INDUSTRIALIZATION AND PRODUCTIVITY

4D

In this issue:

BULLETIN No. 22

UNITED NATIONS

21

Mathematical programming models in industrial development planning

by Janos KORNAI*

Introduction

This article is based on the models employed in Hungary since the late 1950s for the mathematical planning of industrial development. In constructing the models, the special conditions of the country had to be considered, of course. Hungary has a centrally planned economy; the five-year economic plans contain detailed estimates covering most industrial investment projects. It is a country in which agriculture still plays the major role, but in which manufacturing has developed at an accelerated rate over the past 10-20 years. The Government has set itself the task of improving the country's position in the international division of labour and achieving a high growth rate.

The endowments and problems of Hungary are similar to those of a number of other countries. Therefore, the mathematical models employed in Hungary should also be useful beyond its frontiers. Although the concrete structure and numerical data of these models reflect the special characteristics of that country, numerous features of the mathematical planning methodology evolved there are generally applicable. These features of general validity will be stressed in this article. It is the author's intention that the reader should become acquainted, not just with the Hungarian methods but with a tested methodology that may be applied in any country.

This article first describes the programming model of one sector, then explains the linking of the models of several sectors and the methods of their collective planning. Next, the application possibilities of the models are described, and finally, some of the special difficulties encountered in model construction are discussed. It is assumed throughout that the reader has a knowledge of the conceptual apparatus of mathematical programming and of matrix algebra.

*Computing Centre of the Hungarian Academy of Sciences, Budapest.

I. The sector model

Suppose that a group of mathematical planners is assigned to construct a model of the development plans for one of the country's industrial sectors; it does not, at the moment, matter which. They must first ask themselves what questions they are trying to answer with the aid of the model. We shall call these questions "decision problems".

A. *Decision problems*

The type of model to be described below is designed to give a simultaneous answer to the following eight decision problems:

1. What products should the sector be able to produce? What should the product pattern of the sector's total output be? How much of each product should the sector produce?

The problem here is not to determine the production programme of a factory in full detail for years or decades in advance. This will be worked out in due course by the factory's programming department in yearly or monthly breakdowns. In our problem, the products have to be aggregated into product groups that must be distinguished from each other from the point of view of investment decisions. Thus, in the model of the metallurgical sector we must plan the quantity of crude iron, steel, plate etc. that the country should produce. However, it will not be necessary to decide on the proportion of the plates of 5-mm and 6-mm thickness. The former, more general proportions affect investment; the latter, more detailed proportions affect only the specific structure of production carried out on the basis of a given investment capital and so do not fall within the scope of long-term industrial development planning.

Decision problem 1 is thus the determination of the output structure of production.

2. The first problem was what to produce and how much of it. However, the question, How should

production be carried out? must also be answered; the planning of sector development also includes technical development planning. Production could continue with the present technology. In the case of a sector lagging behind international technical standards, "present technology" usually means a comparatively labour-intensive technology. (This may have its advantages from the points of view of securing employment and of requiring comparatively low amounts of capital, however.) Or, a switch to more advanced technologies could be made. In this case productivity performance will increase, but procuring the necessary capital and skilled labour may pose a problem.

It is, of course, not only in the degree of productivity and of labour and capital intensity that these alternatives may differ. It may become necessary to choose among alternative raw materials or semi-finished goods, for example coal, oil or nuclear energy for the production of electricity. But in any case, the technological choice can always be traced to the problem of determining the pattern of the production inputs.

Decision problem 2 is thus the determination of the input structure of production.

3. The problem of selecting the production pattern and technology is closely related to the problem of forming the country's stock of capital, especially fixed capital. The real content of sector development planning is not the estimation of future production and foreign trade but the working out of the potential that is to delimit future economic activity.

The planning of the capital stock involves answering these questions:

(a) Did an initial stock of capital exist at the beginning of the plan period? What should its fate be? Should it be dismantled? Should it continue in operation without transformation? Or should it be transformed and modernized? If so, what technological requirements should the modernized equipment satisfy?

(b) What should the composition and size of any new capital stock be? On what type of technology should the plants it includes be based?

Problem 3, the clarification of the structure of capital stock is thus inseparably linked with decision problems 1 and 2 (production and technology).

4. In connexion with the capital stock it must be decided where the plants should be sited. This is a significant problem in every country. The problem becomes particularly important in those with a large territory; however, it may also be important in small countries in which considerable differences exist between regions. For example, there may be advanced and backward regions in the same country, coastal areas and territories situated far from the sea, regions inhabited by different cultural groups.

Decision problem 4 is thus the determination of the regional structure of production.

5. In the case of long-term planning, the problems of production are usually closely connected with those of external trade. On the one hand, there are certain essential import requirements, such as indispensable agricultural products that cannot be grown in the country or minerals that do not exist there. Their import does not constitute a decision problem and affects the plan only indirectly.

For some products, on the other hand, there is a competition between those produced domestically and imports. More frequently, domestic production has already reached a certain level but the question arises whether additional future demand should be met by increased domestic production, imports, or a combination of the two.

The extent to which the development plans will be affected by competitive imports depends on the characteristics of the sector; on its initial position, for example. In the case of some developing countries the question may arise in the form, Should a new sector be brought into existence at all? For example, Hungary has so far not taken up the production of passenger cars, although modern motor cycles, tractors, dumpers and motor buses are produced there. In other cases, partial decisions must be made in some sectors; for example, with an existing domestic production of machine tools it must be decided which types to manufacture domestically and which types to buy abroad. Or, if a country has no iron ore deposits, at which stage should a domestic ferrous metals industry be started? With the production of crude iron, importing only the iron ore? At the steel production stage, with imported crude iron? Or at the rolling stage, with imported steel?

To sum up: decision problem 5 is the determination of the import structure.

6. Whether competitive or not, imports must be paid for. Imports and exports together will enable the exploitation of the comparative advantages to be derived from the international division of labour and from external trade. It must, therefore, be decided to what extent the capacities to be set up within the sector should be used to meet domestic requirements and to what extent for export. The estimation of a realistic and desirable volume of exports is inseparable from the determination of the level of production.

Decision problem 6 is thus the determination of the export structure.

7. Whereas the question Where? has been raised in connexion with domestic production, Where from? and Where to? must be posed in connexion with external trade. In the case of both imports and exports, the markets must be determined. Besides economic and business considerations, the viewpoints of foreign policy may also play a role.

Decision problem 7 is thus the determination of the market structure.

8. In connexion with all seven decision problems listed above, a further question—When? —must be raised. How should the creation of the various components of the capital stock be scheduled? According to what schedule should the various production and foreign-trade activities be linked with this? What should be realized

earlier and what later? It is not only the final stage that should be planned but also the path leading up to it.

As can be seen, problems 4 and 7 introduce the dimension of space into the ensemble of decision problems, while problem 8 introduces that of time.

Decision problem 8 is thus the determination of the temporal structure of the economic activities.

B. *Delimitations and simplifying assumptions*

Having outlined the questions we wish to answer, we must now consider the problems which we cannot or do not wish to solve and the simplifying assumptions and abstractions on which the model is based. The type of model to be described is based on the following main assumptions.

Structural model

All eight decision problems discussed above are so-called structural decisions. Their effects can be observed in the real processes of the economy—in production, productive input, consumption, investment, export and import. The equations of the model, which will be dealt with later in detail, describe structural relationships, mainly technical and natural connexions between the economic activities.

Our model will, accordingly, not answer questions such as what stimulants the Government should employ to promote or to curb this or that economic activity; what wages could best provide an incentive to workers; what subsidies could help to bring about the suitable pattern of exports; what rates of duty would influence imports. The model will not give any advice regarding such problems as the selection of institutional economic frameworks and the establishment of expedient forms of information flow within the economy.

It is extremely important to model the regularities in behaviour of those participating in economic life and the interrelations between them, but this can hardly be done without adequate observation and the study of long time series, which is particularly difficult in rapidly changing countries. The type of model described below cannot begin to do it. It is not a behavioural but a structural model.

Exogenous parameters

Our model isolates the sector from the country's economic organism in an artificial and deliberate manner. There is no sector that is not linked with a thousand threads to other sectors. These threads are described by what are called "exogenous parameters". When modelling the sector of electrical energy production, for example, we will have to consider as exogenous parameters, on the one hand, the bounds of available inputs (e.g. the available quantity of domestic coal) and, on the other hand, the electric power requirements of the other sectors.

The exogenous parameters may be treated in various ways, according to the purpose of the computation. In the present description, in order to facilitate the treatment of the subject, the problem will be approached in several stages:

(a) In section I (this section), it will be simply assumed that the exogenous parameters are given constants;

(b) In section II, this assumption will be weakened. We shall examine what happens to the parameters connecting the sectors when the computations of the various sectors are linked up and united in a single model. But the exogenous parameters connecting sectors within the model with those outside will still be treated as constant. For example, let manufacturing be in the collectively treated group of sectors. The electric power industry is also in this group. In this case, the energy consumption of manufacturing will cease to be an exogenous parameter; it is now an "endogenous variable". The power requirements of agriculture or of the population will still be considered exogenous parameters;

(c) Finally, in section III, our assumption will be weakened further. Methods that make it possible to treat the exogenous parameters as variables will be introduced.

The lack of uncertainty

The uncertainty of the data used in the computations will be disregarded. Accordingly, the risk involved in the decision will not be taken into account.

The possibility of removing this assumption will be dealt with in section IV.

Aggregation

It was mentioned earlier that it will not be necessary to plan the pattern of output in detail; it will be sufficient to break it down to the extent to which it really affects investment decisions. This, however, is only one form of the more generally valid simplification employed in the construction of the model, namely aggregation. It is not only that ten thousand products are aggregated into a product group; an individual product may often be produced in a hundred ways, of which only some of the most characteristic will be taken into account. In the case of regional division, every country, village and street could be taken into account separately; instead, only a few main regions will be considered. Trade relations may be entered into with a hundred countries, but instead of treating these individually, we consider only a few principal trading regions. Similarly, with time: when planning for a period of 20 years, a schedule could be prepared in terms of years, months or even days. Instead, we shall aggregate in time as well, taking four periods of five years each.

In what follows, we shall treat such concepts as "product", "technology", "region", "market", "period" as aggregates.

We shall return to consider the problem of aggregation in more detail in section IV.

C. *The variables*

The next step is to construct a model suited to solving the decision problems posed in section I.A, taking into consideration the simplifying assumptions outlined in section I.B.

The model will contain a great number of variables, whose values are unknown; it is the purpose of the calculation to find these values.

The basic types of variable will be distinguished by various symbols; to the basic symbols various subscripts will be added. These will now be defined.

The model will cover a plan period divided into T shorter periods. Let the plan period be, say, 20 years, and let $T = 4$; then the computations will cover four five-year periods. The periods will be labelled by the subscript $t = 1, \ldots, T$.

The R regions of the country will be labelled by the subscript $r = 1, \ldots, R$.

The sector will produce I products. For example, in ferrous metallurgy product 1 will be crude iron, product 2 steel, product 3 thin plate and so forth. The product will be labelled by the subscript $i = 1, \ldots, I$.

The domestic production of product i can be carried out by J_i production methods. These differ in *(a)* the source of the capital stock with which production is carried out and *(b)* the type of technology employed in the production.

Item *(a)* refers to investment, maintenance and creation of capital stock during the period, whereas *(b)* refers to production and operation. The two aspects are obviously closely connected. When drawing up the list of variables and when defining and delimiting the production methods taken into account and distinguished from one another in the model, the essential alternatives of sector development must be carefully and logically examined. In any case, the J_i alternatives must include all distinguishable possibilities of investment and technical development that may come into consideration in connexion with the production of product i.

The production method will be labelled by the subscript $j = 1, \ldots, J_i$.

In the country's foreign trade, a total of M markets may come into consideration. The market will be labelled by the subscript $m = 1, \ldots, M$.

In the model there are three main types of variable, namely:

- v_{ijrt} – the production of product i by production method j in region r in the last year of period t
- y_{imt} – the import of product i from market m in the last year of period t
- z_{imt} – the export of product i to market m in the last year of period t

Of these, only the first will require detailed explanation. As will be clear from the definition, it represents primarily production activity carried out in the last year of the period. However, on the basis of what has been said above when defining "production method", it will be clear that it represents also investment activity: the investment actions that must be carried out during the whole period in order to permit production to reach the volume v_{ijrt} in the last year.

Let us define production method 1 as the continued operation of the old capacity already in existence at the beginning of the period, with the original technology A; production method 2 as the technical reconstruction of the old capacity already in existence at the beginning of the period, in such a way that a switch from technology A to technology B is possible; production method 3 as a reconstruction also, but one which permits changing from technology A to technology C; production method 4 as creation of new plant during the period, the new plant operating with technology D; and production method 5 as the creation of a new plant that operates with technology E in the last year.

By fixing the value of the variables v_{ijrt} not only will production in the last year be determined but also investment activity during the period. Variable v_{i1rt} will require maintenance inputs only, variables v_{i2rt} and v_{i3rt}, inputs of maintenance and reconstruction and variables v_{i4rt} and v_{i5rt}, inputs of maintenance and new plant construction.

The example throws light on one of the proposed model's principal "tricks". The variables v_{ijrt} represent both production in the last year and investment over the whole period (interpreting the latter in the broadest sense as gross investment, which includes also replacement and maintenance).

In accordance with the assumption of aggregation, the model does not provide an estimate of the investment schedule within the period. It is only the capacities to be set up by the end of the period that are planned by means of the variables v_{ijrt}.

Let us introduce for the purposes of later discussion the more concise notation:

- $\mathbf{v}$ – the vector composed of all production variables of the sector model
- $\mathbf{y}$ – the vector composed of all import variables of the sector model
- $\mathbf{z}$ – the vector composed of all export variables of the sector model

Let the vector $\mathbf{x}$ denote in summary form the sector's development programme:

$$\mathbf{x} = [\mathbf{v}, \mathbf{y}, \mathbf{z}] \tag{1}$$

The breakdown of the production variables by four subscript types and of the foreign-trade variables by three subscript types ensures that computation of the programme shall give answers to all the decision problems listed in section I.A.

D. *The constraints*

The variables of the model are bounded by certain constraints, which in the present model are expressed as linear equalities or inequalities.

Non-negativity

The first constraint is that the variables must not be negative:

$$x \geq 0 \tag{2}$$

where x stands for any one of the variables introduced in section I.C. The model is constructed in such a way that economic activities of a possibly destructive character are represented by a non-negative value of a separate variable. Should we, for example, wish to take into consideration the direct consequences of the dismantling of an old plant, we simply build into the model a separate "dismantling variable" representing the scrap value of the old equipment.

The constraint in equation (2) applies to all variables. The constraints set forth below each apply to only a certain group of variables.

The balances of output products

The general form of the output product balance for product i in the last year of period t is

$$\sum_{j=1}^{J_i} \sum_{r=1}^{R} v_{ijrt} - \sum_{h=1}^{I} \sum_{j=1}^{J_h} \sum_{r=1}^{R} g_{ihjrt} v_{hjrt} + \sum_{m=1}^{M} y_{imt} - \sum_{m=1}^{M} z_{imt} \geq V_{it} \qquad \begin{matrix} i=1, \ldots, I \\ t=1, \ldots, T \end{matrix} \tag{3}$$

The first term is the total domestic production of i within the sector. The second is the productive input of i within the sector, the technological coefficient g_{ihjrt} representing the input of i per unit of h. The third is total import, and the fourth is total export. Equation (3) says that the external demand for product i outside the sector V_{it} must be at least equal to the amounts of i produced in and imported into the sector less the amounts exported out of and used within the sector to produce other products. (For a better understanding of its economic content, equation (3) has been written in a form such that the first term contains the output of the variables producing product i and the second the "own input" of the same variables for the subscripts $h = i$. When carrying out the programming, the coefficient $(1 - g_{iijrt})$ will, of course, stand at the corresponding place in the matrix of coefficients.)

According to the second assumption in section I.B, V_{it} it a constant exogenous parameter.

There may be several special types of output product balance. A few examples:

(a) Output product balances may be given not only for the country's entire territory but also separately for individual regions by placing a constraint on the input-output flows between the regions;

(b) In the electric power industry, it is not sufficient to stipulate annual total output, because requirements will vary seasonally;

(c) In the chemical industry, it is necessary to prescribe special technological proportionalities (for example mixing proportions) in the form of product balances;

(d) In the case of products not utilized within the sector, equation (3) will be simpler, the second term falling out.

The balances of input products

Certain products turned out by other sectors, called external materials, will be available to the sector only in limited quantities. For example, domestic production of coal and petroleum will be limited by the store of mineral resources, and therefore the part of the output of the coal and petroleum sectors that is available for electric power production is also limited. Various bottle-necks in the economy will also limit the production of certain sectors. In most developing countries bottle-necks exist in railway transport and in the building industry.

The general form of the input product balance valid for external material h in the last year of period t is

$$\sum_{i=1}^{I} \sum_{j=1}^{J_i} \sum_{r=1}^{R} g_{hijrt} v_{ijrt} \leqq G_{ht} \qquad \begin{matrix} h=1, \ldots, H \\ i=1, \ldots, I \end{matrix} \tag{4}$$

Here again, g_{hijrt} is a technological coefficient: the input of h necessary for one unit of i. The constraint G_{ht} on the right-hand side is a constant and represents the external material quota, which is considered an exogenous parameter. This is usually the full quota destined for the sector, irrespective of whether h has been produced domestically or imported. However, it is conceivable that in certain sector models it would be more expedient to establish quotas for external materials of domestic origin only and to represent the import of h by introducing import variables y_{hmt}.

The constraints of labour

The general form of the balance for the last year of the t-th period is the following:

$$\sum_{i=1}^{I} \sum_{j=1}^{J_i} \sum_{r=1}^{R} l_{ijrt} v_{ijrt} \leqq L_t \qquad t=1, \ldots, T \tag{5}$$

Constraint (5) puts a bound on the total labour requirements of the sector. The coefficient l_{ijrt} gives the labour requirements per unit of product and L_t is a constant representing the supply of labour.

Instead of (or in addition to) constraint (5), some other constraints relating to live labour may be given:

(a) Instead of (or in addition to) the upper bound, a lower bound may also be set. The upper bound will prevent the insertion into the programme of excessive labour requirements that cannot be satisfied. The latter will ensure the desired level of employment;

(b) Instead of a global, aggregate labour constraint, more detailed constraints may be applied. The breakdown may be by region, sex, qualification etc., according to the conditions in the country;

(c) It is not only the size of the labour force but also the amount of wages that may be bounded. This may play a role in securing the proper proportionality between commodity funds and purchasing power in the country.

The constraints of natural resources

The production of some sectors is limited because the available amount of natural resources necessary for their operation is limited. This is the case with land in agriculture, ore in the extractive industries, water power in the production of electricity etc.

Let $\bar{H}$ be the number of natural resources essential from the point of view of the sector. The general form of the constraint for natural resource h in the last year of period t, will be:

$$\sum_{i=1}^{I} \sum_{j=1}^{J_i} \sum_{r=1}^{R} n_{hijrt} v_{ijrt} \leqq N_{ht} \qquad \begin{matrix} h=1, \ldots, \bar{H} \\ t=1, \ldots, T \end{matrix} \tag{6}$$

In equation (6), n_{hijrt} denotes natural resource requirements per unit of product and N_{ht} the total available natural resources.

Capacity constraints

From the viewpoint of fixed capital, the constraints may be divided into two categories. To the first belong the constraints relating to the capital stock already in existence at the beginning of the period. These will be described under this heading. The second category comprises the constraints which limit capital formation during the period and will be dealt with next, under the heading "Investment quotas".

We describe some special constraints first and discuss the general form of the capacity constraints later. Let us consider production method $j = 1$, which utilizes existing capacity with unchanged technology A (section I.C), during period t = 1. In this case, the following capacity constraint may be given for the production of product i in region r:

$$v_{i1r1} \leqq Q_{i1r1} \qquad \begin{matrix} i=1, \ldots, I \\ r=1, \ldots, R \end{matrix} \tag{7}$$

Here, Q_{i1r1} denotes a capacity already in existence at the beginning of the period; the constraint is that production carried out with unchanged technology in old plants in the last year of the period cannot exceed this capacity.

We also have the variables

v_{i2r1} –Production with old capacity, but following a change-over from technology A to technology B

v_{i3r1} –production with old capacity, but with a change-over from technology A to technology C

These change-overs not only affect the input pattern but also increase the possible output. We therefore have analogous constraints for $j = 2, 3$:

$$\begin{aligned} v_{i2r1} &\leqq Q_{i2r1} \\ v_{i3r1} &\leqq Q_{i3r1} \end{aligned} \tag{8}$$

where $Q_{i2r1}, Q_{i3r1} \geqq Q_{i1r1}$.

However, in the model it must be ensured that the old capacity should not all be engaged first in production method 1, then in production method 2, and then once more in production method 3. A collective capacity constraint must, therefore, be applied to the three production methods:

$$q_{i1r1} v_{i1r1} + q_{i2r1} v_{i2r1} + q_{i3r1} v_{i3r1} \leq Q_{i1r1} \tag{9}$$

where

$$q_{ijr1} \equiv \frac{Q_{i1r1}}{Q_{ijr1}} \qquad j=1, 2, 3 \tag{10}$$

are the coefficients of capacity engagement. If the reconstruction is accompanied by capacity extension, then q_{ijr1} is smaller than unity; if not, it equals unity. (Note that $q_{i1rt} = 1$.)

So far, we have only been speaking of period 1, in which production is limited by "inheritance", that is, by the situation which prevailed before the plan period. In period 2, however, the real constraint depends on the investment activity in period 1.

Let us again take the most simple case, that of production method 1. The formula analogous with constraint (8), now in general terms for period $t + 1$, is

$$v_{i1r,t+1} \leq v_{i1rt} \qquad \begin{cases} i=1, \ldots, I \\ r=1, \ldots, R \\ t=1, \ldots, T \end{cases} \tag{11}$$

Let us assume that in period 1 the old plant was not kept in operation in unchanged form, but dismantled; then we have $v_{i1r1} = 0$. Constraint (11) then gives $v_{i1r2} = 0$, that is, production based on method 1 cannot be carried out in period 2 either.

Let us now leave the special cases and proceed to the general form of the capacity constraint:

$$\sum_{j \in U_{irf}} q_{ijr,t+1} v_{ijr,t+1} \leq \sum_{j \in W_{irf}} v_{ijrt} \qquad \begin{cases} i=1, \ldots, I \\ r=1, \ldots, R \\ t=0, 1, \ldots, T \\ f=1, \ldots, F \\ v_{ijr0} = Q_{ijr1} \end{cases} \tag{12}$$

The summation on the left-hand side is over all the production methods that are to be utilized by the end of a given period (capacity needs), on the right-hand side over those existing at the end of the preceding period (capacity supply). The set of subscripts labelling the methods in the former group is symbolized by U_{irf}, in the latter by W_{irf}. The new subscript $f = 1, \ldots, F$ allows for establishing several constraints, each involving different sets of production methods.

As can be seen, the constraint is of intertemporal character. It establishes a connexion between one group of production variables belonging to period $t + 1$ and another belonging to the preceding period t. The constraint can be loosely formulated verbally as follows: The demand of the variable group for old capacities in the last year (left-hand side) cannot exceed the old capacity set up by the end of the preceding period (right-hand side).

The special case of constraint (11) is obtained from (12) by setting

$$U_{irf} = \{1\}, \quad W_{irf} = \{1\} \tag{13}$$

and (9) is one of the special cases obtained when $t = 0$.

At first, this formal description of the intertemporal capacity constraints may seem rather intricate. Actually, only the definitions of the inevitable symbols are; the relationships are themselves mathematically simple, and the economic content is easily grasped. Once the definitions are understood, there should be no difficulty in setting up the intertemporal capacity constraints.

Investment quotas

The development of the sector requires investment. One of the characteristics of the model described here is that it takes into account so-called gross investment, which covers, among other things:

(a) The maintenance and replacement costs of old fixed capital, in the case of fixed capital already in existence at the beginning of the period slated for conservation and further operation;

(b) The cost of the technical reconstruction of old fixed capital, in the case of fixed capital already in existence at the beginning of the period and slated for technological transformation and reconstruction;

(c) The cost of new capital investment.

The general form of the investment quota constraint for period t is

$$\sum_{i=1}^{I} \sum_{j=1}^{J_i} \sum_{r=1}^{R} k_{ijrt}\, v_{ijrt} \leq K_t \qquad t=1, \ldots, T \tag{14}$$

Coefficient k_{ijrt} is the gross investment requirement per unit of product. The quantity K_t on the right-hand side, which represents the total gross investment quota for the sector, is an exogenous parameter treated as a constant.

The constraints bounding investment resources in general and constraint (14) in particular are the only type used in the model that limit the investment activities carried out during the whole period instead of co-ordinating inputs and outputs in the last year of the period.

Constraint (14) limits the total gross investment of the sector. In addition to this global quota, partial ones may also be fixed. Their breakdown should correspond to the special conditions of the country and the sector. Some possible breakdowns are:

(a) According to the material composition of the investment. For example, special bounds may be set to the construction part of investment activity, or to capital formation embodied in equipment and machinery;

(b) According to the regional distribution of investment. Separate lower and upper bounds may be set to investment activity carried out in the individual regions;

(c) According to the financial sources of investment. Thus, separate quotas may be established for investment projects that the enterprises are financing by themselves or by bank loans, or for those realized from government funds;

(d) According to the nature of foreign credits. There may be foreign creditors who would grant credit only for some definite purposes. Should the drawing on foreign credit also figure in the model, a limitation of the amount of these credits by separate credit constraints may be justifiable. In such cases, the credit constraint constitutes an exogenous parameter, the magnitude of which may be affected by the points of view of foreign policy.

Foreign-trade constraints

The general forms of the foreign-trade constraints are

$$y_{imt} \leq Y_{imt} \qquad \left\{ \begin{array}{l} i=1, \ldots, I \\ m=1, \ldots, M \\ t=1, \ldots, T \end{array} \right. \tag{15}$$

$$z_{imt} \leq Z_{imt} \tag{16}$$

Upper bounds by product, market and period are thus set to both import and export activities. It is, of course, neither necessary nor whorth while to set the constraints for every product and every market in exactly the above form.

A constraint like (15) will be required when the foreign exporter is not prepared to sell unlimited quantities to the sector of the model. This constraint will not be needed if the scarcity is due to a lack of foreign currency rather than product; the constraints on currency will be determined by the balances of external payments (see below).

The export constraint (16), on the other hand, is almost always necessary, because there is hardly any product that can be sold in foreign markets in unlimited quantities.

In certain sector models a deviation from the form of foreign-trade constraint shown in (15) and (16) may be warranted.

It may be necessary to set lower bounds in addition to (or instead of) the upper bounds.[1] This will be the case when under prior international or interenterprise agreements, the country is obliged to carry out certain foreign transactions.

Sometimes combination may become necessary. For example, constraints may be set, not for each product, but collectively for several products, allowing mutual substitution. Or, instead of setting a constraint for each market, a constraint may be set for several markets collectively.

The balances of external payments

The income from exports to market m and the outgo for imports from the same market in the last year of period t are bounded by the following balance of payments constraint:

$$\sum_{i=1}^{I} (p_{imt}\, z_{imt} - \bar{p}_{imt}\, \bar{y}_{imt}) - \sum_{i=1}^{I} \sum_{j=1}^{J_i} \sum_{r=1}^{R} d_{mijrt}\, v_{ijrt} \geqq D_{mt} \qquad m=1, \ldots, M \tag{17}$$

[1] Actually, lower bounds of imports and exports do not usually require a separate constraint in the model; they may be deducted from or added to the constraint constant on the right-hand side of the product balance.

The coefficients p_{imt} and $\bar{p}_{imt}$ in the first sum denote the unit price of exports and imports, respectively. The coefficient d_{mijrt} in the second sum is the non-competitive import costs per unit product. It covers all outlays for imports that are not represented by the competitive import variables y_{imt} in the model.

If the total income of the sector from market m are positive in sign, the constraint constant D_{mt} will represent the net balance expected from the sector, and the constraint will be set in the form of a lower bound. If the sector shows a loss of foreign exchange, the upper bound of the loss will be set as a constraint and the signs will change accordingly.

Several markets may be combined. For each of these markets separate foreign-trade constraints like (15) and (16) will be prescribed, but their balances of payments will be treated collectively. This is possible only if the weights (exchange rates) necessary for the summation are available.

As has already been mentioned, various credit alternatives may be included among the variables. The effect of these variables on the balance of payments must be taken into account in (17); the coefficients must express the annual cost of capital, that is, the principal and interest payments due in the last year.

The constraints of smooth development

The constraints of smooth development serve a double purpose. On the one hand, they express the degree of "incapacity" of the economy: whatever the wishes of those making the decisions, there is an upper limit to the rate of progress that can be maintained. On the other hand, these constraints serve to slow progress; development that occurs with either great set-backs or exaggerated leaps is generally undesirable.

To give an example, let us describe the general form of the smoothness constraint prescribed for production of product i in period t:

$$\left.\begin{aligned} \sum_{j=1}^{J_i}\sum_{r=1}^{R} \alpha_i v_{ijrt} &\geq \sum_{j=1}^{J_i}\sum_{r=1}^{R} v_{ijr,t-1} \\ \sum_{j=1}^{J_i}\sum_{r=1}^{R} \beta_i v_{ijrt} &\leq \sum_{j=1}^{J_i}\sum_{r=1}^{R} v_{ijr,t-1} \end{aligned}\right\} \quad \begin{array}{l} i=1,\ldots,I \\ t=2,\ldots,T \end{array} \tag{18}$$

where coefficient α_i in the first equation is a number not much smaller than unity, and coefficient β_i in the second is a number not much greater than unity.

The constraints of smoothness may differ in form from (18) in several ways:

(a) For the first period, the member on the right-hand side of the two inequalities in (18) is not variable but constant; it is the production level "inherited" from the preceding plan period;

(b) Constraints of smoothness may be imposed not only on production but also on export and import activities;

(c) Constraint of smoothness may be set collectively for a group of products instead of a single product;

(d) It is possible to set only one bound, either lower or upper.

E. *The objective function*

Various optimization criteria can be prescribed for the sector model. It is not our aim to discuss at this point what the economic content of the optimization criterion should be; some aspects of this problem will be dealt with later. A few examples are given here for the sake of illustration only.

Maximization of total external output. The aim is to increase the additional output over that needed to satisfy the external demand prescribed as a lower bound in the output product balance, equation (3). This can be achieved in two ways:

If the product pattern of the additional external production is fixed in advance a single variable u representing the additional external output is introduced; this variable is then maximized in the objective function. In equation (3) a negative term

$$-\delta_i u \tag{19}$$

is added, where the coefficient δ_i gives the contribution of i to one unit of total additional external output.

Separate additional external consumption variables may be assigned to each product. The model permits these to be freely chosen; it does not fix their proportion in advance but assigns preference weights to them. In a single-period model, the objective function thus has the following form:

$$\sum_{i=1}^{I} \pi_i u_i \rightarrow \max! \tag{20}$$

where π_i is the preference weight of product i and u_i the variable representing its additional external output. The ratio π_i/π_j expresses the socially desirable rate of substitution of product i for product j.

Maximization of employment. Conversely, in case a small supply of labour constitutes a bottle-neck in the economy, the criterion could be the maximum saving of manpower.

Maximization of some positive foreign-exchange balance. Alternatively, the aim could be the minimization of foreign-exchange loss or the collective optimization of several foreign-exchange balances, assigning definite weights (exchange rates) to the various currencies. This criterion will be particularly interesting to countries with foreign-trade difficulties.

Minimization of total costs. This could be the criterion when a sector has a prescribed exogenous output obligation.

Maximization of total productive capacity. In this case we are actually maximizing the final capital stock, but for the sake of simplicity this can be measured by the sum total of all v_{ijrt} production variables. We may also assign capital coefficients to the production variables v_{ijrt} and thus maximize the final capital stock directly.

Whichever criterion is selected, dynamical problems will arise.

Let the five types of economic objective listed above (or similar ones) be called "returns". The returns of period t are a linear function of the programme vector $\mathbf{x}$:

$$C_t = c_t\, x_t \tag{21}$$

where x_t is a component of the programme vector relating to period t, and c_t is a coefficient with the dimensions of returns per unit of x_t.

In mathematical terms, optimization of the plan for period t merely means maximizing the objective function (21). The question remains whether we should maximize the returns for every t or for a specific one, say T. In the former case, we may wish to maximize the simple or weighted sum of the objective functions. We therefore have for the most general form of the objective function

$$C = \sum_{t=1}^{T} \gamma_t\, c_t\, x_t \rightarrow \max! \tag{22}$$

In this equation the weights γ_t express the time preferences of the subperiods. The quotient γ_{t-1}/γ_t is usually greater than unity and expresses the extent to which returns achieved in period $t-1$ are valued more highly than the returns achieved in period t.

F. *Summary*

The formulation is now complete and can be stated as a straightforward linear programming problem:

$$\begin{aligned} \mathsf{A}\,\mathbf{x} &= \mathbf{b} \\ \mathbf{x} &\geqq 0 \\ \mathbf{c}'\,\mathbf{x} &\rightarrow \max! \end{aligned} \tag{23}$$

The programme vector $\mathbf{x}$ has already been defined in equation (1).

The coefficient matrix A contains all the coefficients that appear in the constraints described in equations (3) to (18).

The constraint vector $\mathbf{b}$ is a column vector containing the constraint constants appearing in equations (3) to (18).[2]

Finally, the vector of objective-function coefficients $\mathbf{c}$ contains the objective-function coefficients $c_1, \ldots, c_T$ of all periods.[3]

[2] In the discussion up to now, the constraints have generally been written down in the form of inequalities. However, with the introduction of suitable auxiliary variables, these can be transformed into equations, as has been done in (23).

[3] In this article a prime after a symbol is a sign of transposition; here it indicates the transposition of the column vector c into a row vector.

II. Linking the sector models

A. *Intersectoral two-level structure*

Throughout section I, we were dealing with the construction of the model of a single sector. Let us now proceed to the problem of constructing the development model of several sectors taken together. The question how great a part of the economy should be embraced will be left open: all sectors of the economy; only one major sector, for example manufacturing; or only one industry of that sector, for example the chemical industry. It suffices to state that we have a total of S sectors and label them with the subscript $s = 1, \ldots, S$.

For every sector there is a sector model similar to the one described in section I and summarized in equation (23). Our task now is to unite the sector models.

The result will be a giant linear programming problem of this general form:

$$\mathsf{A}_1^{(1)}\mathbf{x}_1 + \mathsf{A}_2^{(1)}\mathbf{x}_2 + \ldots + \mathsf{A}_S^{(1)}\mathbf{x}_S = \mathbf{b}^{(1)} \tag{24}$$

$$\mathsf{A}_1^{(2)}\mathbf{x}_1 = \mathbf{b}_1^{(2)} \tag{25.1}$$

$$\mathsf{A}_2^{(2)}\mathbf{x}_2 = \mathbf{b}_2^{(2)} \tag{25.2}$$

$$\ldots \qquad \vdots$$

$$\mathsf{A}_S^{(2)}\mathbf{x}_S = \mathbf{b}_S^{(2)} \tag{25.S}$$

$$\mathbf{x}_1 \geq 0, \mathbf{x}_2 \geq 0, \ldots, \mathbf{x}_S \geq 0 \tag{26}$$

$$\mathbf{c}_1'\,\mathbf{x}_1 + \mathbf{c}_2'\,\mathbf{x}_2 + \ldots + \mathbf{c}_S'\,\mathbf{x}_S \rightarrow \max! \tag{27}$$

The variables $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_S$ are the programme vectors of the first, second, ..., S-th sector, respectively. Equation (24), the first row of the system of constraints, contains the so-called central constraints (or, to use another term, the intersectoral constraints). Accordingly, the matrix $\mathsf{A}_s^{(1)}$, $s = 1, \ldots, S$, is the matrix of the coefficients of sector s that figure in the central constraints. On the right-hand side, $\mathbf{b}^{(1)}$ is the vector of the central bounds.

The subsequent rows of the system of constraints, (25.1), (25.2), ..., (25.S), contain the so-called sector constraints. Accordingly, the matrix $\mathsf{A}_s^{(2)}$, $s = 1, \ldots S$, is composed of the coefficients of sector s figuring in the sector constraints. On the right-hand side of each of these rows, each sector has a vector $\mathbf{b}_s^{(2)}$, the vector of the bounds of its own sector constraints.

The equations in (26) represent the conditions of non-negativity, and equation (27) is the objective function. The latter constitutes the sum of the objective functions of the sectors.

Figure I shows equations (24) and (25) written in matrix form. The horizontal strip on the top contains the central constraints; below it, in diagonally arranged blocks, are the sector constraints. The other blocks of the combined matrix of coefficients are null matrices. The explanation is that the sector constraints for sector s can connect only the variables belonging to that sector. The result is that each of the sector constraint rows

Central constraints

$$\begin{bmatrix} A_1^{(1)} & A_2^{(1)} & \cdots & A_S^{(1)} \end{bmatrix} \begin{bmatrix} x \end{bmatrix} = \begin{bmatrix} b^{(1)} \end{bmatrix}$$

Sector constraints

$$\begin{bmatrix} A_1^{(2)} & 0 & \cdots & 0 \\ 0 & A_2^{(2)} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & A_S^{(2)} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_S \end{bmatrix} = \begin{bmatrix} b_1^{(2)} \\ b_2^{(2)} \\ \vdots \\ b_S^{(2)} \end{bmatrix}$$

Figure I. Matrix equations representing the central and sector constraints

contains only a single non-zero block; the rest of the row contains only zeros. Every constraint connecting at least two sectors must be considered a central constraint. Some examples of these two types of constraint follow.

Central constraints, intersectoral relationships

Balances of products produced by more than one sector, or balances in which both the producing sector and one or more of the user sectors are included in the model, for example the sector of electric power production, which is included in the model together with several sectors that use electric power. In this case, a number of product balances that in the separate sector models figured as input product balances of the type (4), bounding the input of an external material, will now become a combined input-output balance. In the sector model of the chemical industry (if this sector was separately programmed), electric power figured as an external material, the input of which was bounded, accordingly, by an input product balance of the type that appears in equation (4). In the combined model, however, the electric power output of the electric power sector and the electric power input of the chemical sector (and of others) will appear in the same constraint

Input product balances of materials that are given from outside as external materials even for the combined model and are used by more than one sector as input. For example, rolled steel, one of the products of the excluded ferrous metallurgy sector, is used in several included sectors (engineering, construction). In such cases, the rolled steel is accounted for as an external material and is allocated to the users by a central constraint of type (4)

Manpower balances that allocate workers generally rather than to individual sectors

Natural resources utilized by several sectors, for example land, if the model covers several agricultural sectors

Investment quotas

Balances of external payments

Smoothness constraints, provided that they affect the production and foreign trade of several sectors at the same time

Sector constraints, intrasectoral relationships

Output product balances of products destined exclusively for external output and not used by any of the model's sectors

Input product balances of external materials used exclusively by one sector

Qualified labour needed for the special requirements of a sector, for example the allocation of miners in the mining sector

Natural resources exploited by only one sector, for example, oil still underground constitutes a bound only for the oil-producing sector

Capacity constraints

Foreign-trade constraints bounding the export and import of the individual products

The model structure described by equations (24)-(27) and figure I is called an intersectoral two-level structure.

If the system of constraints has this structure, the linear programming problem can be solved by one of the many so-called decomposition methods. We shall not describe any of these in detail but merely outline some of their common features.

Instead of treating the problem as a single large linear programming problem, we decompose it into numerous small problems. One of these, the one relating to the intersectoral constraints, will be solved as a central model. The other, relating to the intrasectoral constraints, will be solved as S sector models. Both the central model and the individual sector models are smaller (usually much smaller) than the large combined problem. It is exactly here, in the smaller size of the models, where the advantage of the decomposition method lies.

However, there is a price to be paid for this advantage: the decomposition methods usually involve a lengthy iterative procedure. Computations are carried out alternately, now with the central model, now with the sector models. As a result of the computation carried out with the central model, definite information is obtained that is utilized in the computations carried out with the sector models. Then, from the computations carried out with the sector models, definite information is obtained that is utilized in the next computation with the central model. In this procedure, the optimum solution to the original undecomposed problem is either reached in a finite number of steps or at least approximated with an arbitrarily chosen degree of accuracy.

Now it is clear why the structure described in equations (24)-(27) and figure I is called two-level. The function of planning has been divided between two levels: part of the work is done at the centre, part in the sectors. Between the two levels, there is a flow of information.

The reason for the other adjective, intersectoral, in the designation will now be clear also: the units of the lower level represent the sectors. The centre performs the planning of intersectoral allocation.

Having summed up the common characteristics of the various decomposition methods, let us now briefly mention their differences. They may differ from each other in the type of computation to be carried out at the central and sectoral levels: linear programming, some kind of ordering, calculation of averages or other type. They may differ in the type of information flowing between the two levels: the centre may set prices to the sectors or merely set lower or upper bounds. The procedure may be finite or infinite (but with monotonous or oscillating convergence).

B. *Other types of arrangement*

Interregional arrangement

Suppose that we wish to focus our interest on the problem, not of intersectoral, but of interregional, allocation. In this case, we give the structure described by equations (24)-(27) and figure I a different interpretation. In the case of decomposition, the "lower level" will be the region, the "upper level" the centre that co-ordinates the regions. The variables will be arranged in the individual column groups according to regions instead of sectors. The diagonally arranged small blocks in the lower-level coefficient matrix $A_1^{(2)}, A_2^{(2)}, \ldots, A_R^{(2)}$ now contain the coefficients of the special constraints of regions $1, 2, \ldots, R$, respectively. The blocks $A_1^{(1)}, A_2^{(1)}, \ldots, A_R^{(1)}$ will consist of the coefficients of the common interregional constraints.

Intertemporal arrangement

Another possibility is to focus on the intertemporal distribution of the activities. Then the variables must be arranged in the individual column groups according to periods. All intertemporal bounds will appear among the central constraints. These are, in the case of the constraint system described in section I, the capacity as well as the smoothness constraints. In the latter case, the same inequality applies to the activities of two subsequent periods. All other constraints appear only in the diagonally arranged lower-level blocks $A_1^{(2)}, A_2^{(2)}, \ldots, A_T^{(2)}$. These constraints refer always to the variables of one period only.

Multilevel arrangement

Figure II shows the coefficient matrix of a linear programming problem arranged in a three-level structure. No symbols have been introduced; the non-zero blocks in the matrix are indicated by different shadings.

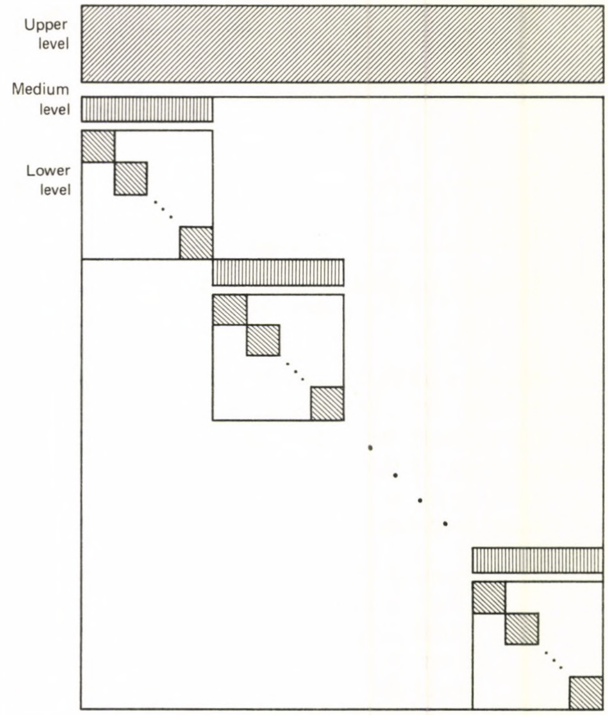

Figure II. Schematic arrangement of coefficient matrices for a three-level intersectoral model

Three-level arrangements can be made with various economic contents; for example, the three levels can represent various degrees of aggregation by products or sectors. On the lower level appear the various subsectors of, say, the chemical industry: the production of fertilizers, plastics, pharmaceuticals, dyestuffs etc. These are united into the main sector of chemical industry (the medium level). Similarly, the subsectors of the engineering industry appear on the lower level: automobile production, machine-tool production, telecommunication equipment production, etc. These are united into the main sector of engineering. Finally, the activities of the main sectors (chemical industry, engineering etc.) are co-ordinated on the upper level.

An arrangement is also conceivable in which the upper level contains the intertemporal relationships, the medium level the intersectoral relationships within the individual periods and the lower level the intrasectoral relationships within the same periods. The arrangements according to the regional and sectoral viewpoints can be similarly combined in a three-level structure.

Structures of more than three levels may also be worked out if required.

There are no general rules for the number of levels to employ, nor for their economic contents and classification. These will depend on the contents of the planning problem, on the decision problems to be answered and on the degrees of aggregation or disaggregation actually used. The structure of the institutions taking an active part in the planning process may also play a role. The answer to the question whether a multilevel model should be of sectoral or regional arrangement, for example, may depend on the institutional structure of the economy and the hierarchy of its administration.

In any case, whichever arrangement is employed, the multilevel model means that we really have, not a single model, but a system of separate submodels linked by information flows. Consider the simplest case, the structure described in section II.A. Here, the elements of the model system are the intersectoral central model and the intrasectoral sector models. It is in the network linking these that the planning process takes place.

III. The possibilities of utilizing the models

A. *Computation series*

It is not just one computation that will be carried out with a model, but a whole series of them. These questions arise:

What is the difference between the individual members of the computation series?

What value will be prescribed for the exogenous parameters? In other words, what numerical value will be given to the individual components of the vector of bounds? Should the output obligations be higher or lower, the investment more or less?

What optimization criterion should be prescribed?

What numerical value should be given to the coefficients figuring in the constraints and in the objective function?

It is possible to work consecutively with different bound vectors $\mathbf{b}^{[1]}, \mathbf{b}^{[2]}, \ldots$, vectors of objective-function coefficients $\mathbf{c}^{[1]}, \mathbf{c}^{[2]}, \ldots$, constraint coefficient matrices $A^{[1]}, A^{[2]}, \ldots$. (The superscripts in the square brackets refer to the number of the computation in the series of computations.) Besides this discrete method of variant calculation, continuous methods can also be employed. This is called parametric programming. For example, the vector of bounds may be written in the following form:

$$\mathbf{b} = \mathbf{b}' + \mathbf{b}''(\lambda) \qquad (28)$$

Here, $\mathbf{b}'$ is the constant part of the vector of bounds, $\mathbf{b}''$ the part depending linearly on parameter λ. In parametric programming the parameter λ is run through an interval $[\lambda, \bar{\lambda}]$ within which the optimum programme will be determined for all values of λ. It may be assumed, for example, that all components of the vector of bounds are constant and only the required level of employment can be varied between two limits. With the aid of parametric programming, one can see how the optimum programme changes as a function of the required employment level.

Parametric programming may be carried out for the objective function in a similar way. The vector of objective-function coefficients is written:

$$\mathbf{c} = \mathbf{c}' + \mathbf{c}''(\lambda) \qquad (29)$$

where $\mathbf{c}'$ is constant and $\mathbf{c}''$ depends linearly on λ. Let the objective function be the minimization of costs, for example. Let us parameterize one of the cost items, say the outlays for imports from some market, using the exchange rate of that market as a parameter. All other cost items figure in the constant member $\mathbf{c}'$.

Any inferences based on the model must be drawn not from a single computation but from the computation series as a whole. This can be done in various ways:

(a) The reactions of the programme to changes in the single factors can be shown, for example, how the structure of the sector changes with the total available investment quota. As a result of the computation series we thus arrive at aggregate functions.

(b) The stability or sensitivity of the programme to changes in the various factors (vector of bounds, objective function, coefficients) can be examined. To do so, we may, among other things, declare every member in the series of computations an experiment, and analyse the experiments with the usual methods of mathematical statistics. Thus, we calculate the averages characteristic of the set of experiments, the dispersion around the averages, the distribution characteristics of the set of data. The sensitivity tests may show which parts of the programme are comparatively stable, which are more sensitive, and what these are sensitive to.

(c) With the aid of the model, it is possible to form various marginal indicators. Indicators of this type will be yielded primarily by the so-called dual solution of the linear programming problem, the optimum system of shadow prices. To every constraint of the model there belongs a shadow price. The shadow price is the answer to the following question: How much would the value of the objective function change with a unit change in the magnitude of the constraint? If the objective function is the maximization of personal consumption and the constraint is the investment quota, the shadow price is the additional consumption made possible with a unit increase in the investment quota.

However, marginal indicators can be derived not only from the dual solution of the programming problem but also by means of parametric programming. For example, the optimization of some balance of payments is prescribed as the objective function. In parametric programming, we let the magnitude of the investment quota run through an interval. From the computation results, it will be easy to determine the improvement in the foreign-exchange balance due to a unit increase in the investment quota.

B. *The role of the models in planning*

The role that mathematical models play in economic planning is not so straightforward as it may seem at first sight. It is not just a matter of setting up a suitable system of equations, compiling the necessary data, feeding the problem into an electronic computer and waiting for it to produce a development plan ready for implementation without further thought. In reality, the role of the models is quite different, and the demands made on them must be more modest and more realistic at the same time. These are some of the ways in which the use of models affects planning:

(a) The construction of the mathematical model has a disciplinary effect on thinking over the problem. It forces the planners to analyse thoroughly the logical structure of the decision problem. What are the alternatives between which we must choose? What interdependence exists between the partial decisions?

(b) The construction of the model has a disciplinary effect on the compilation of data. Computations made outside the framework of a comprehensive mathematical model often cannot be compared. Different definitions, different types of classification, different measurement units may have been used for each of them. The working out of a model system that unifies the sectors, regions or institutions helps standardize the definitions, classifications and measurement units and thus facilitates the comparison of data.

(c) The various sectors and regions will try to assert their own viewpoints when working out recommendations and proposals. Some proposals may conflict with each other. It is one of the advantages of the comprehensive model systems that, even if they do not put an end to the conflict, they facilitate compromise. In fact, within the framework of the model system the various proposals have a common factual basis.

(d) With the aid of the mathematical model, not just one, but several alternative proposals may be worked out, from which the decision makers can choose according to their own considerations. The methodology of working out the variants has already been briefly outlined in the preceding section. With the aid of the computation series carried out with the model, not only can the alternative proposals be worked out, but the differences between the alternatives can also be computed at once. The programming model does not relieve the decision maker of his responsibility; however, it helps him by showing the consequences of all the possible decisions. For example, suppose he computes the programme twice, with the same constraints but first maximizing employment then optimizing the balance of external payments. He can then compare the two programmes and analyse their differences, either on the basis of aggregate macroeconomic functions (see *(a)* in section III.A), with different differential indicators (see *(b)* in section III.A) or item by item, going into the details of the programme.

(e) Through its system of constraints, the model helps make the variants self-consistent and clear them of contradictions, disproportionalities and disturbances of equilibrium.

(f) The model makes it possible to select efficient programmes. Let us compare, for example, three hypothetical programmes on the basis of two different criteria, increase in personal consumption and increase in the balance of foreign exchange:

	Consumption	*Foreign-exchange balance*
	(percentage increase)	
Programme 1	200	50
Programme 2	150	80
Programme 3	130	80

From the point of view of consumption, programme 1 is definitely more advantageous than programme 2; from that of the foreign-exchange balance, on the other hand, programme 2 is undoubtedly more advantageous than programme 1. Neither can be said to "dominate" the other. When comparing programmes 2 and 3, the situation is entirely different. The two programmes ensure the same balance of foreign exchange; however, from the point of view of consumption, programme 3 is definitely less advantageous than programme 2. Programme 2 "dominates" programme 3; it is not less advantageous than the latter from any point of view and is more advantageous from at least one point of view.

The understanding of this concept of domination leads to this definition of efficiency: a programme is "efficient" if it is not dominated by any other programme. The requirement of efficiency is more modest than that of optimality. With the aid of the mathematical model, whole series of efficient programmes can be produced, but only one can be optimum. On the other hand, whichever of the objective functions is employed, the programme that optimizes it must also be an efficient plan proposal.

IV. Some problems of model construction

The construction of a workable planning model is an intricate intellectual task. The conditions and decision problems of the field of investigation will, of course, set a limit to the phantasy of the model constructor. Nevertheless, there usually are numerous open questions. In fact, some requirements of model construction may be in contradiction with each other. The model constructor will have to find an acceptable compromise between the contradictory requirements. Some of his difficulties in this will be discussed below.

A. *Delimitation of exogenous and endogenous spheres*

The problem what should be "inside" and what "outside" exists even when the model is a partial one. Suppose it concerns the development plans of only the chemical industry. Where and how should the model be "cut off" from the other sectors of the economy—from agriculture, a user of fertilizers; from the textile industry, a user of synthetic fibres; and so forth? This cutting off is always arbitrary. The distorting effects of arbitrariness can be mitigated by the methods described in section III.A; however, they cannot be wholly eliminated.

It is particularly difficult to plan the structure of the economy as a whole. There will be certain spheres of operation in the economy that can be squeezed into the type of model described above only forcibly. In section I.B it was emphasized that the model is a structural one, that is, it is a model that describes the real processes in the economy—production, investment, consumption—and not the behaviour of the producers, investors and consumers. Any optimization of personal consumption would therefore be incompatible with the character of the model, since it would require the modelling of the consumers' reactions to the various impulses. The consumers' requirements should therefore be given from outside as an exogenous parameter.

The model may plan the allocation of the given wages fund to the various branches, on the basis of wage-cost coefficients set from outside. But the model cannot give an answer to the question what the wage proportions serving as a basis for the wage-cost coefficients should be. Actually, this would require the modelling of the reactions of labour to the various wage proportions, that is, the short- and long-term supply of labour as a function of the wage tariff. The description of these types of relationship is not included in the model; it ought, therefore, to be treated as an exogenous parameter.

This does not mean that no model of the behaviour of individuals in the economy can be constructed. Nor can these problems be considered as unimportant from the point of view of planning long-term industrial development. All we wish to emphasize is that this is not the purpose of the model described here and that the effects of these factors on the model must, therefore, be given from outside, as an exogeneous parameter.

B. *Other variable classifications*

In the models described in sections I and II the variables were classified according to:

Character of product (i)
Technology and production method (j)
Region (r)
Market (m)
Period (t)

In the course of actual application, we may deviate from this classification in two ways:

(a) Fewer classifications may be used. We may decide, for example, not to break down the economy into regions, foreign trade into markets, all activities into periods etc;

(b) More classifications may be used; the model may be further broken down, according to viewpoints not discussed until now. We give only one example. Production may be broken down according to the form of ownership characteristic of the producing plant. Ownership form 1 could be the small family plant; form 2, the large privately owned (usually corporate) plant; form 3, the co-operative plant; form 4, the municipal plant; form 5, the State-owned plant. An even more detailed breakdown that takes domestic and foreign ownership into account is conceivable.

When a new viewpoint of breakdown and arrangement is introduced into the model, all its consequences must be carried through; that is, different subscripts must be given to the new variables, further constraints must be built into the system of constraints and so forth.

The model constructor's difficulty is that, on the one hand, he would like to take into account the greatest possible number of viewpoints and, on the other, he wants to avoid overdimensioning the model. (The main disadvantage of an oversized model is that the organizational and technical difficulties of data compilation and of computation are very great.)

C. *Degree of aggregation*

The problem of aggregation must be logically separated from that of breakdown and arrangement. For example, with respect to the regional breakdown the first question is whether the regional aspect (subscript r) should be taken into account at all. If the answer is yes, the second question is how many regions the country should be divided into. Similarly, should technological choice (j) be taken into account at all? (It is neglected in many models.) If yes, how many of the almost infinite number of technological choices should be included? (That is, what should the value of J_i be?)

In summing up the assumptions of the model type analysed here, it was pointed out that each of the elements in the model (product, technology, region etc.) is necessarily an aggregate; the model does not reflect the most detailed specifications of reality (all products, all towns and villages etc.). However, the requirements of some model users tend to enforce a higher degree of disaggregation. Enterprises, for example, cannot use a plan model that contains only large macro-aggregates such as the growth rate of national income or total investment.

Some problems arising in connexion with the compilation of data will also compel us to work with more disaggregated models. The technological coefficients quantified on the basis of the engineers' estimates can usually be determined more easily in a detailed breakdown and in concrete, technically definable form relating to some actually existing plant than in the form of highly aggregated sectoral averages. The case is similar with the estimates of export and import prices.

There are factors, on the other hand, that speak in favour of more aggregated models. Some of those using the models, especially higher government organs, will usually demand highly aggregated estimates. Some of the data, particularly those based on statistical observation, are more easily available in aggregated form.

Again, it is easy to see that the model constructor is faced with requirements of a contradictory character. The solution to his problem lies in recognizing that both kinds of model will be needed. As far as possible, models with different degrees of aggregation should be united in an organic multilevel model system.

D. *Full-* versus *limited-range models*

Two contradictory requirements are usually placed on plan models. One is that the model should be full-range; it should embrace all activities of the sector or group of sectors concerned, its total output and external trade. There should be no residual inputs and outputs not represented by variables in the model. To enforce this requirement, it is necessary to describe fully all input-output flows in the system of product balances. As a statistical starting-point we may use for this purpose the so-called Leontief models, the input-output tables describing the intersectoral flow of products. Moreover, the inputs of all resources (natural resources, labour etc.) must be built into the model. The other requirement is that the model should not include activities that have a negligible effect on the economic structure; in other words, the model should embrace only the outputs that involve the entire activity of a whole range of other servicing branches and the inputs that constitute the principal bottle-necks in the development of the national economy.

Both requirements are reasonable and neither can be dismissed offhandedly. Which one to enforce to a higher degree, is a difficult decision for the model constructor. It is possible to operate the two models alongside each other. It would be even more effective to enforce both aspects within one and the same model. In this case, however, the relationships that connect the full-range and the priority activities must be found and formally described, for example the connexion between the total output of the sector and the production volume of some of its major products. It will be necessary to quantify the residuum between the total output of the sector and the output of the priority products. All this involves rather intricate, though not insoluble, difficulties of planning methodology.

E. *Units of measurement*

The problem discussed in the preceding section leads to another: What units of measurement should be applied to the inputs and outputs appearing in the model? There are two principal kinds, value units and physical units. But with either, several questions of detail may arise. For example, in the case of measurement in value units, one must first decide whether to use domestic or foreign monetary units and then which value to measure with them, domestic or world-market prices, current prices or estimated future prices. In the case of measurement in physical units, one can apply several units of measurement to the same product or resource. Labour input, for example, may be measured in number of persons employed or in work hours.

Measurement in physical units has the advantage that it will better follow the actual trend of the real processes of production, consumption etc. Moreover, it is free of the fluctuations, arbitrariness and distorting effects of prices. On the other hand, different units cannot be added together.

The completeness requirement mentioned in the preceding section can generally be met only when measurement is in value units. However, in the case of certain main processes and structure-determining activities, it is not only possible but also more reasonable to apply physical units. Here again, the model constructor faces a dilemma. It is impossible to take a firm stand for either of the two basic methods of measurement; the application of both is desirable. Should there be no other solution, then let us have models measuring in value units and in physical units side by side. It will be even more effective to use both methods of measurement within the same model. In this case, however, a relationship must be established between the two, such as equations linking the total output of the sector in terms of value to the output of some principal products expressed in physical units.

F. *Cases of diminishing returns to scale*

We have said that we are describing a linear model. However, certain relationships between real economic processes are non-linear. One type of non-linearity is connected with the phenomenon which economists call "diminishing returns to scale". Instead of a precise definition, we give an example that will at the same time show a practical way of dealing with this phenomenon.

In the model described in section I, the export of product i to market m is represented by a single variable z_{imt}. Associated with this variable are a single upper export bound Z_{imt} and a single export price p_{imt}. It is possible, however, that the export price will actually depend on the exported quantity. In market m there will be several purchasers who may be inclined to pay a higher price. In such cases, the activity relating to the export of product i to market m will actually consist of several subactivities, represented by the variables $z^{(\mathrm{I})}_{imt}, z^{(\mathrm{II})}_{imt}, z^{(\mathrm{III})}_{imt}, \ldots$. The price will not be uniform:

$$p^{(\mathrm{I})}_{imt} > p^{(\mathrm{II})}_{imt} > p^{(\mathrm{III})}_{imt} > \ldots \tag{30}$$

This case may also be interpreted as one in which the returns from the export of product i to market m decrease as a function of the volume.

The phenomenon can also be taken into account with good approximation in the linear model. We need only treat the subactivities I, II, III, . . . as separate variables with separate upper export bounds:

$$\begin{gathered} z^{(\mathrm{I})}_{imt} \leq Z^{(\mathrm{I})}_{imt} \\ z^{(\mathrm{II})}_{imt} \leq Z^{(\mathrm{II})}_{imt} \\ \ldots \end{gathered} \tag{31}$$

Whenever we meet the phenomenon of diminishing returns to scale, we can always linearize it in this way. The programme will certainly first increase the first subactivity to its upper bound (provided it is worth while), and only then will it draw in the second-best subactivity. After exhausting the upper bound of the latter, the third one will be drawn into the programme, and so forth.

By means of this method of periodical linearization, the original non-linear function (in our example, the export price as a function of volume) can be approximated with the desired degree of accuracy. Unfortunately, the method involves an increase in the number of constraints and variables, that is, in the size of the model, and this works against the accuracy of the approximation. Here again, the model constructor's usual problem arises–a better approximation of reality is obtained at the cost of increased difficulty of computation.

G. *Indivisibility and increasing returns to scale*

There is another frequent kind of non-linearity that is not so simple to treat as the one just discussed. To this category belong the phenomena the economist calls "indivisibility", "non-convexity" and "increasing returns to scale". It is not intended to give exact definitions of these terms here; they can be found in the literature dealing with non-linear programming. Instead, we give some characteristic examples.

Certain actions cannot be carried out half way; they are realized either fully or not at all. It is not possible to build only half a bridge, from one bank of the river to its middle. The bridge itself may have several quantitative characteristics: it may be wider or narrower, it may take a long or short time to build etc. But the answer to the question whether there will be a bridge at all must be answered yes or no. It is not possible to represent a bridge of a given type by a continuous variable; it must be represented instead by a variable that can have one of two values, 1 or 0, meaning, respectively, that either the bridge will be constructed or it will not.

The dimensions of a plant cannot assume any arbitrary positive magnitude, particularly in such modern, highly concentrated industries as the chemical and motor-vehicle industries. The size has a lower limit that is determined by the design engineer. We either do not construct a fertilizer plant or construct one with a minimum capacity of 100,000 tons per year. A fertilizer plant with a capacity of 1,000 tons per year is practically unrealizable. The capacity variable is discontinuous: it cannot have values between zero and 100,000.

It is partly because of the phenomenon of indivisibility that mass production, larger plant size and the production of large series are accompanied by comparative cost savings. Some items, such as operating and investment costs, are fixed and more or less independent of the volume of production. The more units that are produced, the smaller the amount of such costs that can be assigned to each. Unfortunately, this phenomenon of increasing returns to scale cannot be treated with the simple method of linearization described in the preceding section for the case of diminishing returns. Nor will we deal here with the reasons–it suffices to point out the regrettable fact.

If the phenomena outlined above are ignored, the linear programming models presented in sections I and II, which employ continuous variables and linear equations, may lead to various inaccuracies. This applies especially to cases in which the phenomena of indivisibility and increasing returns to scale are prominent, such as infrastructural investment projects and studies of the comparative advantages of the international division of labour. To improve accuracy, the model constructor can use the so-called mixed model instead of the continuous linear model. Besides continuous non-negative variables, this model also has discrete variables that can have only certain values, such as variables that can be only 0 or 1, or only 0, 50,000 or 100,000. With this formalism he can represent exactly, or at least to a fairly good approximation, the phenomena of increasing returns to scale, non-convexity and indivisibility. However, mixed models involve considerably greater computational difficulties than those containing continuous variables only.[4]

In many countries, the enormous computer capacities required for solving such major mixed problems are not available. In such cases, the mathematical planner will have to content himself with naive heuristic methods. Of course, the phenomena in question will cause fewer difficulties when more highly aggregated models are used. The phenomena of indivisibility and of increasing returns will present themselves on the level of the individual plant rather than on that of an entire sector.

Moreover, it will be necessary to combine linear programming with manual checking and correction of the results obtained. Let us take the problem of indivisibility. Should a "too small" plant appear in the optimum programme, it should be left out of the recommendations and the estimates decreased accordingly. The variable which is too low may be left out of the model and the computation repeated. Conversely, the same variable can be forced to reach the normal plant size by prescribing a suitable lower bound, and the two programmes can be compared to see which yields a more favourable value for the objective function, the one in which the value of the variable is 0, or the one in which it has the normal plant size.

The input-output coefficients of the production and investment variables are usually estimated by the experts on the basis of some presumed plant size. Should the plant size obtained in the optimum programme differ strongly from this presumed size, the computation can be repeated, this time with the new coefficients

[4] The problems containing discrete variables only are usually simpler than those containing both discrete and continuous variables. If necessary, the model can be reformulated into a purely discrete problem, the only disadvantage being that this may to some extent decrease the accuracy of the approximation.

corresponding to the plant size obtained in the previous optimum programme. The result will show how sensitively the programme reacts to the change.

These methods of trial and error will help diminish the distortions caused by the continuity and linearity of the model. The amount of computational work will, of course, increase accordingly. This is why the decision makers should not base their decisions solely on multivariable mathematical models consisting of large simultaneous equation systems but complement the latter with project evaluation and profitability calculations that examine the possible major investment projects individually. True, in the course of these calculations many relationships that connect the individual investment projects with the rest of the actions planned must be neglected. The calculations facilitate, however, taking into account the indivisibilities and non-linear relationships characteristic of the individual project.

H. *Uncertainty*

In section III.A it was pointed out that it is not a single computation that must be carried out with our model but a whole series of them. Now, however, let us consider but one computation of the series. We proceed as if the matrix of coefficients A, the constraint vector **b** and the vector of objective function coefficients **c** were constant. We suppose them to be given and known, that is, certain, at least as far as the individual computation in question is concerned; in other words, we make the "lack of certainty" assumption (see section I.B.)[5] It must be obvious to every mathematical planner that the data are not certain, but at least he tries to get the best data he can.

Further efforts to relieve uncertainty can also be made. Some were mentioned in section III.A. Although within an individual computation the data are treated as constant, many computations are carried out in succession, the data introduced in the model being systematically changed each time. With the aid of such a sequence it is possible to examine the sensitivity of the programme to the uncertainty of the data.

It is obvious that there are practical limits to these experiments. Suppose our model contains a total of n data and each can independently assume m different values. It is uncertain which of the m different values is the "real", or correct, value. In this case, the computation should be repeated as many times as necessary to take into account all combinations of all the data with each of the m values. This would mean doing m^n computations, which is practically impossible with any linear programming model of realistic dimensions. We must therefore select from this enormous mass of computations a comparatively small number of the most characteristic cases, those that are most interesting from the economico-political point of view, and content ourselves with carrying out the computation of variants on the basis of these.

[5] In the literature, such a model is inappropriately called "deterministic" to distinguish it from the stochastic model.

The variant calculations and the sensitivity tests do not constitute the only solution. What is called stochastic programming may also be carried out. Here, some or all of the data for the model (the constraint coefficients, the objective function coefficients, the bounds) are treated as random variables instead of being characterized by a single constant. Determining the type of probability distribution of the random variables as well as the parameters characterizing the distribution (expected value, standard deviation etc.), we will, as a result of the computation, arrive at an optimum programme that is itself a probability variable. As a result of stochastic programming, the distribution of the optimum programmes may be established.

The promising methods of stochastic programming are, unfortunately, still at the beginning stage. They are at present applied to very special cases only, since their application involves a whole range of mathematical and computing difficulties. However, when researchers work out models and algorithms that lend themselves better to practical treatment than those known at present, stochastic programming is bound to become one of the regular tools of long-term planning. In view of the inevitable uncertainty of the data, this will be the most natural type of model in the field.

Another important problem is how to determine the risk involved in the decisions. The point is not simply that all data are more or less uncertain; it is that some alternatives of economic development involve less risk than others. For example, one way to increase production may be to count primarily on the normal expansion of the domestic market and on a low level of exports, at the same time restricting imports to a bare minimum and employing technologies already known. Another alternative may be to create new plants and even entire new industries that are mainly export-oriented, equip them with up-to-date and hitherto untested technologies and maintain a high level of imports from the export returns of their products. The latter alternative involves greater risks; both the foreign-trading and the technological data are uncertain. However, it may prove to be the more advantageous variant.

The mathematical formulation of the problems of this type is the task of what is called decision theory. There are many concepts, methods and models of decision theory that can be used in long-term planning, but they have not yet been extensively drawn upon. This, too, is a promise of the future rather than a tested practice of the present.

I. *Concluding remarks*

What has been said in this section allows certain general conclusions to be drawn.

The first is that no universally valid formula exists for the construction of industrial development models. Whichever problem he is approaching, the model constructor will be under pressure from two quarters. One demands a model of the greatest possible accuracy, reflecting reality as truly as possible and providing a direct answer to the greatest possible variety of questions. The other demands that the model should be produced within the shortest possible time with the least possible cost; it should require the least amount of data and computing work, and the results produced should be easy to survey. Every subsection in this section has contributed to making it clear that the construction of models involves a compromise between these two demands.

A further conclusion is that there are no miraculous models, not even adequate ones. Every reasonably constructed model will have both advantages and disadvantages. The mathematical planner should acquaint the decision makers with the assumptions and weaknesses of his models. The decision makers, on the other hand, must not take any model for a panacea, but for what it actually is: just one of many useful planning tools.

Finally, planning should be carried out whenever possible, not with a single model, but with a system of models that complement and compete with one another. Only by employing several model types collectively will it be possible to clarify with adequate reliability all essential aspects of the intricate problem of industrial development.