADE) are the same. In practice, this is not the case and additional conversion factors or summation operations are necessary to match different time-slice levels.

Remark: to establish the peak capacity, two cases must be distinguished in equation EQ_PEAK.
– For production processes where the peaking commodity is the only commodity in the primary commodity group (denoted c=pcg), the capacity of the process may be assumed to contribute to the peak.
– For processes where the peaking commodity is not the only member of the pcg, there are several commodities included in the pcg. Therefore, the capacity as such cannot be used in the equation. In this case, the actual production is taken into account in the contribution to the peak, instead of the capacity. For example, in the case of CHP only the production of electricity contributes to the peak electricity supply, not the entire capacity of the plant, because the activity of the process consists of both electricity and heat generation in either fixed or flexible proportions, and, depending on the modeler's choice, the capacity may represent either the electric power of the turbine in condensing or back-pressure mode, or the sum of power and heat capacities in back-pressure mode.

Note also that in the peak equation (4-7), it is assumed that imports of the commodity are contributing to the peak of the importing region (thus, exports are of the *firm power* type).

4.4.8 Constraints on commodities

In TIMES variables are optionally attached to various quantities related to commodities, such as total quantity produced. Therefore it is quite easy to put constraints on these quantities, by simply bounding the commodity variables at each period. It is also possible to impose cumulative bounds on commodities over more than one period, a particularly useful feature for cumulatively bounding emissions or modeling reserves of fossil fuels. By introducing suitably naming conventions for emissions the user may constrain emissions from specific sectors. Furthermore, the user may also impose global emission constraints that apply to several regions taken together, by allowing emissions to be traded across regions. Alternatively or concurrently a tax or penalty may be applied to each produced (or consumed) unit of a commodity (energy form, emission), via specific parameters.

A specific type of constraint my be defined to limit the share of process (p) in the total production of commodity (c). The constraint indicates that the flow of commodity (c) from/to process (p) is bounded by a given fraction of the total production of commodity (c). In the present implementation, the same given fraction is applied to all time slices.

4.4.9 User Constraints

In addition to the standard TIMES constraints discussed above, the user interested in developing reference case projections of energy market behavior typically introduces additional constraints to express these special conditions. For example, there may a user-defined constraint limiting investment in new nuclear capacity (regardless of the type of reactor), or dictating that a certain percentage of new electricity generation capacity must be powered by renewable energy sources. User constraints may be employed across time periods, to for example model options for retrofitting and life extension for processes.

In order to facilitate the creation of a new user constraint TIMES provides a *template* for indicating a) the set of variables involved in the constraint, and b) the user-defined coefficients needed in the constraint.

4.4.10 Representation of oil refining in MARKAL

Two alternative approaches are available in MARKAL to represent oil refining. Under the simplified approach that is adopted in the majority of MARKAL models, the refinery is treated as a set of one or more standard MARKAL technologies. But a more sophisticated approach is available where the modeler wishes to specify *bona fide* quality requirement constraints for each refined product, such as: Octane Rating, Sulfur Content, Flash Index, Density, Cetane Number, Viscosity, Reid Vapor Pressure, etc. This approach is embodied in the special oil-refining module of MARKAL – whose main features are outlined in this section – and requires additional parameters, variables, and constraints.

4.4.10.1 New sets and parameters

First, in order to properly apply the specific constraints to this sector, the model requires that several sets and other parameters that are unique to this sector be defined, as follows:

- Constant **REFUNIT**: specifies in what units the refinery streams are defined (volume, weight, or energy)
- Set **REF**: contains the list of refining processes (a subset of **PRC**)
- Set **OPR**: contains the intermediate energy carriers (refinery streams) that are produced by the members of **REF**, plus the available crude oils. Each stream will enter the production of one or more RPP. The members of **OPR** are expressed in units specified by REFUNIT (volume, weight, or energy).
- Set BLE: contains the product commodity of the blending activity
- Parameter **CONVERT**: contains the density and energy content (by weight or by volume) of each blending stream. These parameters are used as coefficients of the blending equations to convert them to correct units (see below).
- Set **SPE:** contains the names of the specifications that must be imposed on refined petroleum products (RPPs), such as octane rating, sulfur content, etc.

- Parameter BL_COM: contains the values of the blending specifications SPE for the blending streams OPR
- Parameter BL_SPEC: contains the value of the specification SPE of the blending product BLE

4.4.10.2 New variables

Once these constants, sets and tables are defined by the user, the model automatically creates blending variables as follows:

$\boldsymbol{BLND_{t,,ble,opr}}$: is the amount of blending stock OPR entering the production of the refined product ***ble*** at time period ***t***.

4.4.10.3 New blending constraints

These variables and input parameters are finally used to express the following two types of blending constraints:

Blending by volume (e.g. octane rating)

$$\sum_{opr \in OPR} (oct_{opr} \cdot BLND_{t,ble,opr}) = oct_{ble} \cdot \sum_{opr \in OPR} BLND_{t,ble,opr}$$

where

- $\boldsymbol{oct_{opr}}$ is the octane content by volume of one unit of stream ***opr*** (itself expressed in units **REFUNIT**. If **REFUNIT** is not equal to *'volume'*, some conversion coefficients (specified in table **CONVERT**) must be applied to the variables of the equation.
- $\boldsymbol{Oct_{ble}}$ is the required volume octane rating of the refined product ***enc***
- ***BLND*** variables are expressed in volume units.

Blending by weight (e.g. sulfur content)

$$\sum_{opr \in OPR} (sulf_{opr} \cdot BLND_{t,ble,opr}) = sulf_{ble} \cdot \sum_{opr \in OPR} BLND_{t,ble,opr}$$

where

- $\boldsymbol{sulf_{opr}}$ is the sulfur content by weight of one unit of stream ***opr*** (itself expressed in units **REFUNIT**. If **REFUNIT** is not equal to *'weight'*, then conversion coefficients (specified in table **CONVERT**) must be applied to the variables of the equation.
- $\boldsymbol{sulf_{blec}}$ is the weight sulfur content of the refined product ***enc***
- ***BLND*** variables are expressed in weight units.

4.5 Linear Programming complements

This section is not strictly needed for a basic understanding of the TIMES model and may be skipped on a first reading. However, it provides additional insights into the microeconomics of the TIMES equilibrium. In particular, it contains a review of the theoretical foundation of Linear Programming and Duality Theory. This knowledge may help the user to better understand the central role shadow prices and reduced costs play in the economics of the TIMES model. More complete treatments of Linear Programming and Duality Theory may be found in several standard textbooks such as Chvátal (1983) or Hillier and Lieberman (1990 and subsequent editions). Samuelson and Nordhaus (1977) contains a treatment of micro-economics based on mathematical programming.

4.5.1 A brief primer on Linear Programming and Duality Theory

4.5.1.1 Basic definitions

In this subsection, the superscript *t* following a vector or matrix represents the transpose of that vector or matrix. A Linear Program may always be represented as the following *Primal Problem* in canonical form:

$$Max\ c^t x \qquad \textbf{(4-7)}$$
$$s.t. \quad Ax \le b \qquad \textbf{(4-8)}$$
$$x \ge 0 \qquad \textbf{(4-9)}$$

where x is a vector of *decision variables*, $c^t x$ is a linear function representing the *objective* to maximize, and $Ax \le b$ is a set of inequality *constraints*. Assume that the LP has a finite optimal solution, x^*.

Then each decision variable, x^*_j falls into one of three categories. x^*_j may be:

- equal to its lower bound (as defined in a constraint), or
- equal to its upper bound, or
- strictly between the two bounds.

In the last case, the variable x^*_j is called *basic*. Otherwise it is *non-basic*.

For each primal problem, there corresponds a *Dual problem* derived as follows:

$$Min\ \ b^t y \qquad \textbf{(4-10)}$$
$$s.t. \quad A^t y \ge c \qquad \textbf{(4-11)}$$
$$y \ge 0 \qquad \textbf{(4-12)}$$

Note that the number of dual variables equals the number of constraints in the primal problem. In fact, each dual variable y_i may be assigned to its corresponding primal constraint, which we represent as: $A_i x \leq b_i$, where A_i is the i^{th} row of matrix A.

4.5.1.2 Duality Theory

Duality theory consists mainly of three theorems[26]: weak duality, strong duality, and complimentary slackness.

Weak Duality Theorem
If x is any feasible solution to the primal problem and y is any feasible solution to the dual, then the following inequality holds:

$$c^t x \leq b^t y \qquad \textbf{(4-13)}$$

The weak duality theorem states that the value of a feasible dual objective is never smaller than the value of a feasible primal objective. The difference between the two is called the *duality gap* for the pair of feasible primal and dual solutions (x,y).

Strong duality theorem
If the primal problem has a *finite, optimal* solution x^*, then so does the dual problem (y^*), and both problems have the same optimal objective value (their duality gap is zero):

$$c^t x^* = b^t y^* \qquad \textbf{(4-14)}$$

Note that the optimal values of the dual variables are also called the *shadow prices* of the primal constraints.

Complementary Slackness theorem

At an optimal solution to an LP problem:

- If y^*_i is > 0 then the corresponding primal constraint is satisfied at equality (i.e. $A_i x^* = b_i$ and the i^{th} primal constraint is called *tight*. Conversely, if the i^{th} primal constraint is *slack* (not tight), then $y^*_i = 0,$
- If x^*_j is basic, then the corresponding dual constraint is satisfied at equality, (i.e. $A^t_j {}^* y = c_j$, where A^t_j is the j^{th} row of A^t, i.e. the j^{th} column of A. Conversely, if the j^{th} dual constraint is slack, then x^*_j is equal to one of its bounds.

Remark: Note however that a primal constraint may have zero slack and yet have a dual equal to 0. And, a primal variable may be non basic (i.e. be equal to one of its bounds), and yet the corresponding dual slack be still equal to 0. These situations are differentcases of the so-called degeneracy of the LP. They ofetn occur when constraints are over specified (a trivial case occurs if a constraint is repeated twice in the LP)

[26] Their proofs may be found in most textbooks on Linear Programming, such as Chvatal (1983) or Hillier and Lieberman (1990).

4.5.2 Sensitivity analysis and the economic interpretation of dual variables

It may be shown that if the j^{th} RHS b_j of the primal is changed by an infinitesimal amount d, and if the primal LP is solved again, then its new optimal objective value is equal to the old optimal value plus the quantity $y_j^* \bullet d$, where y_j^* is the optimal dual variable value.

Loosely speaking[27], one may say that *the partial derivative of the optimal primal objective function's value with respect to the RHS of the i^{th} primal constraint is equal to the optimal shadow price of that constraint.*

4.5.2.1 Economic Interpretation of the Dual Variables

If the primal problem consists of maximizing the surplus (objective function $c^t x$), by choosing an activity vector x, subject to upper limits on several resources (the b vector) then:

- Each a_{ij} coefficient of the dual problem matrix, A, then represents the consumption of resource b_j by activity x_i;
- The optimal dual variable value y^*_j is the unit price of resource j, and
- The total optimal surplus derived from the optimal activity vector, x^*, is equal to the total value of all resources, b, priced at the optimal dual values y^* (strong duality theorem).

Furthermore, each dual constraint $A^t_j * y \geq c_j$ has an important economic interpretation. Based on the Complementary Slackness theorem, if an LP solution x^* is optimal, then for each x^*_j that is not equal to its upper or lower bound (i.e. each basic variable x^*_j), there corresponds a *tight* dual constraint $y^*A'_j = c_j$, which means that the revenue coefficient c_j must be exactly equal to the cost of purchasing the resources needed to produce one unit of x_j. In economists' terms, *marginal cost equals marginal revenue, and both are equal to the market price of* $\boldsymbol{x^*_j}$. If a variable is not basic, then by definition it is equal to its lower bound or to its upper bound. In both cases, the unit revenue c_j need not be equal to the cost of the required resources. The technology is then either non-competitive (if it is at its lower bound) or it is super competitive and makes a surplus (if it is at its upper bound).

Example: The optimal dual value attached to the balance constraint of commodity c represents the change in objective function value resulting from one additional unit of the commodity. This is precisely the internal price of that commodity.

4.5.2.2 Reduced Surplus and Reduced Cost

In a maximization problem, the difference $y^*A'_j - c_j$ is called the *reduced surplus* of technology j, and is available from the solution of a TIMES problem. It is a useful indicator of the competitiveness of a technology, as follows:

---

[27] Strictly speaking, the partial derivative may not exist for some values of the RHS, and may then be replaced by a directional derivative.

- if x^*_j is at its lower bound, its unit revenue c_j is *less* than the resource cost (i.e. its reduced surplus is positive). The technology is not competitive (and stays at its lower bound in the equilibrium);
- if x^*_j is at its upper bound, revenue c_j is *larger* than the cost of resources (i.e. its reduced surplus is negative). The technology is super competitive and produces a surplus; and
- if x^*_j is basic, its reduced surplus is equal to 0. The technology is competitive but does not produce a surplus.

We now restate the above summary in the case of a Linear Program that minimizes cost subject to constraints:

$$\begin{aligned} & Min\ c^t x \\ s.t.\quad & Ax \geq b \\ & x \geq 0 \end{aligned}$$

In a minimization problem (such as the usual formulation of TIMES), the difference $c_j - y^*A'_j$ is called the *reduced cost* of technology *j*. The following holds:

- if x^*_j is at its lower bound, its unit cost c_j is *larger* than the value created (i.e. its reduced cost is positive). The technology is not competitive (and stays at its lower bound in the equilibrium);
- if x^*_j is at its upper bound, its cost c_j is *less* than the value created (i.e. its reduced cost is negative). The technology is super competitive and produces a profit; and
- if x^*_j is basic, its reduced cost is equal to 0. The technology is competitive but does not produce a profit

The reduced costs/surpluses may thus be used to rank all technologies, *including those that are not selected by the model.*

5 A comparison of the TIMES and MARKAL models

This chapter contains a point-by-point comparison of the TIMES and MARKAL models. It is of interest primarily to modelers already familiar with MARKAL, and to modelers who are considering adoption of either model. The descriptions of the features given below are not detailed, since they are repeated elsewhere in the documentation. Rather, the function of this chapter is to guide the reader, by mentioning the features that are present in one model and not in the other.

5.1 Similarities

The TIMES and the MARKAL models share the same basic modeling paradigm. Both models are technology explicit, dynamic partial equilibrium models of energy markets. In both cases the equilibrium is obtained by maximizing the total surplus of consumers and suppliers via Linear Programming. The two models also share the multi-regional feature, which allows the modeler to construct geographically integrated (even global) instances. These fundamental features were described in chapter 3 of this documentation, and Section 1.3, PART I of the MARKAL documentation, and constitute the backbone of the common paradigm. However, there are also significant differences in the two models, which we now outline. These differences do not affect the basic paradigm common to the two models, but rather some of their technical features and properties.

5.2 TIMES features not in MARKAL

5.2.1 Variable length time periods

MARKAL has fixed length time periods. However TIMES allows the user to define period lengths in a completely flexible way. This is a major model difference, which indeed required a complete re-definition of the mathematics of most TIMES constraints and of the TIMES objective function. The variable period length feature is very useful in two instances: first if the user wishes to use a single year as initial period (handy for calibration purposes), and second when the user contemplates long horizons, where the first few periods may be described in some detail by relatively short periods (say 5 years), while the longer term may be regrouped into a few periods with long durations (perhaps 20 or more years).

5.2.2 Data decoupling

This somewhat misunderstood feature does not confer additional power to TIMES, but it greatly simplifies the maintenance of the model database and allows the user great flexibility in modifying the new definition of the planning horizon. In TIMES all input data are specified by the user independently from the definition of the time periods employed for a particular model run. All time-dependent input data are specified by the year in which the data applies. The model then takes care of matching the data with the periods, wherever required. If necessary the data is interpolated (or extrapolated) by the model preprocessor code to provide data points at those time periods required for the current model run. In addition, the user has control over the interpolation and extrapolation of each time series.

The general rule of data decoupling applies also to past data: whereas in MARKAL the user had to provide the residual capacity profiles for all existing technologies in the initial period, and over the periods in which the capacity remains available, in TIMES the user provides technical and cost data at those past years when the investments actually took place, and the model takes care of calculating how much capacity remains in the various modeling periods. Thus, past and future data are treated essentially in the same manner in TIMES. One instance when the data decoupling feature immensely simplifies model management is when the user wishes to change the initial period, and/or the lengths of the periods. In TIMES, there is essentially nothing to do, except declaring the dates of the new periods. In MARKAL, such a change represents a much larger effort requiring a substantive revision of the database.

5.2.3 Flexible time slices and storage processes

In MARKAL, only two commodities have time-slices: electricity and low temperature heat, and their time slices are rigidly defined (six time-slices for electricity and three for heat). In TIMES, any commodity and process may have its own, user-chosen time-slices. These flexible time-slices are segregated into three groups, seasonal (or monthly), weekly (weekday vs weekend), and daily (day/night), where any level may be expanded (contracted) or omitted.

The flexible nature of the TIMES time-slices is supported by storage processes that ’consume’ commodities at one time-slice and release them at another. MARKAL only supports night-to-day (electricity) storage.

Note that many TIMES parameters may be time-slice dependent (such as availability factor (AF), basic efficiency (FLO_FUNC), etc).

5.2.4 Process generality

In MARKAL processes in different RES sectors are endowed with different (data and mathematical) properties. For instance, end-use processes do not have activity variables (activity is then equated to capacity), and source processes have no investment variables. In TIMES, every process has the same basic features, which are activated or not solely via data specification.

5.2.5 Flexible processes

In MARKAL processes are by definition rigid, except for some specialized processes which permit flexible output (such as limit refineries or pass-out turbine CHPs), and thus outputs and inputs are in fixed proportions with one another. In TIMES, the situation is reversed, and each process starts by being entirely flexible, unless the user specifies certain coefficients to rigidly link inputs to outputs. This feature permits better modeling of many real-life processes as a single technology, where MARKAL requires several technologies (as well as dummy commodities) to achieve the same result. A typical example is that of a boiler that accepts any of 3 liquid fuels as input, but whose efficiency depends on the fuel used. In MARKAL, to model this situation requires four processes (one per possible fuel plus one that carries the investment cost and other parameters), plus one dummy fuel. In TIMES one process is sufficient, and no dummy fuel is required. Note also that TIMES has a number of parameters that limit the input share of each fuel, whereas in MARKAL, imposing such limits requires that the user define several user constraints.[28]

5.2.6 Investment and dismantling lead-times and costs

New TIMES parameters allow the user to model the construction phase and dismantling of facilities that have reached their end-of-life. These are: lead times attached to the construction or to the dismantling of facilities, capital cost for dismantling, surveillance costs during the dismantling lead-time. Like in MARKAL, there is also the possibility to define flows of commodities consumed at construction time, or released at dismantling times, thus allowing the representation of life-cycle energy and emission accounting.

5.2.7 Vintaged processes and age-dependent parameters

The variables associated with user declared vintaged processes employ both the time period p and vintage period v (in which new investments are made and associated input data is obtained). The user indicates that a process is to be modeled as a vintaged process by using a special vintage parameter. Note that in MARKAL vintaging is possible only for demand devices (for which there is no activity variable) or via the definition of several replicas of a process, each replica being a different vintage. In TIMES, the same process name is used for all vintages of the same process.[29]

In addition, some parameters can be specified to have different values according to the age of the process. In the current version of TIMES, these parameters include the availability factors, the in/out flow ratios (equivalent to efficiencies), and the fixed cost

[28] In the end the two models use equivalent mathematical expressions to represent a flexible process. Only TIMES reduces the user's effort to a minimum, while MARKAL requires the user to manually define the multiple processes, dummy fuels and user constraints.
[29] The representation of vintage as a separate index helps eliminate a common confusion that existed in MARKAL, namely the confusion of vintage with the age of a process. For instance, if the user defines an annual cost for a car equal to 10 in 2005 and only 8 in 2010, the decrease would not only apply to cars purchased in 2010, but also to cars purchased in 2005 and earlier when they reach the 2010 period.

parameters only. Several other parameters could, in principle, be defined to be age-dependent, but such extensions have not been implemented yet.

5.2.8 Commodity related variables

MARKAL has very few commodity related variables, namely exports/imports, and emissions. TIMES has a large number of commodity-related variables such as: total production, total consumption, but also specific variables representing the flows of commodities entering or exiting each process. This allows the user many "handles" to put limits, and costs on commodities.

5.2.9 More accurate and realistic depiction of investment cost payments

In MARKAL each investment is assumed to be paid in its entirety at the beginning of some time period. In TIMES the timing of investment payments is quite detailed. For large facilities (e.g. a nuclear plant), capital is progressively laid out in yearly increments over the facility's construction time, and furthermore, the payment of each increment is made in installments spread over the economic life of the facility. For small processes (e.g. a car) the capacity expansion is assumed to occur regularly each year rather than in one large lump, and the payments are therefore also spread over time. Furthermore, when a time period is quite long (i.e. longer that the life of the investment), TIMES has an automatic mechanism to repeat the investment more than once over the period. These features allow for a much smoother (and more realistic) representation of the stream of capital outlays in TIMES than in MARKAL.

Moreover, in TIMES all discount rates can be defined to be time-dependent, whereas in MARKAL both the general and technology-specific discount rates are constant over time.

5.2.10 Climate equations

TIMES now possesses a set of variables and equations that endogenize the concentration of CO2 and also calculate the radiative forcing and global temperature change resulting from GHG emissions and accumulation in the atmosphere. This new feature is described in chapter 7 of PART II.

5.3 *MARKAL features not in TIMES*

Over the years, several extensions were added to the MARKAL model. Most of these extensions have been implemented in TIMES, for instance elastic demands, multi-regional trading, lumpy investments, and endogenous technological learning. Four extensions of MARKAL have not (yet) been implemented in TIMES, namely:

- Stochastic programming,
- Integration of damage costs in the objective function,
- Myopic energy markets (SAGE), and
- MARKAL-MACRO.

There are plans to implement most of these missing options in the future.

The interested reader is referred to the MARKAL documentation[30] for the descriptions of these four options, and the full description of the model.

[30] http://www.etsap.org/documentation.asp, with links to Standard MARKAL, MARKAL-MACRO and SAGE write-ups.

6 Elastic demands and the computation of the supply-demand equilibrium

In the preceding chapters, we have seen that TIMES does more than minimize the cost of supplying energy services. Instead, it computes a supply-demand equilibrium where both the supply options and the energy service demands are computed by the model. The equilibrium is driven by the user-defined specification of demand functions, which determine how each energy service demand varies as a function of the market price of that energy service. The TIMES code assumes that each demand has constant own-price elasticity in a given time period, and that cross price elasticities are zero. Economic theory establishes that the equilibrium thus computed corresponds to the maximization of the net total surplus, defined as the sum of the suppliers and of the consumers' surpluses (Samuelson, 1952, Takayama and Judge, 1972). The total net surplus has been often considered a valid metric of societal welfare in microeconomic literature, and this fact confers strong validity to the equilibrium computed by TIMES.

The TIMES model is normally run in two contrasted modes: first to simulate some reference case, and then to simulate alternate scenarios, each of which departs in some way from the reference case assumptions and parameters. For instance, an alternate scenario may make different assumptions on the availability or the cost of some new technologies. Or, it may assume that certain energy or environmental policies are being implemented (e.g. emission taxes, or portfolio standards, or efficiency improvements). Or again, a scenario may assume that a certain goal must be reached (such as a cap on emissions), leaving the model free to achieve that goal at least cost. In almost any such alternate scenario, some strain is put on some sectors, resulting in increases in the marginal values of at least some energy services (for example, severe emission reductions may increase the price of auto transportation). In TIMES demands self-adjust in reaction to changes (relative to the reference case) of their own price, and therefore the model goes beyond the optimization of the energy sector only. Thus although TIMES falls short of computing a general equilibrium, it does capture a major element[31] of the feedback effects not previously accounted for in bottom-up energy models.

In this chapter, we explain how Linear Programming computes the equilibrium. Additional technical details may be found in Tosato (1980) and in Loulou and Lavigne (1995). One of the first large scale application of these methods was realized in the Project Independence Energy System (PIES, Hogan, 1975), although in the context of demands for final energy rather than for energy services as in TIMES or MARKAL.

6.1 Theoretical considerations: the Equivalence Theorem

The computational method is based on the equivalence theorem presented in chapter 3, which we restate here:

---

[31] It has been argued, based on strong circumstantial evidence, that the change in demands for energy services is indeed the main feedback economic effect of energy system policies (Loulou and Kanudia, 2002)

"A supply/demand economic equilibrium is reached when the sum of the producers and the consumers surpluses is maximized"

Figure 3.2 of Chapter 3 provides a graphical illustration of this theorem in a case where only one commodity is considered.

6.2 Mathematics of the TIMES equilibrium

6.2.1 Defining demand functions

For each demand category, define a demand curve, i.e. a function determining demand as a function of price. In TIMES, a constant elasticity relationship is used, represented as:

$$\boldsymbol{DM_i(p\ ) = K_i \bullet p_i^{E_i}} \qquad (6-1)$$

where DM_i is the i^{th} demand, p_i is its price, taken to be the marginal cost of procuring the i^{th} commodity, and E_i is the own price elasticity of that demand. Note that although the region and time indexes *r, t* have been omitted in this notation, all quantities in Equation (1.6-1), including the elasticities are region (if appropriate) and time dependent. Constant K_i may be obtained if one point (p^0_i, DM^0_i) of the curve is known (the reference case). Thus Equation (1.6-1) may be rewritten as:

$$\boldsymbol{DM_i / DM_i^0 = (p_i / p_i^0)^{E_i}} \qquad (6-2)$$

Or its inverse:

$$\boldsymbol{p_i = p_i^0 \cdot (DM_i / DM_i^0)^{1/E_i}}$$

where the superscript ‘0’ indicates the reference case, and the elasticity E_i is negative. Note also that the elasticity may have two different values, one for upward changes in demand, the other for downward changes.

6.2.2 Formulating the TIMES equilibrium

With inelastic demands, the TIMES model may be written as the following Linear Program

$$\boldsymbol{Min}\quad \boldsymbol{c \cdot X} \qquad (6-3)$$

$$\boldsymbol{s.t.}\quad \sum_{\boldsymbol{k}} \boldsymbol{CAP_{k,i}(t) \geq DM_i(t)} \qquad \boldsymbol{i} = 1,2,..,\boldsymbol{I};\ \boldsymbol{t} = 1,..,\boldsymbol{T} \qquad (6-4)$$

$$\boldsymbol{and}\quad \boldsymbol{B \cdot X \geq b} \qquad (6-5)$$

where $\boldsymbol{X}$ is the vector of all variables and $\boldsymbol{I}$ is the number of demand categories. In words:

- (6-3) expresses the total discounted cost to be minimized.
- (6-4) is the set of demand satisfaction constraints (where the ***CAP*** variables are the capacities of end-use technologies, and the ***DM*** right-hand-sides are the exogenous demands to satisfy).
- (6-5) is the set of all other constraints.

With elastic demands the role of TIMES is to compute a supply/demand equilibrium among equations (1.6-3) through (1.6-5) where both the supply side and the demand side adjust to changes in prices, and the prices charged by the supply side are the marginal costs of the demand categories (i.e. p_i is the marginal cost of producing demand DM_i). A priori this seems to be a difficult task, because the prices used on the demand side are computed as part of the solution to equations (1.6-3), (1.6-4), and (1.6-5). The Equivalence Theorem, however, states that such an equilibrium is reached as the solution of the following mathematical program, where the objective is to maximize the net total surplus:

$$\boldsymbol{Max} \quad \sum_{i}\sum_{t}\left(\boldsymbol{p}_{i}^{0}(\boldsymbol{t})\cdot\left[\boldsymbol{DM}_{i}^{0}(\boldsymbol{t})\right]^{-1/\boldsymbol{E}_i}\bullet \int_{\boldsymbol{a}}^{\boldsymbol{DM}_i(\boldsymbol{t})} \boldsymbol{q}^{1/\boldsymbol{E}_i}\cdot \boldsymbol{dq}\right) - \boldsymbol{c}\cdot \boldsymbol{X} \qquad (6-6)$$

$$\boldsymbol{s.t.} \qquad \sum_{k} \boldsymbol{CAP}_{k,i}(\boldsymbol{t}) - \boldsymbol{DM}_i(\boldsymbol{t}) \ge 0 \qquad \boldsymbol{i}=1,..,\boldsymbol{I};\ \boldsymbol{t}=1,..,\boldsymbol{T} \qquad (6-7)$$

$$\boldsymbol{and} \qquad \boldsymbol{B}\cdot\boldsymbol{X} \ge \boldsymbol{b} \qquad (6-8)$$

Where ***X*** is the vector of all TIMES variables with associated cost vector ***c***, (6-6) expresses the total net surplus, and ***DM*** is now a vector of *variables* in (6 -7), rather than fixed demands.

The integral in (6-6) is easily computed, yielding the following maximization program:

$$\boldsymbol{Max} \quad \sum_{i}\sum_{t}\left(\boldsymbol{p}_{i}^{0}(\boldsymbol{t})\cdot\left[\boldsymbol{DM}_{i}^{0}(\boldsymbol{t})\right]^{-1/\boldsymbol{E}_i}\bullet \boldsymbol{DM}_i(\boldsymbol{t})^{1+1/\boldsymbol{E}_i}/(1+1/\boldsymbol{E}_i)\right) - \boldsymbol{c}\cdot \boldsymbol{X} \qquad (6-6)'$$

$$\boldsymbol{s.t.} \qquad \sum_{k} \boldsymbol{CAP}_{k,i}(\boldsymbol{t}) \ge \boldsymbol{DM}_i(\boldsymbol{t}) \qquad \boldsymbol{i}=1,..,\boldsymbol{I};\ \boldsymbol{t}=1,..,\boldsymbol{T} \qquad (6-7)'$$

$$\boldsymbol{B}\cdot\boldsymbol{X} \ge \boldsymbol{b} \qquad (6-8)'$$

6.2.3 Linearization of the Mathematical Program

The Mathematical Program embodied in (6-6)', (6-7)' and (6-8)' has a non-linear objective function. Because the latter is separable (i.e. does not include cross terms) and concave in the DM_i variables, each of its terms is easily linearized by piece-wise linear functions which approximate the integrals in (6-6). This is the same as saying that the inverse demand curves are approximated by staircase functions, as illustrated in figure 6.1. By so doing, the resulting optimization problem becomes linear again. The linearization proceeds as follows.

a) For each demand category $\boldsymbol{i}$, the user selects a range within which it is estimated that the demand value $\boldsymbol{DM_i(t)}$ will always remain, even after adjustment for price effects (for instance the range could be equal to the reference demand $\boldsymbol{DM^o_i(t)}$ plus or minus 50%). The smallest range value is denoted $\boldsymbol{DM(t)_{min}}$.

b) Select a grid that divides each range into a number n of equal width intervals. Let $ß_i(t)$ be the resulting common width of the grid, $\beta_i(t)= R_i(t)/n$. See Figure 6.1 for a sketch of the non-linear expression and of its step-wise constant approximation. The number of steps, n, should be chosen so that the step-wise constant approximation remains close to the exact value of the function.

c) For each demand segment $\boldsymbol{DM_i(t)}$ define n step-variables (one per grid interval), denoted $s_{1,i}(t), s_{2,i}(t), \ldots, s_{n,i}(t)$. Each s variable is bounded below by 0 and above by $\beta_i(t)$. One may now replace in equations (6-6)' and (6-7)' each $\boldsymbol{DM_i(t)}$ variable by the sum of the n-step variables, and each non-linear term in the objective function by a weighted sum of the n step-variables, as follows:

$$DM_i(t) = DM(t)_{\min} + \sum_{j=1}^{n} s_{j,i}(t) \qquad 6-9$$

and

$$DM_i(t)^{1+1/E_i} \cong DM(t)_{\min}^{1+1/E_i} + \sum_{j=1}^{n} A_{j,s,i}(t) \bullet s_{j,i}(t) / \beta_i(t) \qquad 6-10$$

The resulting Mathematical Program is now fully linearized.

Remark: instead of maximizing the linearized objective function, TIMES minimizes its negative, which then has the dimension of a cost. The portion of that cost representing the negative of the consumer surplus is akin to a *welfare loss.*

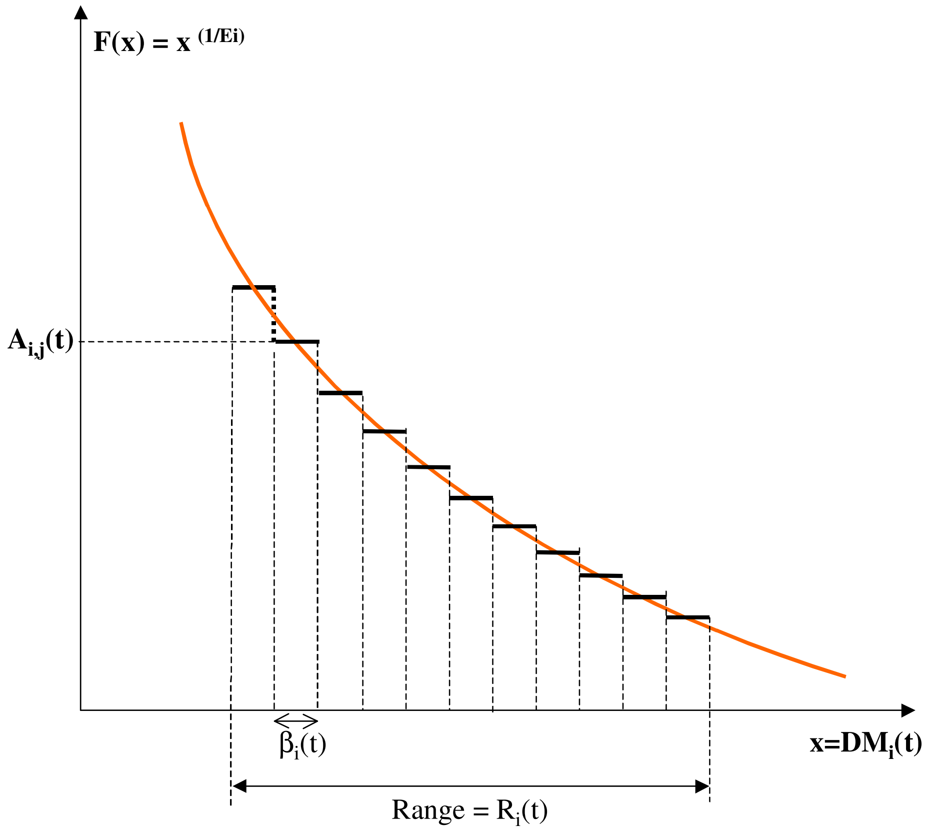

Figure 6.1. Step-wise approximation of the non-linear terms in the objective function

6.2.4 Calibration of the demand functions

Besides selecting elasticities for the various demand categories, the user must evaluate each constant $K_i(t)$. To do so, we have seen that one needs to know one point on each demand function in each time period, $\{ p^0_i(t), DM^0_i(t) \}$. To determine such a point, we perform a single preliminary run of the inelastic TIMES model (with exogenous $DM^0_i(t)$), and use the resulting shadow prices $p^0_i(t)$ for all demand constraints, in all time periods for each region.

6.2.5 Computational considerations

Each demand segment that is elastic to its own price requires the definition of as many variables as there are steps in the discrete representation of the demand curve (both upward and down if desired), for each period and region. Each such variable has an upper bound, but is otherwise involved in no new constraint. Therefore, the linear program is augmented by a number of variables, but does not have more constraints than the initial inelastic LP (with the exception of the upper bounds). It is well known that with the modern LP codes the number of variables has little or no impact on computational time in Linear Programming, whether the variables are upper bounded or not. Therefore, the inclusion in TIMES of elastic demands has a very minor impact on computational time or on the tractability of the resulting LP. This is an important observation in view of the very large LP's that result from representing multi-regional and global models in TIMES.

6.2.6 Interpreting TIMES costs, surplus, and prices

It is important to note that, instead of maximizing the net total surplus, TIMES minimizes its negative (plus a constant), obtained by changing the signs in expression (6-6). For this and other reasons, it is inappropriate to pay too much attention to the meaning of the *absolute* objective function values. Rather, examining the difference between the objective function values of two scenarios is a far more useful exercise. That difference is of course, the negative of the difference between the net total surpluses of the two scenario runs.

Note again that the popular interpretation of shadow prices as the *marginal costs* of model constraints is inaccurate. Rather, the shadow price of a constraint is, by definition, the incremental value of the objective function per unit of that constraint's right hand side (RHS). The interpretation is that of an amount of *surplus loss* per unit of the constraint's RHS. The difference is subtle but nevertheless important. For instance, the shadow price of the electricity balance constraint is not necessarily the marginal cost of producing electricity. Indeed, when the RHS of the balanced constraint is increased by one unit, one of two things may occur: either the system produces one more unit of electricity, or else the system consumes one unit less of electricity (perhaps by choosing more efficient end-use devices or by reducing an electricity-intensive energy service, etc.). It is therefore correct to speak of shadow prices as the marginal *system value* of a resource, rather than the marginal *cost* of procuring that resource.

7 The Lumpy Investment option

In some cases, the linearity property of the TIMES model may become a drawback for the accurate modeling of certain investment decisions. Consider for example a TIMES model for a relatively small community such as a city. For such a scope the *granularity* of some investments may have to be taken into account. For instance, the size of an electricity generation plant proposed by the model would have to conform to an implementable minimum size (it would make no sense to decide to construct a 50 MW nuclear plant). Another example for multi-region modelling might be whether or not to build cross-region electric grid(s) or gas pipeline(s) in discrete size increments. Processes subject to investments of only specific size increments are described as “lumpy” investments.

For other types of investments, size does not matter: for instance the model may decide to purchase 10,950.52 electric cars, which is easily rounded to 10,950 without any serious inconvenience. The situation is similar for a number of residential or commercial heating devices, or for the capacity of wind turbines or industrial boilers, or for any technologies with relatively small minimum feasible sizes. Such technologies would not be candidates for treatment as “lumpy” investments.

It is the user’s responsibility to decide that certain technologies should (or should not) respect the minimum size constraint, weighing the pros and cons of so doing. This chapter explains how the TIMES LP is transformed into a Mixed Integer Program (MIP) to accommodate minimum or multiple size constraints, and states the consequences of so doing on computational time and on the interpretation of duality results.

The lumpy investment option available in TIMES is slightly more general than the one described above. It insures that investment in technology $\boldsymbol{k}$ is equal to one of a finite number $\boldsymbol{N}$ of pre-determined sizes: $\boldsymbol{0, S_1(t), S_2(t), \ldots, S_N(t)}$. As implied by the notation, these discrete sizes may be different at different time periods. Note that by choosing the $\boldsymbol{N}$ sizes as the successive multiples of a fixed number $\boldsymbol{S}$, it is possible to invest (perhaps many times) in a technology with fixed standard size.

Imposing such a constraint on an investment is unfortunately impossible to formulate using standard LP constraints and variables. It requires the introduction of *integer variables* in the formulation. The optimization problem resulting from the introduction of integer variables into a Linear Program is called a Mixed Integer Program (MIP).

7.1 Formulation and Solution of the Mixed Integer Linear Program

Typically, the modeling of a lumpy investment involves Integer Variables, i.e. variables whose values may only be non-negative integers (0, 1, 2, …). The mathematical formulation is as follows

$$NCAP(p,t) = \sum_{i=1}^{N} S_i(p,t) \times Z_i(p,t) \qquad each\ t = 1,..,T$$

with

$$Z_i(p,t) = 0\ or\ 1$$

and

$$\sum_{i=1}^{N} Z_i(p,t) \leq 1$$

The second and third constraints imply that at most one of the ***Z*** variables is equal to 1. Thereforew, the first constraint now means that ***NCAP*** is equal to one of the preset sizes or is equal to 0, which is the desired result.

Although the formulation of lumpy investments *looks* simple, it has a profound effect on the resulting optimization program. Indeed, MIP problems are notoriously more difficult to solve than LPs, and in fact many of the properties of linear programs discussed in the preceding chapters do not hold for MIPs, including duality theory, complementary slackness, etc. Note that the constraint that ***Z(p,t)*** should be 0 or 1 departs from the *divisibility* property of linear programs. This means that the *feasibility domain* of integer variables (and therefore of some investment variables) is no longer contiguous, thus making it vastly more difficult to apply purely algebraic methods to solve MIP's. In fact, practically all MIP solution algorithms make use (at least to some degree) of partial enumerative schemes, which tend to be time consuming and less reliable[32] than the algebraic methods used in LP.

The reader interested in more technical details on the solution of LPs and of MIPs is referred to references (Hillier and Lieberman, 1990, Nemhauser et al. 1989). In the next section we shall be content to state one important remark on the interpretation of the dual results from MIP optimization.

7.2 Important remark on the MIP dual solution (shadow prices)

Using MIP rather than LP has an important impact on the interpretation of the TIMES shadow prices. Once the optimal MIP solution has been found, it is customary for MIP solvers to fix all integer variables at their optimal (integer) values, and to perform one

[32] A TIMES LP program of a given size tends to have fairly constant solution time, even if the database is modified. In contrast, a TIMES MIP may show some erratic solution times. One may observe reasonable solution times (although significantly longer than LP solution times) for most instances, with an occasional *very* long solution time for some instances. This phenomenon is predicted by the theory of complexity as applied to MIP, see Papadimitriou and Steiglitz (1982)

additional iteration of the LP algorithm, so as to obtain the dual solution (i.e. the shadow prices of all constraints). However, the interpretation of these prices is different from that of a LP. Consider for instance the shadow price of the natural gas balance constraint: in a pure LP, this value represents the price of natural gas. In MIP, this value represents the price of gas *conditional on having fixed the lumpy investments at their optimal integer values.* What does this mean? We shall attempt an explanation via one example: suppose that one lumpy investment was the investment in a gas pipeline; then, *the gas shadow price will not include the investment cost of the pipeline, since that investment was fixed when the dual solution was computed.*

In conclusion, when using MIP, only the primal solution is fully reliable. In spite of this major caveat, modeling lumpy investments may be of paramount importance in some instances, and may thus justify the extra computing time and the partial loss of dual information.

8 Endogenous Technological Learning (ETL)

In a long-term dynamic model such as TIMES the characteristics of many of the future technologies are almost inevitably changing over the sequence of future periods due to *technological learning*.

In some cases it is possible to forecast such changes in characteristics as a function of time, and thus to define a time-series of values for each parameter (e.g. unit investment cost, or efficiency). In such cases, technological learning is *exogenous* since it depends only on time elapsed and may thus be established outside the model.

In other cases there is evidence that the pace at which some technological parameters change is dependent on the *experience* acquired with this technology. Such experience is not solely a function of time elapsed, but typically depends on the cumulative investment (often global) in the technology. In such a situation, technological learning is *endogenous*, since the future values of the parameters are no longer a function of time elapsed alone, but depend on the cumulative investment decisions taken by the model (which are unknown). In other words, the evolution of technological parameters may no longer be established outside the model, since it depends on the model's results. ETL is also named *Learning-By-Doing* (LBD) by some authors.

Whereas exogenous technological learning does not require any additional modeling, endogenous technological learning (ETL) presents a tough challenge in terms of modeling ingenuity and of solution time. In TIMES, there is a provision to represent the effects of endogenous learning on the unit investment cost of technologies. Other parameters (such as efficiency) are not treated, at this time.

8.1 The basic ETL challenge

Empirical studies of unit investment costs of several technologies have been undertaken in several countries. Many of these studies find an empirical relationship between the unit investment cost of a technology at time $\boldsymbol{t}$, $\boldsymbol{INVCOST_t}$, and the cumulative investment in that technology up to time $\boldsymbol{t}$, $\boldsymbol{C_t} = \sum_{j=-1}^{t} \boldsymbol{VAR\_INV_j}$.

A typical relationship between unit investment cost and cumulative investments is of the form:

$$INVCOST_t = a \cdot C_t^{-b} \qquad (8-1)$$

where $\boldsymbol{a}$ is the initial unit investment cost (when $\boldsymbol{C_t}$ is equal to 1) and $\boldsymbol{b}$ is the learning index, representing the speed of learning[33]. As experience builds up, the unit investment

[33] It is usual to define, instead of $\boldsymbol{b}$, another parameter, $\boldsymbol{pr}$ called the *progress ratio*, which is related to $\boldsymbol{b}$ via the following relationship:

$$pr = 2^{-b}$$

cost decreases, and thus may make investments in the technology more attractive. It should be clear that near-sighted investors will not be able to detect the advantage of investing early in learning technologies, since they will only observe the high initial investment cost and, being near-sighted, will not anticipate the future drop in investment cost resulting from early investments. In other words, tapping the full potential of technological learning requires far-sighted agents who accept making initially non-profitable investments in order to later benefit from the investment cost reduction.

With regard to actual implementation, simply using (8-1) as the objective function coefficient of $\boldsymbol{VAR\_INV_t}$ will yield a non-linear, non-convex expression. Therefore, the resulting mathematical optimization is no longer linear, and requires special techniques for its solution. In TIMES, a Mixed Integer Programming (MIP) formulation is used, that we now describe.

8.2 The TIMES formulation of ETL

8.2.1 The Cumulative Investment Cost

We follow the basic approach described in Barreto, 2001

The first step of the formulation is to express the total investment cost, i.e. the quantity that should appear in the objective function for the investment cost of a learning technology in period $\boldsymbol{t}$. $\boldsymbol{TC_t}$ is obtained by integrating expression (8-1):

$$TC_t = \int_0^{C_t} a \cdot y^{-b} * dy = \frac{a}{1-b} \cdot C_t^{-b+1} \qquad (8-2)$$

$\boldsymbol{TC_t}$ is a concave function of $\boldsymbol{C_t}$, with a shape as shown in figure 8.1

Hence, $\boldsymbol{1-pr}$ is the cost reduction incurred when cumulative investment is doubled. Typical observed $\boldsymbol{pr}$ values are in a range of .75 to .95

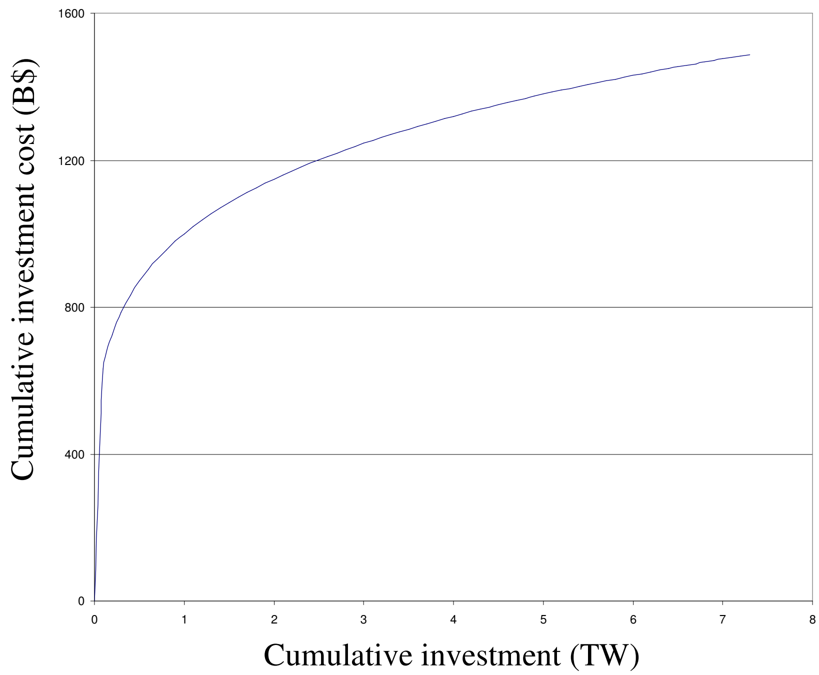

Figure 8.1. **Example of a cumulative learning curve**

With the Mixed Integer Programming approach implemented in TIMES, the cumulative learning curve is approximated by linear segments, and binary variables are used to represent some logical conditions. Figure 8.2 shows a possible piecewise linear approximation of the curve of Figure 8.1. The choice of the number of steps and of their respective lengths is carefully made so as to provide a good approximation of the smooth cumulative learning curve. In particular, the steps must be smaller for small values than for larger values, since the curvature of the curve diminishes as total investment increases. The formulation of the ETL variables and constraints proceeds as follows (we omit the period, region, and technology indexes for notational clarity):

1. The user specifies the set of learning technologies (TEG).
2. For each learning technology, the user provides:
   a) The progress ratio $\boldsymbol{pr}$ (from which the learning index $\boldsymbol{b}$ may be inferred)
   b) One initial point on the learning curve, denoted (C_0, TC_0)
   c) The maximum allowed cumulative investment C_{max} (from which the maximum total investment cost TC_{max} may be inferred)
   d) The number N of segments for approximating the cumulative learning curve over the (C_0, C_{max}) interval (note that N may be different for different technologies).
3. The model automatically selects appropriate values for the N step lengths, and then proceeds to generate the required new variables and constraints, and the new objective function coefficients for each learning technology. The detailed formulae are shown and briefly commented on below.

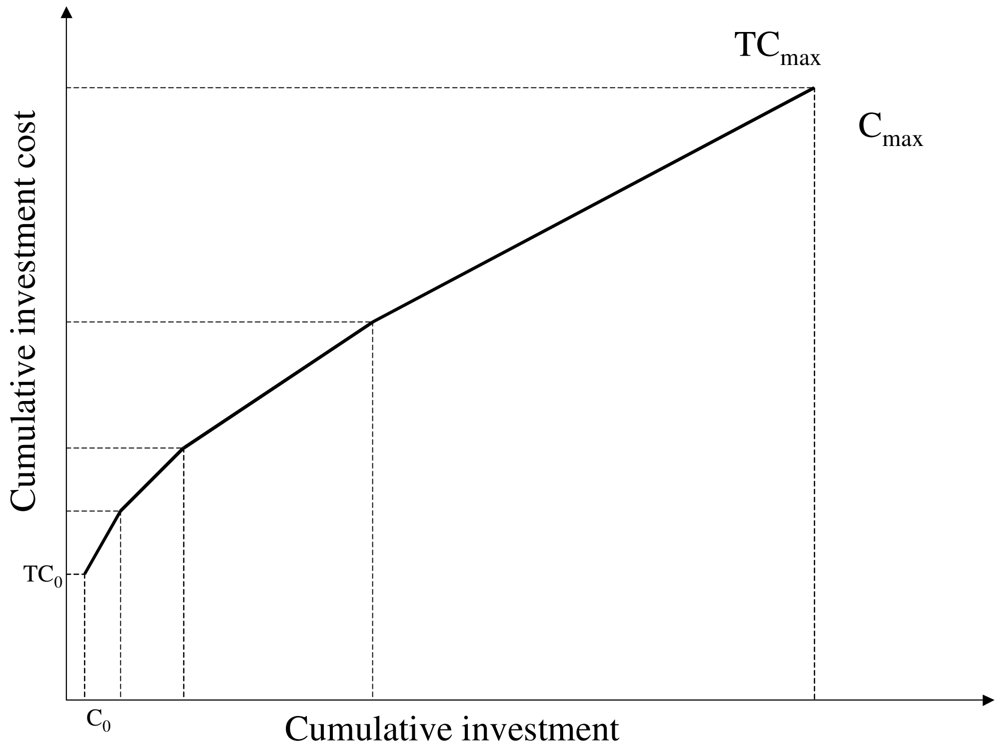

Figure 8-2. Example of a 4-segment approximation of the cumulative cost curve

8.2.2 Calculation of break points and segment lengths

The successive interval lengths on the vertical axis are chosen to be in geometric progression, each interval being twice at wide as the preceding one. In this fashion, the intervals near the low values of the curve are smaller so as to better approximate the curve in its high curvature zone. Let $\{TC_{i-1}, TC_i\}$ be the i^{th} interval on the vertical axis, for $i = 1, \ldots, N\text{-}1$. Then:

$$TC_i = TC_{i-1} + 2^{i-N-1}(TC_{\max} - TC_o)/(1-0.5^N)\ , \qquad i = 1,2,\cdots,N$$

Note that TC_{max} is equal to TC_N.

The break points on the horizontal axis are obtained by plugging the TC_i 's into expression (1.10-2), yielding:

$$C_i = \left(\frac{(1-b)}{a}(TC_i)\right)^{\frac{1}{1-b}}\ , \qquad i = 1,2,\ldots,N$$

8.2.3 New variables

Once intervals are chosen, standard approaches are available to represent a concave function by means of integer (0-1) variables. We describe the approach used in TIMES. First, we define N continuous variables $\boldsymbol{x_i}$, $i= 1,\ldots,N$. Each $\boldsymbol{x_i}$ represents the portion of cumulative investments lying in the $\boldsymbol{i^{th}}$ interval. Therefore, the following holds:

$$C = \sum_{i=1}^{N} x_i \qquad 8-3$$

We now define $\boldsymbol{N}$ integer (0-1) variables $\boldsymbol{z_i}$ that serve as indicators of whether or not the value of $\boldsymbol{C}$ lies in the $\boldsymbol{i^{th}}$ interval. We may now write the expression for $\boldsymbol{TC}$, as follows:

$$TC = \sum_{i=1}^{N} a_i z_i + b_i x_i \qquad 8-4$$

where $\boldsymbol{b_i}$ is the slope of the $\boldsymbol{i^{th}}$ line segment, and $\boldsymbol{a_i}$ is the value of the intercept of that segment with the vertical axis, as shown in figure 8.3. The precise expressions for $\boldsymbol{a_i}$ and $\boldsymbol{b_i}$ are:

$$\begin{aligned} b_i &= \frac{TC_i - TC_{i-1}}{C_i - C_{i-1}} \qquad && i = 1,2,\ldots,N \\ a_i &= TC_{i-1} - b_i \cdot C_{i-1} \qquad && i = 1,2,\ldots,N \end{aligned} \qquad 8-5$$

8.2.4 New constraints

For (1.8-4) to be valid we must make sure that exactly one $\boldsymbol{z_i}$ is equal to 1, and the others equal to 0. This is done (recalling that the $\boldsymbol{z_i}$ variables are 0-1) via:

$$\sum_{i=1}^{N} z_i = 1$$

We also need to make sure that each $\boldsymbol{x_i}$ lies within the i^{th} interval whenever $\boldsymbol{z_i}$ is equal to 1 and is equal to 0 otherwise. This is done via two constraints:

$$\boldsymbol{C_{i-1} \cdot z_i \le x_i \le C_i \cdot z_i}$$

8.2.5 Objective function terms

Re-establishing the period index, we see that the objective function term at period $\boldsymbol{t}$, for a learning technology is thus equal to $\boldsymbol{TC_t - TC_{t-1}}$, which needs to be discounted like all other investment costs.

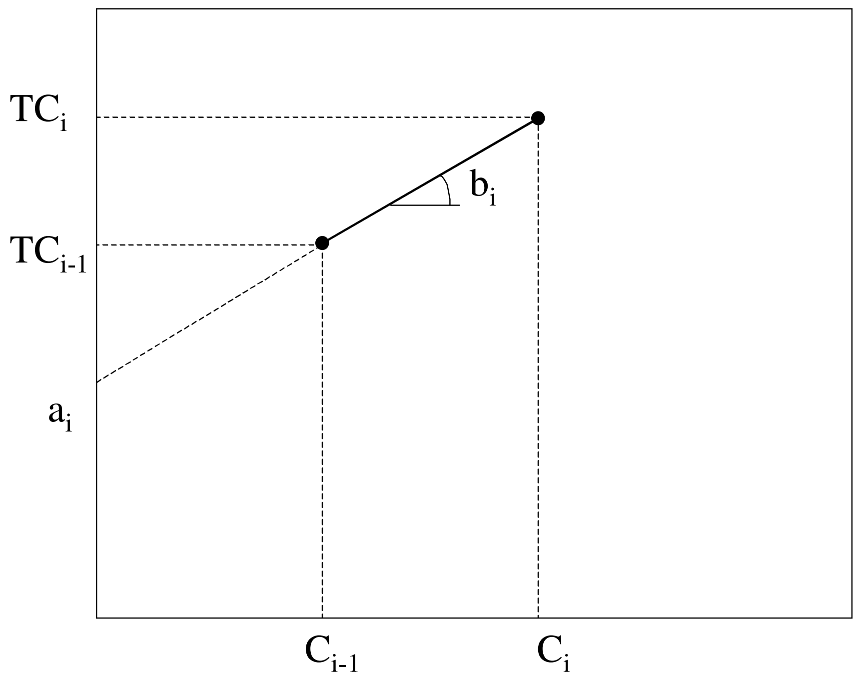

Figure 8.3. The ith segment of the step-wise approximation

8.2.6 Additional (optional) constraints

Solving integer programming problems is facilitated if the domain of feasibility of the integer variables is reduced. This may be done via additional constraints, that are not strictly needed but that are guaranteed to hold. In our application we know that experience (i.e. cumulative investment) is always increasing as time goes on. Therefore, if the cumulative investment in period t lies in segment i, it is certain that it will not lie in segments *i-1, i-2, .., 1* in time period *t+1*. This leads to two new constraints (re-establishing the period index $\boldsymbol{t}$ for the z variables):

$$\sum_{j=1}^{i} z_{j,t} \geq \sum_{j=1}^{i} z_{j,t+1}$$

$$\boldsymbol{i} = 1,2,...,\boldsymbol{N}-1,\ \boldsymbol{t} = 1,2,...,\boldsymbol{T}-1$$

$$\sum_{j=i}^{N} z_{j,t} \leq \sum_{j=i}^{N} z_{j,t+1}$$

Summarizing the above formulation, we observe that each learning technology requires the introduction of ***N*T*** integer (0-1) variables. For example, if the model has 10 periods and a 5-segment approximation is selected, 50 integer (0-1) variables are created for that learning technology, assuming that the technology is available in the first period of the model. Thus, the formulation may become very onerous in terms of solution time, if many learning technologies are envisioned, and if the model is of large size to begin with. In section 8.5 we provide some comments on ETL, as well as a word of warning.

8.3 Clustered learning

An interesting variation of ETL is also available in TIMES, namely the case where several technologies use the same key technology (or component), itself subject to learning. For instance, table 8.1 lists 11 technologies using the key Gas Turbine technology. As experience builds up for gas the turbine, each of the 11 technologies in the cluster benefits. The phenomenon of clustered learning is modeled in TIMES via the following modification of the formulation of the previous section.

Let ***k*** designate the key technology and let ***l* = 1, 2, …, *L*** designate the set of clustered technologies attached to ***k***. The approach consists of three steps:

i) Step 1: designate ***k*** as a learning technology, and write for it the formulation of the previous section;

ii) Step 2: subtract from each $\boldsymbol{INVCOST_l}$ the initial investment cost of technology ***k*** (this will avoid double counting the investment cost of ***k***);

iii) Step 3: add the following constraint to the model, in each time period. This ensures that learning on ***k*** spreads to all members of its cluster:

$$\boldsymbol{VAR\_INV_k} - \sum_{l=1}^{L} \boldsymbol{VAR\_INV_l} = 0$$

Table 8-1: Cluster of gas turbine technologies (from A. Sebregts and K. Smekens, unpublished report, 2002)

Description
Integrated Coal gasification power plant
Integrated Coal Gasification Fuel Cell plant
Gas turbine peaking plant
Existing gas Combined Cycle power plant
New gas Combined Cycle power plant
Combined cycle Fuel Cell power plant
Existing gas turbine CHP plant
Existing Combined Cycle CHP plant
Biomass gasification: small industrial cog.
Biomass gasification: Combined Cycle power plant
Biomass gasification: ISTIG+reheat

8.4 Learning in a Multiregional TIMES Model

Technological learning may be acquired via global or local experience, depending on the technology considered. There are examples of techniques that were developed and perfected in certain regions of the World, but have tended to remain regional, never fully spreading globally. Examples are found in land management, irrigation, and in household heating and cooking devices. Other technologies are truly global in the sense that the same (or close to the same) technology becomes rather rapidly commercially available globally. In the latter case, global experience benefits users of the technology world wide. Learning is said to *spillover* globally. Examples are found in large electricity plants, in steel production, wind turbines, and many other sectors.

The first and obvious implication of these observations is that the appropriate model scope must be used to study either type of technology learning.

The formulation described in the previous sections is adequate in two cases: a) learning in a single region model, and b) regional learning in a multiregional model. It does *not* directly apply to *global learning* in a multiregional global model, where the cumulative investment variable must represent the sum of all cumulative investments in all regions together. We now describe an approach to global learning that may be implemented in TIMES, using only standard TIMES entities.

The first step in modeling multiregional ETL (MRETL) is to create one additional region, region 0, which will play the role of the Manufacturing Region. This region's RES consists only of the set of (global) learning technologies (LT's). Each such LT has the following specifications:

a) The LT has no commodity input
b) The LT has only one output, a new commodity ***c*** representing the 'learning'. This output is precisely equal to the investment level in the LT in each period.
c) Commodity ***c*** may be exported to all other regions

Finally, in each ‘real’ region, the LT is represented with all its attributes *except the investment cost NCAP_COST* Furthermore, the construction of one unit of the LT requires an input of one unit of the learning commodity ***c*** (using the *NCAP_ICOM* parameter see chapter 3 of PART II). This ensures that the sum of all investments in the LT in the real regions is exactly equal to the investment in the LT in region 0, as desired.

8.5 Endogenous vs. Exogenous Learning: Discussion

In this section, we formulate a few comments and warnings that may be useful to potential users of the ETL feature.

We start by stating a very important caveat to the ETL formulation described in the previous sections: if a model is run with such a formulation, it is very likely that the model will select some technologies, and *will invest massively at some early period* in these technologies unless it is prevented from doing so by additional constraints. Why this is likely to happen may be qualitatively explained by the fact that once a learning technology is selected for investing, two opposing forces are at play in deciding the optimal timing of the investments. On the one hand, the discounting provides an incentive for postponing investments. On the other hand, investing early allows the unit investment cost to drop immediately, and thus allows much cheaper investments in the learning technologies in the current and all future periods. Given the considerable cost reduction that is usually induced by learning, the first factor (discounting) is highly unlikely to predominate, and hence the model will tend to invest massively and early in such technologies, or not at all. Of course, what we mean by “massively” depends on the other constraints of the problem (such as the extent to which the commodity produced by the learning technology is in demand, the presence of existing technologies that compete with the learning technology, etc.). However, there is a clear danger that we may observe unrealistically large investments in some learning technologies.

ETL modelers are well aware of this phenomenon, and they use additional constraints to control the penetration trajectory of learning technologies. These constraints may take the form of upper bounds on the capacity of or the investment in the learning technologies in each time period, reflecting what is considered by the user to be realistic penetrations. These upper bounds play a determining role in the solution of the problem, and it is most often observed that the capacity of a learning technology is either equal to 0 or to the upper bound. This last observation indicates that the selection of upper bounds (or capacity/investment growth rates) by the modeler is the predominant factor in controlling the penetration of successful learning technologies.

In view of the preceding discussion, a fundamental question arises: is it worthwhile for the modeler to go to the trouble of modeling e*ndogenous* learning (with all the attendant computational burdens) when the results are to a large extent conditioned by *exogenous* upper bounds? We do not have a clear and unambiguous answer to this question; that is left for each modeler to evaluate.

However, given the above caveat, a possible alternative to ETL would consist in using exogenous learning trajectories. To do so, the same sequence of ‘realistic’ upper bounds on capacity would be selected by the modeler, and the values of the unit investment costs (INVCOST) would be externally computed by plugging these upper bounds into the learning formula (8-1). This approach makes use of the same exogenous upper bounds as the ETL approach, but avoids the MIP computational burden of ETL. Of course, the running of exogenous learning scenarios is not entirely foolsafe, since there is no absolute guarantee that the capacity of a learning technology will turn out to be exactly equal to its exogenous upper bound. If that were not the case, a modified scenario would have to be run, with upper bounds adjusted downward. This trial-and-error approach may seem inelegant, but it should be remembered that it (or some other heuristic approach) might prove to be necessary in those cases where the number of learning technologies and the model size are both large (thus making the rigorous ETL formulation computationally intractable).

References for Part I

- Altdorfer, F., "Introduction of price elasticities on energy demand in MARKAL", *Memorandum No 345, KFA*, Julich, July 1981.
- Barreto, L. (2001), "Technological Learning in Energy Optimisation Models and the Deployment of Emerging Technologies", Ph.D. Thesis No 14151, Swiss Federal Institute of Technology Zurich (ETHZ). Zurich, Switzerland.
- Berger, C., R. Dubois, A. Haurie, E. Lessard, R. Loulou, and J.-Ph. Waaub, "Canadian MARKAL: an Advanced Linear Programming System for Energy and Environment Modelling", *INFOR*, Vol. 20, 114-125, 1992.
- Birge, J. R., and Rosa, C. H. (1996), "Incorporating Investment Uncertainty into Greenhouse Policy Models," *The Energy Journal*, 17/1, 79-90.
- Chvátal, V., *Linear Programming*, Freeman and Co, New-York, 1983
- Dantzig, G.B., *Linear programming and Extensions*, Princeton University Press, Princeton, New-Jersey, 1963.
- Department of Energy/Energy Information Administration, US Government, "The National Energy Modeling System : An overview (2000)", *DOE/IEA-0581*, Washington. DC, March 2000
- Dorfman, R., P.A. Samuelson, and R.M. Solow, *Linear programming and Economic Analysis*, McGraw-Hill, New-York, 1958.
- Edmonds, J.A., H.M. Pitcher, D. Barns, R. Baron, and M.A. Wise, "Modelling future greenhouse gas emissions: The second generation model description", in *Modelling global change*, L.R. Klein and Fu-Chen Lo eds, United Nations University Press, Tokyo-New-York-Paris, 1991, 292-362.
- Fankhauser, S. (1994), "The social cost of GHG emissions: an expected value approach", *Energy Journal 15/2*, 157-184.
- Fishbone L G, Giesen G, Hymmen H A, Stocks M, Vos H, Wilde, D, Zoelcher R, Balzer C, and Abilock H. (1983), "Users Guide for MARKAL: A Multi-period, Linear Programming Model for Energy Systems Analysis", BNL Upton, NY, and KFA, Julich, Germany, BNL 51701.
- Fishbone, L.G., and H. Abilock, (1981), "MARKAL, A Linear Programming Model for Energy Systems Analysis: Technical Description of the BNL Version", *International Journal of Energy Research*, Vol. 5, 353-375.
- Fragnière, E. and Haurie, A. (1996), "MARKAL-Geneva: A Model to Assess Energy-Environment Choices for a Swiss Canton", in C. Carraro and A. Haurie (eds.), *Operations Research and Environmental Management*, Kluwer Academic Books.
- Gale, D., *The Theory of Linear Economic Models*, McGraw-Hill, New-York, 1960.
- Gerking. H., and Voss, A. (1986), "An Approach on How to Handle Incomplete Foresight in Linear Programming Models," International Conference on Models and Uncertainty in the Energy Sector, Risö National Laboratory, Denmark, 11-12 February, 1986, p. 145.
- Gorenstin, B. G. et al. (1993), "Power System Expansion Planning Under Uncertainty," IEEE Transactions on Power Systems, 8, 1, pp 129-136.
- Grubb, M., (1993), "Policy Modelling for Climate Change: The Missing Models," Energy Policy, 21, 3, pp 203-208.

- Hillier, F. S, and G. L. Lierberman, *Introduction to Operations Research, Fifth Edition,* McGraw-Hill, New-York, 1990
- Hogan, W.W., (1975), "Energy Policy Models for Project Independence", *Computers and Operations Research,* 2, 251-271.
- Jaccard, M., J. Nyboer, C. Bataille, and B. Sadownik (2003). "Modelling the Cost of Climate Policy: Distinguishing between alternative cost definitions and long-run cost dynamics." Energy Journal 24(1): 49-73.
- Kanudia, A., and R. Loulou, "Robust Responses to Climate Change via Stochastic MARKAL: the case of Québec", *European Journal of Operations Research*, vol. 106, pp. 15-30, 1998
- Kunsch, P. L. and Teghem, J., Jr. (1987), "Nuclear Fuel Cycle Optimization Using Multi-Objective Stochastic Linear Programming," *European Journal of Operations Research 31/2*, 240-249.
- Larsson, T. and Wene, C.-O. (1993), "Developing Strategies for Robust Energy Systems. I: Methodology," *International Journal of Energy Research 17*, 503-513.
- Larsson, T. (1993), "Developing Strategies for Robust Energy Systems. II: Application to CO_2 Risk Management," *International Journal of Energy Research 17*, 505-535.
- Loulou, R., and A. Kanudia, "Minimax Regret Strategies for Greenhouse Gas Abatement: Methodology and Application", *Operations Research Letters,* 25, 219-230, 1999.
- Loulou, R. and A. Kanudia, "Using Advanced Technology-rich models for Regional and Global Economic Analysis of GHG Mitigation", in *Decision and Control: Essays in honor of Alain Haurie,* G. Zaccour ed., Kluwer Academic Publishers, Norwell, USA, pp.153-175, 2000. A condensed version of this article was published electronically in the *Proceedings of the International Energy Agency International conference on Climate Change Modelling and Analysis*, held in Washington DC, June15-17, 1999.
- Loulou, R., and Lavigne, D., "MARKAL Model with Elastic Demands: Application to GHG Emission Control", in *Operations Research and Environmental Engineering*, C. Carraro and A. Haurie eds., Kluwer Academic Publishers, Dordrecht, Boston, London, 1996, pp. 201-220.
- Manne, A., Mendelsohn, R. and Richels, R. (1995), "MERGE: a Model for Evaluating Regional and Global Effects of GHG Reduction Policies", *Energy Policy* 23 (1), 17-34.
- Manne, A.S., and Wene, C-O., "MARKAL-MACRO: A Linked Model for Energy-Economy Analysis", *BNL-47161 report, Brookhaven National Laboratory*, Upton, New-York, February 1992.
- Meade, D. (1996) "LIFT User's Guide", Technical report, INFORUM, IERF, College Park, Md (July 1996).
- Messner, S. and Strubegger, M. (1995), User's Guide for MESSAGE III, WP-95-69, IIASA, Luxembourg, Austria.
- Moomaw, W.R. and J.R. Moreira. 2001. « Technological and Economic Potential of Greenhouse Greenhouse Gas Emissions Reduction ». In *Climate Change 2001: Mitigation*, edited by Metz, B., Davidson, O., Swart, R. and J. Pan, Intergovernmental Panel on Climate Change (IPCC), Third Assessment Report, Working Group III, p.167-299. Cambridge : Cambridge University Press.

- Nakicenovic, N. (ed.). 2000. *Special Report on Emissions Scenarios. A Special Report of Working III of the Intergovernmental Panel on Climate Change*. Cambridge : Cambridge University Press, p.599.
- Nordhaus, W. (1993), "Rolling the DICE: an optimal transition path for controlling GHG's", *Resources and Energy Economics 15/1*, 27-50.
- Nemhauser, G.L., A.H.G. Rinnooy Kan, and M.J. Todd (editors), *Handbooks in Operations Research and Management Science*, Vol I: Optimization, North-Holland, 1989.
- Papadimitriou, C.H., and K. Steiglitz, *Combinatorial Optimization -- Algorithms and Complexity*, Prentice-Hall, New-Jersey, 1982.
- Peck, S. C., and Teisberg,T. J. (1995), "Optimal CO 2 Control Policy with Stochastic Losses from Temperature Rise," *Climatic change*, 31/1, 19-34.
- Raiffa, H., *Decision Analysis*, Addison-Wesley, Reading, Mass., 1968
- Samuelson, P.A., "Spatial Price Equilibrium and Linear Programming", *American Economic Review*, 42, 283-303, 1952.
- Samuelson, P.A., and W. Nordhaus, *Economics (17th edition)*, John Wiley, 1977
- Takayama, T., and Judge G.G., *Spatial and Temporal Price and Allocation Models*, North Holland, Amsterdam, 1971.
- Tosato, G.C., "Extreme Scenarios in MARKAL LP Model: use of Demand Elasticity", presented at the *5th Italian-Polish Symposium on Applications of Systems Theory to Economics and Technology*, Torun, June 11-16 1980.
- Voort, E. van der, Donni, E., Thonet, C., Bois d'Enghien, E., Dechamps, C. & Guilmot, J.F. 1984. Energy Supply Modelling Package EFOM-12C Mark I, Mathematical description. Louvain-la-Neuve, Cabay: Commission of the European Communities, EUR-8896.