

Universität Bielefeld/IMW

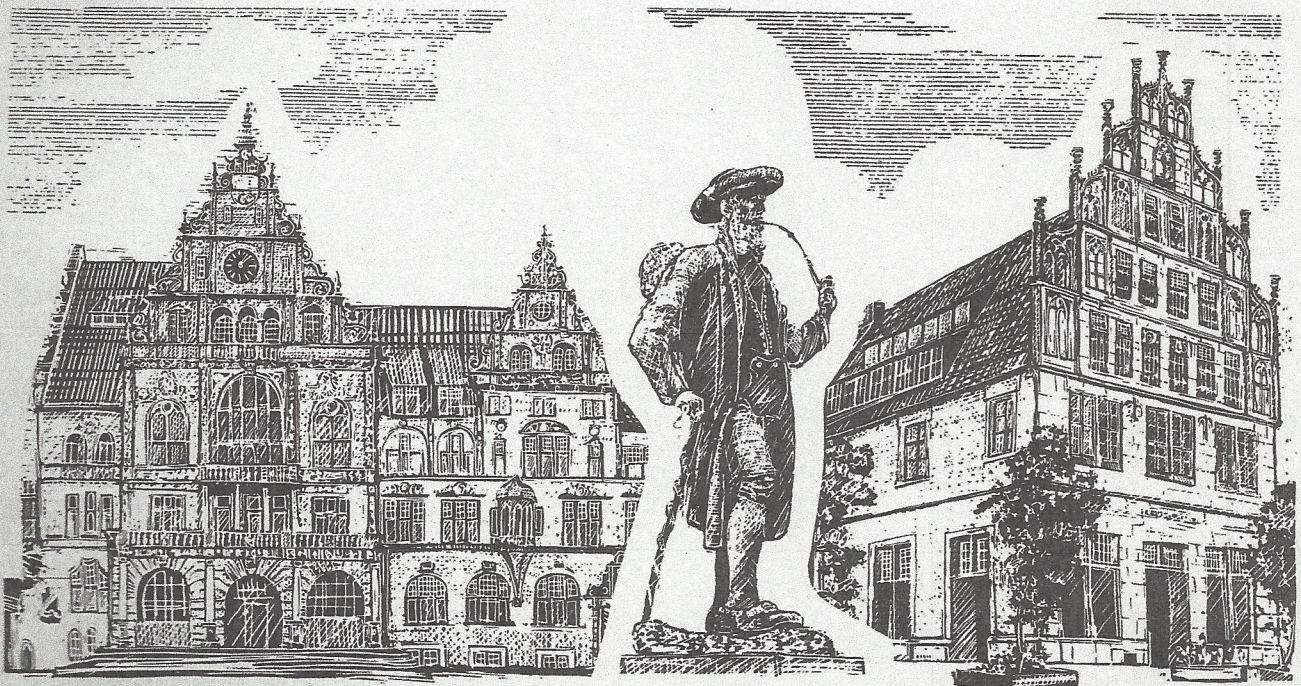
Working Papers  
Institute of Mathematical Economics

Arbeiten aus dem  
Institut für Mathematische Wirtschaftsforschung

Nr. 4

Carl Christian von Weizsäcker  
Long Term Global Optimization  
in Educational Planning. A Simple Example

September 1972



H. G. Bergenthal

Institut für Mathematische Wirtschaftsforschung  
an der  
Universität Bielefeld  
Adresse/Address:  
Universitätsstraße  
4800 Bielefeld 1  
Bundesrepublik Deutschland  
Federal Republic of Germany

Carl Christian von Weizsäcker

Long Term Global Optimization in Educational Planning.  
A Simple Example.

---

Introduction. Long term planning of the public sector has become increasingly important in most industrialized countries. The methods which are used so far are different from the sophisticated mathematical optimization techniques which have been developed in the different disciplines of operations research. While the application of cost-benefit analysis (which is also an optimization technique) for problems of a more local character has quite frequently been a success, more global tasks, such as planning for an entire national educational system, have not yet effectively been assisted by optimization techniques.

I have been involved in recent years in research related to educational planning in Germany. A team, of which I am a member, has developed a planning model and has provided data for it, so that we were able to make projections for the period from 1970 to the year 2000.<sup>1</sup> This work was closely related to work on an official general plan for the educational system of Western Germany and thus we were not free to choose our model. The model does not imply any optimization of an objective function and in many instances it introduces as fixed parameters what in optimization models would become a variable. It restricts itself to parameters and variables which are easily measurable. It is thus heavily input oriented, since the true outputs of educational processes are difficult to measure.

In the following I want to propose an approach to the problems of measurement of variables and parameters which are important but difficult to measure. This approach involves in an essential way the use of optimization techniques for educational

---

<sup>1</sup> C.C. von Weizsäcker, W.Konrad, H.Kurth, K.Uh Oh, W.Sutter, H.Vollet, Simulationsmodell für Bildungssysteme, Weinheim, Germany, 1972.

planning. I believe that the principle of the approach is quite general, but I do not want at the present moment to talk too much about generalities. I rather prefer to present a simple example which is of special interest to me because it is a first attempt (so far not a practical one) to overcome the deficiencies of educational planning models which I have mentioned above. But the generality of the principle underlying the simple example is mentioned as a justification for its presentation at this conference.

## I. The Model

To make things transparent we have chosen a rather simple model. It has four main parts:

1. the population and students model
2. the teacher supply model
3. the educational production model
4. the objective function.

### 1. The population and students model

Let  $n(j,t)$  be the number of persons of age  $j$  at time  $t$  in the society under consideration. It may be useful to explain  $n(j,t)$  by means of the following equation:

$$n(j,t) = p(j,t) n(o,t-j)$$

$n(o,t-j)$  is of course the birth rate at time  $t-j$ .

Let  $x(j,t)$  be the average amount of time used for learning of a person of age  $j$  at time  $t$ . We assume  $n(j,t)$  or  $n(o,t)$  and  $p(j,t)$  to be exogenously given. But  $x(j,t)$  is one of the important variables of the model. In this simple model time expenditure for learning is not divided up according to different fields or institutions of learning. The index  $j$  runs from  $o$  to  $J$ .

### 2. The teacher supply model

Let  $J_o$  be the age at which teachers normally enter the teaching profession.  $J_o$  is exogenously given. Let  $K_o$  (also exogenously given) be the normal length of time it takes to be trained as a teacher. Let  $Z(t)$  be the number of teachers leaving teaching training colleges and entering the teaching profession in period  $t$ .

To have a realistic model of teacher supply we must take into account that the tendency to leave or to reenter the teaching profession is strongly age dependent. It is therefore appropriate to differentiate teachers by age. Now for empirical and computational reasons it is perhaps sensible to distinguish only a few age groups each of which comprises several unit period age cohorts of teachers. Let  $k$  be the index of the teacher age groups,  $k$  goes from 1 to  $m$ . Let  $L(k,t)$  be the

number of teachers in age group  $k$  in period  $t$ . We then introduce the following system of equations

$$L(1,t+1) = Z(t) + [1 - \lambda - \lambda(1) - \lambda(1,2)] L(1,t)$$

$$L(2,t+1) = \lambda(1,2) L(1,t) + [1 - \lambda - \lambda(2) - \lambda(2,3)] L(2,t)$$

$$L(k,t+1) = \lambda(k-1,k) L(k-1,t) + [1 - \lambda - \lambda(k) - \lambda(k,k+1)] L(k,t)$$

for  $k = 2, 3, \dots, m-1$

$$L(m,t+1) = \lambda(m-1,m) L(m-1,t) + [1 - \lambda - \lambda(m)] L(m,t)$$

The  $\lambda(k,k+1)$  are the transition proportions between the different age groups. Naturally they are inversely related to the number of unit period age cohorts contained in age group  $k$ .  $\lambda + \lambda(k)$  is the proportion of net flows from teachers of group  $k$  into the outside world. It consists of the constant  $\lambda(k)$  and of a variable part  $\lambda$  whose determinants will be discussed below. It is an interesting question how to construct age groups optimally, if we want a model in which each age group is to be considered homogeneous in the way we have done above.

Let us now consider the influences on the teacher flows. Let  $w(t)$  be the wage rate of teachers at time  $t$  as a proportion of the average wage rate in the economy. Let  $d(t)$  be the teaching load of a teacher at  $t$ . Let  $\hat{w}(t)$  be an exponentially weighted average of past values of  $w$ , so that

$$\hat{w}(t) = \gamma w(t) + (1-\gamma)\hat{w}(t-1)$$

Similarly  $\hat{d}(t)$  is defined by

$$\hat{d}(t) = \gamma d(t) + (1-\gamma)\hat{d}(t-1)$$

We now assume that  $\lambda$  is a function of  $\hat{w}$  and  $\hat{d}$

$$\lambda = \lambda(\hat{d}, \hat{w})$$

Also  $Z(t)$  is supposed to depend on the working conditions of teachers. Remember that  $K_0$  is the time it takes to train a teacher. We then assume

$$Z(t) = q(t)n (J_0 - K_0, t - K_0)$$

where  $q(t) = q(\hat{w}(t), \hat{d}(t))$

The proportion of an age group who want to become teachers,  $q$ , depends on the working conditions of teachers. The variables  $w(t)$  and  $d(t)$  are control parameters of the planners. Many reasons point to the assumption that the rates of change of  $w$  and  $d$  cannot be too large, hence we introduce the restriction

$$|\Delta w(t)| = |w(t) - w(t-1)| \leq \Delta_1$$

$$|\Delta d(t)| = |d(t) - d(t-1)| \leq \Delta_2$$

where  $\Delta_1$  and  $\Delta_2$  are constants.

### 3. The educational production model

This model defines the "production functions" of the educational process. The total supply of teacher inputs, measured in efficiency units,  $A(t)$ , is determined by the total number of active teachers

$$L(t) = \sum_k L(k, t)$$

and the average teaching load  $d(t)$ . We assume

$$A(t) = L(t) g(d(t))$$

Effective teacher input is proportional to the number of teachers and a function  $g$  of the teaching load. For sufficiently large values of  $d$  a law of diminishing (at last perhaps even negative) marginal efficiency of the teaching load operates. We thus may assume  $g''(d) < 0$  for sufficiently large  $d$  and  $g'(d) > 0$  for sufficiently small  $d$ .

Let  $l(j, t)$  be the teacher intensity of the education of age group  $j$ . The total teacher input requirement at time  $t$  is therefore

$$\sum_{j=0}^J n(j, t) x(j, t) l(j, t)$$

It cannot be larger than  $A(t)$ .

Let  $B(j,t)$  be the level of education already attained by group  $j$  at time  $t$ . Here we shall not discuss how to measure  $B(j,t)$ . More will be said on this point later in II. Let  $D(j,t)$  be a measure of what has been learned by age group  $j$  during period  $t$ . We then assume

$$B(j+1,t+1) = (1-h) B(j,t) + D(j,t), \quad 0 < h < 1$$

This means: The level of knowledge of the group born in period  $t-j$  in period  $t+1$  is explained by the level of knowledge of that group in period  $t$  and the amount which has been learnt during period  $t$ . Because of subjective processes of forgetting and objective processes of obsolescence there arises a phenomenon of "depreciation" of knowledge which is represented in the equation above by the rate of obsolescence  $h$ .

The output of the learning process of age group  $j$ ,  $D(j,t)$ , is now linked to the inputs into that learning process by means of an "educational production function"

$$D(j,t) = \alpha \Psi(j, x(j,t)) \varphi(j, \beta(j,t))$$

where  $\beta(j,t)$  represents the teacher inputs and physical inputs (teaching material, equipment, buildings, etc.) per unit of time used by age group  $j$  for learning purposes. In general we can assume that  $\varphi$  and  $\Psi$  are functions exhibiting diminishing marginal returns, at least for sufficiently large values of their arguments. The number  $\alpha$  is basically a parameter for the relative importance attached to learning by society. The variable  $\beta(j,t)$  is defined by means of a linear homogeneous production function whose inputs are  $l(j,t)$ , the teacher input intensity and  $s(j,t)$ , the physical input intensity.

$$\beta(j,t) = F(j, s(j,t), l(j,t))$$

where  $F$  is homogeneous of degree one in  $s(j,t)$  and  $l(j,t)$ . This assumption is less restrictive than one may perhaps suppose. Basically it means that the set of indifference curves for a given educational output and given other inputs on the  $s(j,t), l(j,t)$  diagram is homothetic. If that is the case a linear homogeneous function exists which can

be taken as a representation of joint teacher and physical inputs such that it reflects accurately the effect of these two inputs on the output.

I may refer the reader to the theory of "true"-indices, such as the true cost of living index or a "true" index of real income. Any nonproportionality between inputs and output can be captured by the properties to be assumed about  $\varphi(j, \beta(j, t))$ .

#### 4. The objective function

We want to maximise the value of the function

$$V = \sum_{t=t_0}^{\infty} R^{t-t_0} \left[ \sum_{j=0}^j (B(j, t) - x(j, t) s(j, t) - v(j)x(j, t)n(j, t)) \right] - \sum_{t=t_0}^{\infty} R^{t-t_0} w(t)L(t)$$

The number  $R$  is a discount factor which we introduce here as an exogenous parameter. It would be of particular interest to observe the sensitivity of the optimal solution to changes in  $R$ . As the last term of the objective function indicates we are measuring all values in units which are equivalent to the average wage rate in the economy, since  $w(t)$  was defined in these units. If the average wage rate rises through time the discount factor  $R$  is different from what it would be if we had chosen money units to measure the values. Indeed, if the rate of growth of the average wage rate is equal the rate of interest (a situation not very far off from reality in many western countries), the discount factor would be equal to unity. The first term in square brackets,  $\sum_t \sum_j B(j, t)$ , represents the output of the educational system in terms of the achievements in training of the population. The second term,  $-\sum_t R^{t-t_0} \sum_j x(j, t) s(j, t) n(j, t)$  represents the costs of physical inputs into the educational system, the third term,  $-\sum_t \sum_j x(j, t) v(j) n(j, t)$  represents the opportunity costs of the time spent by pupils in the educational process. We



assume here that the opportunity costs per unit of time,  $v(j)$ , only depend on  $j$  and not on  $t$ . The fourth term  $-\int_t^{\infty} R^{t-t_0} w(t) L(t)$  represents the costs of teachers.

## II. Steady State Analysis

Our main purpose in the present paper is the development of a method which allows us to obtain reasonable data for long term optimization models such as the example discussed here. The method is quite general and it can be used for other planning problems too. In particular, it can be used to obtain approximate data for variables measuring the global benefit of activities such as education, health services, traffic, etc. where conventional methods of measurement are difficult to apply. On the other hand, I must warn the reader not to expect revolutionary advances from this method. On the contrary, in a way it is antirevolutionary, because it starts with the assumption that there are good reasons for the real world processes corresponding to our model processes to take on the values of the variables they are observed to take on. It does not mean that these processes develop optimally, but it means that realistically we should not strive for more than gradual and piecemeal improvements.<sup>1</sup>

Our method proceeds as follows. There are quite a few functional relations whose parameters are not known to us. Now we look at stationary (steady state) optimal solutions of our model, which have rather convenient mathematical properties. It is fairly easy to compute optimal steady state solutions for given parameter values of the functional relations. We thus can study the implications of different sets of parameters on the optimal values of control and state variables of the system without having to solve a rather complex dynamic program. The control variables and state variables (such as

---

<sup>1</sup> cf Popper's concept of piecemeal engineering in K. Popper, *The Open Society and its Enemies*, Princeton, 1950, or Braybrooke's and Lindblom's concept of disjointed incrementalism in D. Braybrooke, Ch. Lindblom, *A Strategy of Decision*, Glencoe, London 1963.

teachers' salaries, teaching loads, teacher inputs, physical inputs, time expenditure for education, etc.) frequently have real world counterparts for which statistical data are available or can with some reasonable effort be made available. We thus can ask the question: which sets of (not directly measurable) parameters of the model are consistent with optimal values of control and state variables corresponding to real world values. Assuming that the real world values are not very far from a steady state optimum we get **hints at** the values of the parameters of the functional relations in the model. Not all parameters can be estimated in this way. The steady state analysis leaves us with a few degrees of freedom with respect to parameter values consistent with the steady state optimality hypothesis of observed data. Indeed these degrees of freedom are crucial to make the whole analysis worthwhile, ~~since~~ it is our hypothesis that the properties of nonsteady state optimal paths are quite sensitive to assumptions made which allow us to get rid of the indeterminacy of the parameter values. It is then our final purpose to study the implications of the choice of the remaining free parameters on the optimal path. Here we no longer assume that the real world development of the past has been optimal, since, of course, this real world did not solve a complicated dynamic program before starting on its course. The computational simplicity of policy decisions in the real world probably makes our approach justifiable: the attempts to optimize, which implicitly are made by policy makers, are of a computational nature similar to the analysis of steady states. For instance we may refer to discussions whether it is worthwhile to reduce pupil-teacher ratios, to raise the school leaving age, to expand the university systems. Thus we may think of observed variables to be the result of an "as-if-steady-state-optimization." On the other hand, experience shows that the intricacies of the dynamics of an educational system have been beyond the computational capabilities of policy makers and public opinion. As an example we may point to the worldwide phenomenon of overreacting with respect to shortages of qualified manpower causing the surpluses or tendencies towards surplus observed in almost all industrialized western countries.

Thus the real path will not be optimal dynamically and it is here that models of the type proposed in this paper could help.

In our steady state analysis we put  $R = 1$  and we assume a stationary population. There arises the difficulty that the objective function may diverge in this case. We therefore do not work with the function  $V$  but rather with the function  $\bar{V}$  giving the average value of the components whose sum (over  $t$ ) is equal to  $V$ . Since we look at steady states the values of the variables do not change with  $t$  and we may thus write (dropping the index  $t$ )

$$\bar{V} = \sum_{j=0}^J n(j) [B(j) - x(j) s(j) - x(j)v(j)] - wL$$

It is reasonable to assume  $B(0) = 0$ . We then can write, using the equation of the educational production model,

$$B(j) = \sum_{i=0}^{j-1} (1-h)^{j-i-1} D(i)$$

Also

$$A = Lg(d) = \sum_j n(j) x(j) l(j)$$

implies

$$\frac{\partial L}{\partial l(j)} = \frac{n(j)x(j)}{g(d)}$$

We have to recognize that there exists a functional relationship between the stationary values of  $w$  and  $d$  and the stationary value of  $L$ . We can write  $L=L(w,d)$  and we assume

$$\frac{\delta L}{\delta w} > 0, \frac{\delta L}{\delta d} < 0.$$

To incorporate the restriction concerning teacher input into our analysis we form the Lagrange expression

$$\begin{aligned} Lg\bar{V} = & \sum_j n(j) B(j) - wL(w,d) - \sum_j n(j) x(j) (s(j) + v(j)) \\ & + \mu(L(w,d) g(d) - \sum_j n(j) x(j) l(j)) \end{aligned}$$

We differentiate with respect to  $w$  and  $d$  and put the derivation equal to zero,

$$\frac{\delta Lg \bar{V}}{\delta w} = -L(w,d) - \frac{\delta L}{\delta w} w + \mu g(d) \frac{\delta L}{\delta w} = 0$$

$$\frac{\delta Lg \bar{V}}{\delta d} = -w \frac{\delta L}{\delta d} + \mu g(d) \frac{\delta L}{\delta d} + \mu L(w,d) g'(d) = 0$$

Let us now assume that we know the elasticities  $\epsilon(w) =$

$$\frac{\delta L}{\delta w} \cdot \frac{w}{L} \quad \text{and} \quad \epsilon(d) = \frac{\delta L}{\delta d} \cdot \frac{d}{L} \quad \text{describing the long run (i.e.}$$

steady state) supply behaviour of teachers. The first equation can be written

$$1 + \frac{\delta L}{\delta w} \frac{w}{L} = 1 + \epsilon(w) = \frac{\delta L}{\delta w} \frac{\mu g}{L} = \frac{\mu g}{w} \epsilon(w)$$

or

$$\frac{g(d) \mu}{w} = \frac{1 + \epsilon(w)}{\epsilon(w)}$$

$\mu$  is the shadow price of an efficiency unit of teacher input, thus  $\mu g(d)$  is the shadow price of a teacher. The ratio of shadow price and market price of a teacher is determined by the wage elasticity of supply of teachers. It is, of course, the Cournot point of the monopsonist "educational system" on the teacher market. Let us now look at the second equation where we replace  $g\mu$  by  $w \frac{1 + \epsilon(w)}{\epsilon(w)}$ . After dividing by  $wL$  and multiplying with  $d$  we have

$$- \frac{d}{L} \frac{\delta L}{\delta d} + \frac{1 + \epsilon(w)}{\epsilon(w)} \frac{d}{L} \frac{\delta L}{\delta d} + \frac{1 + \epsilon(w)}{\epsilon(w)} \frac{dg'(d)}{g(d)} = 0$$

$$- \epsilon(d) + \frac{1 + \epsilon(w)}{\epsilon(w)} \epsilon(d) + \frac{1 + \epsilon(w)}{\epsilon(w)} \frac{dg'(d)}{d} = 0$$

$$\frac{\epsilon(d)}{\epsilon(w)} + \frac{1 + \epsilon(w)}{\epsilon(w)} \frac{dg'(d)}{g(d)} = 0$$

$$\frac{dg'(d)}{g(d)} = - \frac{\epsilon(d)}{1+\epsilon(w)} = \frac{|\epsilon(d)|}{1+\epsilon(w)}, \text{ since } \epsilon(d) < 0$$

The elasticity of teacher input with respect to teaching load in the optimum is proportional to the absolute value of the long run teacher supply elasticity with respect to the teaching load. These formulas are of some interest. Assuming that they are approximately fulfilled in the real world they allow us to compute certain parameters, if others are known. If, for example, we have estimates for  $w, d, \epsilon(w), \epsilon(d)$  we can compute  $\frac{dg'(d)}{g(d)}$  and  $\mu g(d)$ , magnitudes indicating something about the effects of teaching and variations of teaching loads on the output of the educational system (as seen by decision makers).

The effect of  $w$  and  $d$  on the long run value of  $L$  is intermediated by the dependence of  $q$  and  $\lambda$  on  $w$  and  $d$ . We shall not discuss the steady state properties of this functional relationship. Let us just say that this analysis gives us insights about the parameters of the functions  $\lambda(\hat{w}, \hat{d})$  and  $q(\hat{w}, \hat{d})$  which are important for a dynamic analysis.

We now turn to a discussion of the "production function." For this purpose we differentiate and put the derivative equal to zero:

$$\frac{\delta Lq\bar{v}}{\delta l(i)} = \sum_j n(j) \frac{\delta B(j)}{\delta l(i)} - un(i)x(1) = 0$$

or, remembering the formula for  $B(j)$ ,

$$\frac{\delta D(i)}{\delta l(i)} \left[ \sum_{j=i+1}^J n(j) (1-h)^{j-i-1} \right] - un(i)x(i) = 0$$

$$\text{We define } H(i) = \sum_{j=i+1}^J n(j) (1-h)^{j-i-1}$$

and obtain, remembering the structure of  $D(i)$ ,

$$H(i) \propto \Psi(i, x(i)) \varphi'(i, \beta(i)) \frac{\delta F(i)}{\delta l(i)} - \mu n(i) x(i) = 0$$

$$\text{where } \varphi'(i, \beta(i)) = \frac{\delta \varphi(i, \beta(i))}{\delta \beta(i)}$$

Similarly optimization with respect to  $s(i)$  yields

$$H(i) \propto \varphi(i, x(i)) \varphi'(i, \beta(i)) \frac{\delta F(i)}{\delta s(i)} - n(i) x(i) = 0$$

These two equations imply

$$\frac{\delta F(i)}{\delta l(i)} = \mu \frac{\delta F(i)}{\delta s(i)}$$

But this is the condition for maximisation of  $F(i)$  under the constraint of given "costs" of inputs

$$C(i) = s(i) + \mu l(i)$$

where teacher inputs are weighted with their shadow price  $\mu$ . Hence the problem of choosing the right point on the isoquant corresponding to the function  $F(i)$  can be solved whenever  $\mu$  is known and without regard to similar problems for other age groups  $j$ . This property should facilitate the computation of a dynamic optimal program. Given  $\mu$  we can now consider  $F(i)$  as a function of the "costs"  $c(i)$  and hence also  $\varphi(i)$  can be interpreted as a function of costs:  $\varphi(i) = \varphi(i, c(i), \mu)$ .

There remains the problem of optimization with respect to  $x(i)$ . We differentiate and put the derivative equal to zero

$$\frac{\delta Lg\bar{V}}{\delta x(i)} = \sum_i n(j) \frac{\delta B(j)}{\delta x(i)} - v(i)n(i) - s(i)n(i) - \mu n(i)l(i) = 0$$

or

$$H(i) \propto \Psi'(i, x(i)) \varphi(i) = n(i) [v(i) + s(i) + \mu l(i)]$$

$$\text{where } \Psi'(i, x(i)) = \frac{d\Psi(i, x(i))}{dx(i)}$$

Obviously we have  $\frac{\delta \varphi(i, c(i), \mu)}{\delta c(i)} = \varphi'(i) \frac{\delta F(i)}{\delta s(i)} =$

$$= \frac{1}{\mu} \varphi'(i) \frac{\delta F(i)}{\delta l(i)} \text{ and hence the optimization condi-}$$

tions with respect to  $s(i)$  or  $l(i)$  yield

$$H(i) \propto \varphi(i) \frac{\delta \varphi(i)}{\delta c(i)} = n(i) x(i)$$

Rewriting this equation and the optimization condition with respect to  $x(i)$ , we obtain

$$\frac{\Psi'(i)x(i)}{\Psi(i)} = n(i) x(i) \frac{v(i) + c(i)}{H(i) \propto \varphi(i) \Psi(i)}$$

$$\frac{\delta \varphi(i)}{\delta c(i)} \frac{c(i)}{\varphi(i)} = n(i) x(i) \frac{c(i)}{H(i) \propto \varphi(i) \Psi(i)}$$

This implies

$$\frac{\delta \varphi(i)}{\delta c(i)} \frac{c(i)}{\varphi(i)} = \frac{c(i)}{c(i)+v(i)} \frac{\Psi'(i)x(i)}{\Psi(i)}$$

This is an interesting equation since it yields a simple relation between the elasticity of educational output with respect to direct costs of education (teachers and physical inputs) on the one side and the elasticity of educational output with respect to pupils' time expenditure on the other side. The values of  $c(i)$  and  $v(i)$  are not difficult to measure or estimate in principle. It is then possible to get estimates for these elasticities of educational output from the last three equations, if we assume that real world variables are similar to steady state optimal conditions. In preparing a numerical treatment of the problem we would choose functions  $\varphi$  and  $\Psi$  which are characterized by a small number of parameters and then use the derivation above to get some indication about the numerical values of the parameters. I have made some investi-

gations in this direction. I shall not describe them here in order to avoid having to write a rather lengthy paper.

There is one additional interesting problem which so far has not been discussed. It is the extent to which it is possible to substitute physical inputs for teacher inputs or vice versa. To treat this problem it is best if we assume that the functions  $F(j)$  are of the CES-type

$$F(j) = [\delta(j) l(j)^{-\rho} + (1-\delta(j)) s(j)^{-\rho}]^{-\frac{1}{\rho}}$$

where  $\sigma = \frac{1}{1+\rho}$  is the elasticity of substitution.

Steady state analysis does not allow us to obtain indications about the value  $\sigma$ . Unless there are other indications about the elasticity of substitution we shall have to try out how sensitive the optimal dynamic program reacts on changes in  $\sigma$ . For any given  $\sigma$  we are able to find indications about  $\delta(j)$ ; we are able to derive in a straight forward fashion the following relation between  $c(j)$  and  $s(j)$

$$c(i) = \left[ \mu^{1-\sigma} \left( \frac{\delta(j)}{1-\delta(j)} \right)^{\sigma} + 1 \right] s(j)$$

Since we have estimates for  $s(j)$ ,  $c(j)$  and  $\mu$ , this equation and the choice of the parameter  $\sigma$  determine the parameter  $\delta(j)$ .

These are a few examples for finding empirically reasonable values of the parameters of the model by means of steady state analysis.

### III. Outlook on the dynamic analysis

I shall not discuss here the technical and mathematical problems of setting up a dynamic program to solve the optimization problems. I only want to make a few remarks concerning the purpose of this exercise. Since it will take some effort to carry this optimization through it is worthwhile to speculate about possible results. I shall give an example: the problem of forecasting teacher requirements:



In the sixties and seventies quite substantial quantitative and qualitative changes in the educational systems of most countries took place and will take place. They are thus far from a steady state. In their plans for the educational system planners and governments have tried to forecast the teacher requirements over a period of 10 to 20 years into the future. Their method has been to lay down teacher-pupil ratios and then to forecast the number of pupils. In the language of our model, the  $l(j)$  and the  $x(j)$  have been fixed exogenously and then the teacher requirements have been computed. Is this a good method to plan the educational system? We want to study the dynamics of our model among other things in order to answer this question.

The following two hypotheses draw on the intuition of the economist. The hypotheses, if warranted by a more thorough investigation, would be a negative answer to this question. The two hypotheses apply to two different parameter constellations, which we call the high elasticity and the low elasticity case. The high elasticity case is characterized by a high elasticity of substitution  $\sigma$  in the production function and a high sensitivity of the optimal cost level  $c(j)$  with respect to changes of parameters, say,  $\alpha$ . In the low elasticity case the opposite is true. Our hypotheses are:

- 1) In the high elasticity case a substantial change in, say,  $\alpha$  will induce substantial changes in the input mix between physical inputs and teachers in the direction of higher physical inputs. Costs  $c(j)$  per unit of pupil time will rise substantially, basically by raising  $s(j)$  and  $\mu$ , not so much by raising  $l(j)$ . The relative wage rate of teachers will rise. The additional supply of teachers effected thereby will mainly be used for raising education time of pupils ( $x(j)$ ) not so much for raising  $l(j)$ , which may actually fall. The optimal value of  $l(j)$  is mainly affected by the teacher supply conditions and by the size of the system ( $x(j)$ ). It is thus not reasonable to introduce  $l(j)$  as a parameter which is independent of the teacher wage rate, the teacher supply rate and the size of the system (number and time expenditure of pupils).

- 2) In the low elasticity case the  $l(j)$  are rather **insensitive**

to changes of parameters like  $\alpha$ . Their introduction as parameters in a conventional planning model is therefore not a severe mistake. But the low elasticity case also implies that  $l(j)$  should remain rather stable over time which it usually does not in conventional planning models. Also the rate of expansion of the system (rate of change of  $x(j)$ ) should not be introduced as an independent parameter. Rather the optimal development of  $x(j)$  depends very much on the supply conditions of teachers and the parameters  $l(j)$ .

Thus in both cases the traditional planning approach is not justified, if our hypotheses are validated, by the computation of optimal programs.