AN INTERACTIVE SYSTEM FOR REGIONAL ANALYSIS OF INDUSTRIAL SECTORS

Börje Johansson*
Bertil Marksjö**

WP-83-54
May, 1983

*Integrated Regional & Urban Development Group, International Institute for Applied Systems Analysis (IIASA), Laxenburg, Austria

**Committee for Regional Planning and Commerce, Stockholm County Council, Stockholm, Sweden

PREFACE

This paper has been written as a contribution to the book "Information Systems for Integrated Regional Planning", edited by P. Nijkamp and P. Rietveld. It reports on two interactive information systems for national and regional analysis of industrial sectors--their internal structure, and its change.

The methods for designing the systems are described, and special attention is given to the confidentiality criteria which such systems must satisfy in Sweden, if they are to be used publically.

In order to describe the reliability of the decision support generated by the systems, the paper describes the underlying economic theory, the statistical reliability of the processed information and the time aspects of the reliability. In addition, the paper outlines the practical use of the systems as well as the dialogue and graphical options they contain.

One of the systems is used by the Industrial Board of the Swedish Ministry of Industry. Information from this system was utilized in IIASA's Skåne case study.

Börje Johansson
Acting Leader
Regional & Urban Development Group

May, 1983

CONTENTS

1. INTRODUCTION

This paper presents two information systems which generate regional information about industrial plants (SIND-data)[1] and firms (SUPPORT) in Sweden. SIND-data is installed in the National Industrial Board; SUPPORT has a stronger orientation towards business corporation decision-makers. Both systems are used for regional planning, impact studies and forecasting.

The presentation is focused on the methods and theories utilized to construct and design information systems like SIND-data and SUPPORT. It also reports on some of the different ways in which the systems and their dialogue and graphical options may be utilized.

The SIND-data system may be used to illustrate the content of the paper. This system contains, for example, information about how production techniques, productivity and profitability of establishments adhering to a given sector are distributed in a region. In Sweden, like in many other countries, there exists a confidentiality criterion which prohibits that information about individual units is exposed to the public. To satisfy this criterion, establishments have to be grouped together. Section 2 outlines how this problem has been approached by means of a filter process. Section 3 describes this process in mathematical form and illuminates its general features. The method arranges the original data in groups which satisfy requirements of (i) confidentiality, (ii) aggregation consistency, (iii) group homogenity, and (iv) distribution accuracy.

We also describe in Section 2 how the design of the systems has been guided by a principle requiring that (i) the information shall be reliable in view of the purpose it is used for, (ii) its expected degree of uncertainty shall be measurable and known. This is obtained by organizing the information in the form of distributions which are structurally invariant over long-time periods.

We may say that statistical and economic theory, together with estimated invariances, become integral parts of the information system by being introduced already at the design stage.

---

[1] A system similar to SIND-data exists in Norway. It is called NIND-data.

2. USER REQUIREMENTS, THEORY AND ESTIMATED INVARIANCES

2.1 Principles for Designing an Information System

Inforation systems may be looked upon and evaluated from different perspectives. In the subsequent presentation we shall interpret them as systems for decision support (see e.g., Marschak (1974)). The following ways of utilizing an information system are identified in Nijkamp (1982):

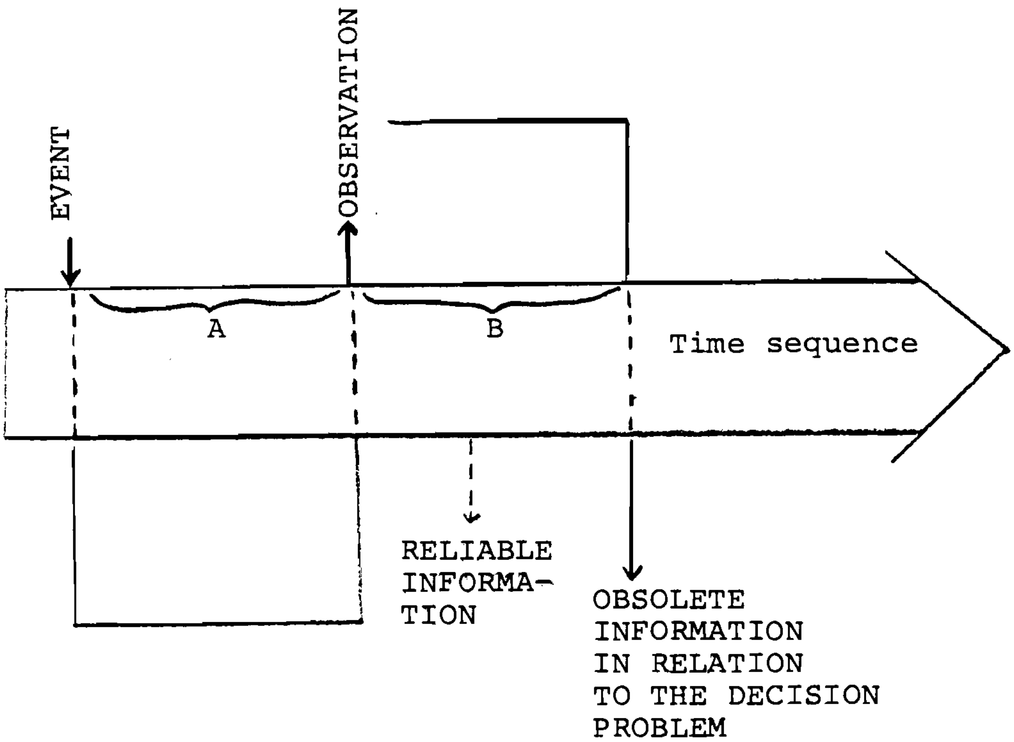
- o Description, monitoring, and ex-post evaluation
- o Impact analysis, scenario projections, and forecasts
- o Decision analysis, evaluation of options, planning

If the user is an organization which makes diagnoses and forecasts, these activities will represent the decision-making of the user. Basic user requirements of an information system are: accessibility to the information, flexibility of the system (multipurpose use of the information), decision relevance and reliability of the information.

User of information systems often reveal an anxious fixation on the problem of obtaining information with a high degree of "actuality". This property does not have to depend only on the observation lag, i.e. the time span between the occurrence of an event and the observation/recording of the event in the information system. In fact one may distinguish between "perishable" and "durable" information, where the latter retains its actuality over a longer time interval. Durable information will generally refer to decision-relevant structural features of a regional system. We may refer to this as information about structural invariances.

Although time invariant information is usually obtained by processing tome sequences of observations, we may still illustrate the basic idea by Figure 1. This figure depicts the trade off between reducing the observation lag A and increasing the invariance interval B within which the information can be regarded as reliable.

We shall present two different information systems both of which were set up with the ambition of satisfying two



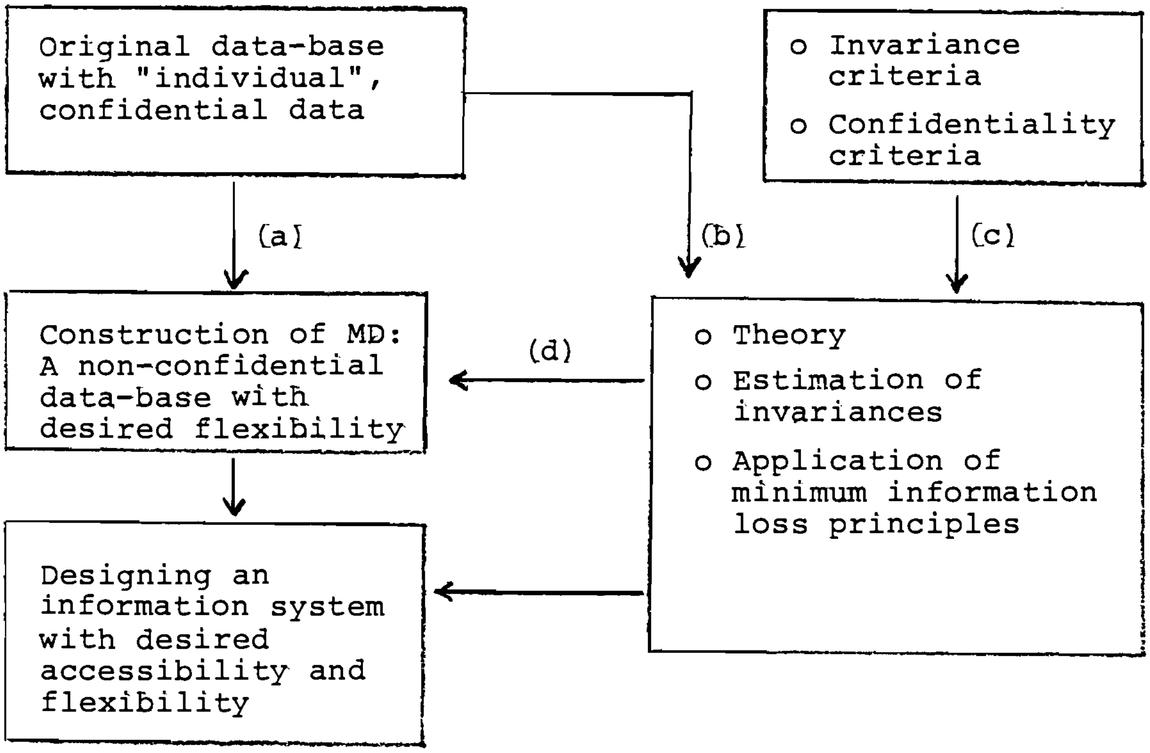
A = Observation lag
B = Reliability-invariance interval

Figure 1. Trade-off between observation lag and reliability interval.

criteria. It was required that the information should have a form which made it decision-relevant and in addition gave it a high degree of invariance so that interval B in Figure 1 could be extended. Moreover, the degree of reliability and invariance was measured and analyzed for the type of information that can be generated by the systems.

The following approach was selected. The original data set is transformed to distributions for which the degree of invariance can be evaluated over time. In one of the systems the distributions refer to characteristics of industrial establishments (production units). This system is called SIND-data. The other system, SUPPORT, refers to industrial firms, where a firm (as distinct from a production unit) represents a decision-making and financially defined unit. For each firm one can always identify one or several establishments belonging to the firm. Data referring to individual establishments and firms are regarded as confidential in Swedish statistical systems.

In order to simultaneously satisfy the invariance and confidentiality criteria[1] the systems are up-dated annually inside the Bureau of Statistics (SCB). Inside the bureau one is allowed to use confidential data to evaluate invariances and to transform the original data to a model of these data (MD). The MD is designed in such a way that it satisfies all confidentiality criteria at a "minimum loss of information". Thereafter the MD can be transferred to users outside the bureau as an integral part of the Info-system. In this sense the MD is an input to the production of information with the information system. The above procedure is illustrated in Figure 2a.



Remark: When confidential data are used, operations (a)-(d) have to be done inside the Bureau of Statistics.

Figure 2a. Designing and up-dating the information system.

---

1) The confidentiality criterion requires that observation units are ordered in groups. Such a group can never contain less than three units. For a definition of the criterion, see Section 3.

In order to be more specific about the structure of transformations underlying Figure 2a, one may consult Figure 2b to envisage the path from "raw" data to decision support information. This illustrates different states in which data may exist and the associated transformations.

Figure 2b illustrates the difficulties of designing a multipurpose information system. To have some success it is essential to know the decision context before designing the filtering according to confidentiality rules and invariance criteria. The SIND-system mentioned above may illustrate some of these aspects. In that case sorting is made with regard to the ratio between wage and value added (wage share) of each unit, and the filter is used for each regional level. To be more specific, in order to include region A, B, and $A \cup B$ in data set (3), three separate "filtering runs" are necessary. If the non-confidential aggregates in set (3) are required to contain units with similar capital costs instead of wages shares, then another filter process would have to be applied.[1]

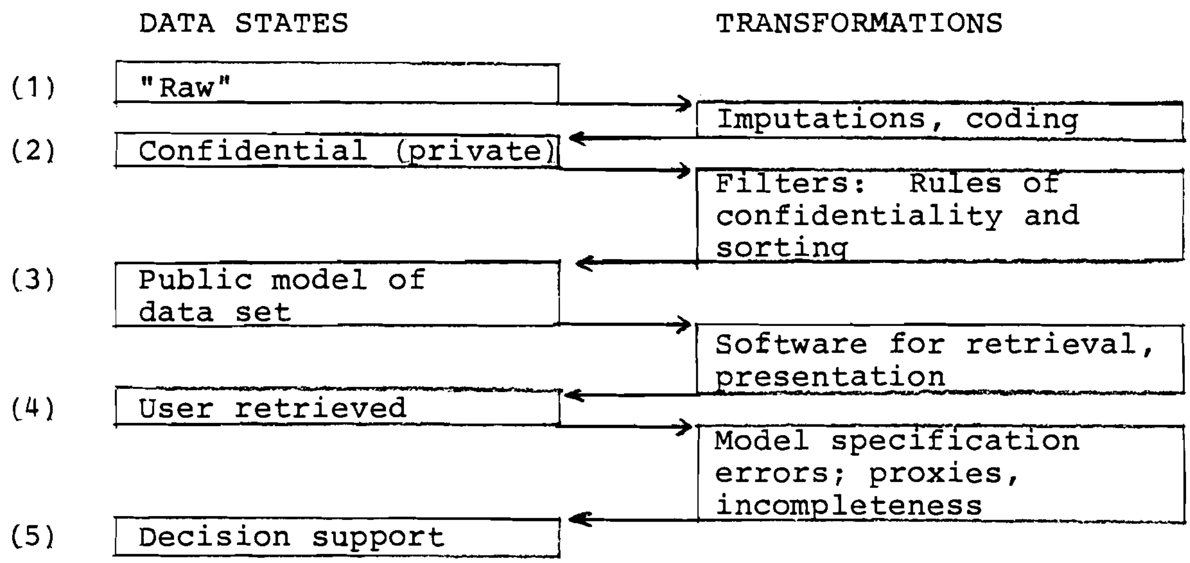


Figure 2b. Stages of data transformation.

---

[1]The SUPPORT system contains separate models of the data set based on (i) wage share-sorting, and (ii) rate of return-sorting; from both models capital cost information can be retrieved.

2.2 Variable Specification of Regions and Sectors

In the preceding section we described a method utilized to design two information systems in such a way that certain criteria were satisfied. Both systems have also been designed in such a way that information is generated

- for varying spatial aggregation levels,
- for different sectoral aggregation levels,
- in a form which makes it possible to separate short-term variations and structural changes.

Three types of geographical subdivisions exist. These are (i) national, (ii) industrial region, and (iii) county level. There are six industrial regions; each of them consists of several counties. The total number of counties is 24.

The SIND-system contains ca. 100 subsectors of the manufacturing and mining industries. These sectors can be aggregated in several different combinations. With certain combinations between region and sector it is not possible to generate any information due to confidentiality restrictions. Then the system offers two options. One is to retain the fine sector level and enlarge the region through spatial aggregation. The other is to aggregate several sectors at the initially given regional level.

Each type of distribution generated by the systems may either be presented in the form of a table or a diagram. A diagram has the principle form described in Figure 3. For each diagram one may obtain two distribution curves. Each curve may be selected in continuous, smooth form or in a discrete (step-function) form.

The curves in Figure 3 will represent one of the following two types of comparisons:

- o Curve (a) and (b) refer to two different years for the same region
- o Curve (a) and (b) reflect the distributions in the same year for two different regions; this includes, for example, county compared with nation.

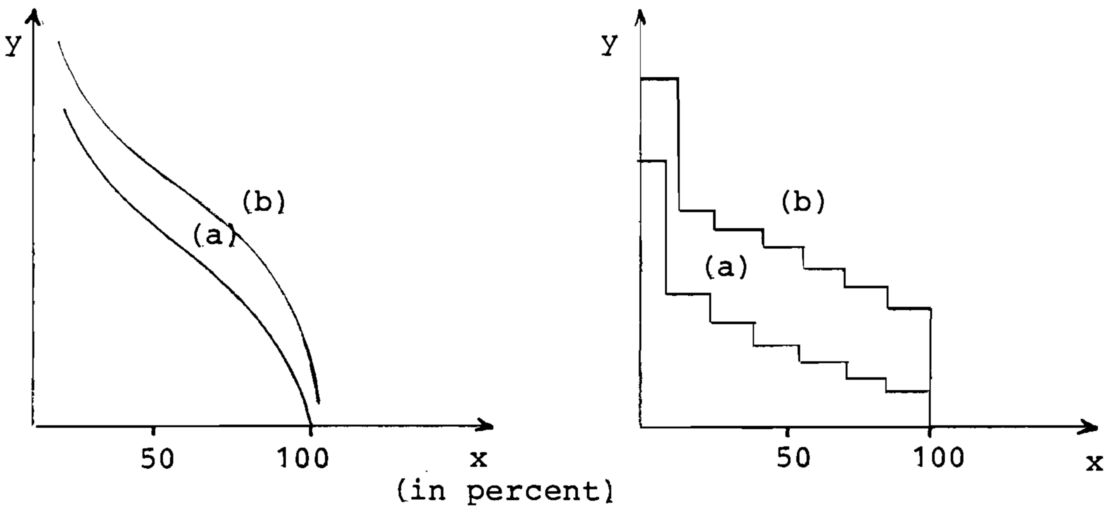


Figure 3. Illustration of the two types of diagrams.

In this way a series of comparisons may be obtained which makes it possible to separate fluctuations from structural changes.

For both information systems there exist several options to choose between as regards the selection of variables reflected by the y- and x-axes. With regard to the SIND-data system the following alternatives are used frequently:

y = productivity (value added per person employed)
y = profit margin (gross profit per sales value)
y = wage share (total labor costs per value added)[1)]
x = cumulative share in percent of (i) persons employed, (ii) value added, (iii) sales value in respective sector and region.

Separately, the absolute value of the x-variable may be obtained. The cumulative share scale has been chosen in order to facilitate comparison of structural properties.

The SUPPORT system contains a rich variety of options as regards the selection of y- and x-variables. In addition to the options in the SIND-data system, the variable y may for example, represent investment, rate of return on total

---

1) Observe that if y denotes wage share, then 1-y represents profit share.

capital and owners' capital, proportion of foreign capital, interest rate with regard to foreign capital, turnover rate etc., and the x-variable may also represent value of total capital, owners' capital, number of firms or persons employed, etc.

2.3 Production Theory and Productivity Measures

Since many features are similar for the two systems we shall concentrate on the SIND-data system, in which establishments are observation units. In this case it is straightforward to apply a vintage type of production theory.

The theoretical background may be summarized as follows. A production unit is characterized by its different types of durable resources which are formed into an organization embodying the operation technique (production, distribution, etc.). The technique represents a time invariant property of the production unit; it can only be changed by means of investment efforts. This also means that value added, at fixed prices, per number of persons employed (labor productivity) remains unchanged in the absence of investments.

In general the renewal investment in existing units is a slow process. This means that the productivity parameter $\bar{\mu}^{\tau}_{jr}$ will also change slowly; $\bar{\mu}^{\tau}_{jr}$ denotes the productivity of units in sector j and region r which are applying technique τ when the units operate at full capacity. However, the annually observed productivity μ^{τ}_{jr} will fluctuate due to variations in capacity utilization.

The structural invariance of the distribution of productivity over the employment in each sector and region has been examined by means of the following type of combined cross-section and time-series estimation:

$$z = [1 + \exp \{a_0 + a_1\mu + a_2t\}]^{-1} \qquad (1)$$

where $z = z(\mu,t)$ denotes the share of the number of persons employed in establishments with a productivity equal to or higher than μ, where t is a time index denoting years, a_0, a_1 and a_2 are estimated parameters. The logistic type of function in (1) gives an acceptable description of the actual distributions over the time period 1968-1978 for specific sectors and regions. When eliminating observations of extremely high values of μ[1)] and values of z such that $z \geq 0.9999$ $(0 \leq z \leq 1)$, it was possible for most sectors to obtain R^2 values ranging between 0.8 and 0.9 at the national level and between 0.6 and 0.85 with regard to industrial regions (Appendix 1). When evaluating the invariance property one should observe that three parameters for each z-function are generating ten different productivity curves of the type illustrated in Figure 4. Certain sectors display in some regions an invariance below the average. In those cases the growth rate of μ is unevenly distributed over different segments of the curve.

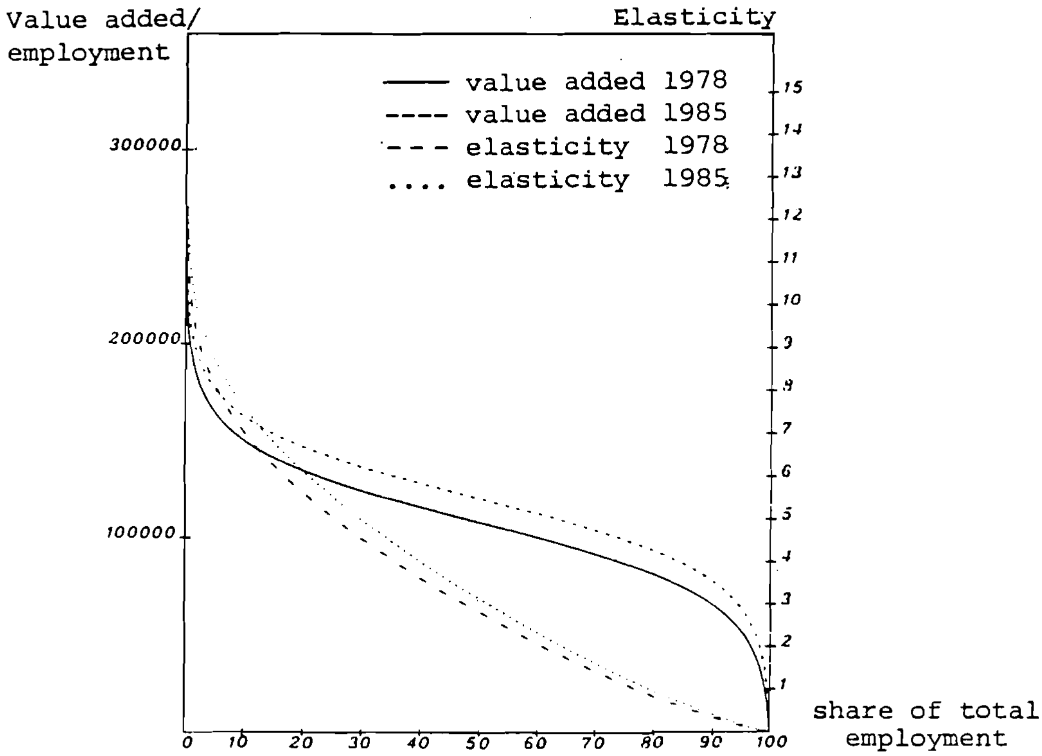


Figure 4. Continuous demand and elasticity function for the graphical industry in south Sweden 1968-1978 (1975 prices).

---

1) For most sectors values of μ exceeding 300 were eliminated. This corresponds to 300,000 in Figure 4.

2.4 Structural Invariance and Gross Profits

For each establishment we may write

$$\mu(t) = W(t)/S(t) + B(t)/S(t) \tag{2}$$

where W(t) denotes labor costs, B(t) gross profits and S(t) the number of persons employed. If $\mu(t)$ exhibits invariance then also the fixed price value of W(t) and B(t) should change slowly on the average. There are two reasons for examining the variance of B(t) and in particular B(t)/F(t) which denotes profit share, where $F(t) = \mu(t)S(t)$ denotes value added in an establishment. One reason is that one of the curves generated by the Info-system describes the distribution of profit share over value added. The other reason is that for the SIND-system, the model of the original data set is constructed by forming groups of establishments which have approximately the same profit share. If the share B/F is approximately invariant in the production units, then the units will have the same relative position on a given type of curve over time.

The degree of invariance for different profit levels displays the same pattern for the different subsectors of industries in Sweden. As illustrated for the whole manufacturing industry in Figure 5, the invariance is large at high and medium profit levels. When the profit is reduced and becomes negative, the invariance decreases. One reason for this is, of course, that when such units are not shut down but retained, then it is desirable for the owners to invest in new techniques so that the profit situation is improved. The ratio B/F is characterized in Figure 5 as approximately invariant if it varies less than ±10 percent between consecutive years.

The examination of invariance properties of the information which can be processed through the SIND-data system is to a large extent based on a complete transition analysis showing how productivity and profit share changes for units between years (see e.g. Johansson, Holmberg, 1982).

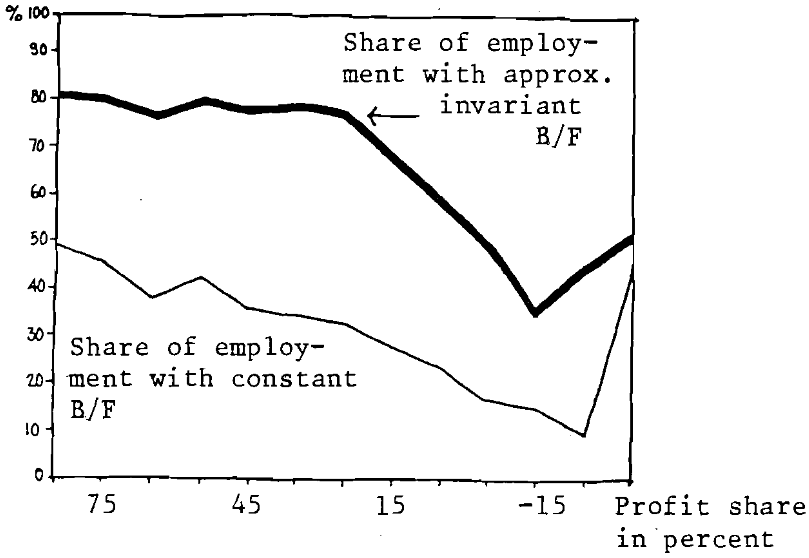


Figure 5a. Share of persons employed in units for which B/F varies less than ±10 percent between years. Manufacturing industry in Sweden 1968-1978. (Source: Strömqvist (1983))

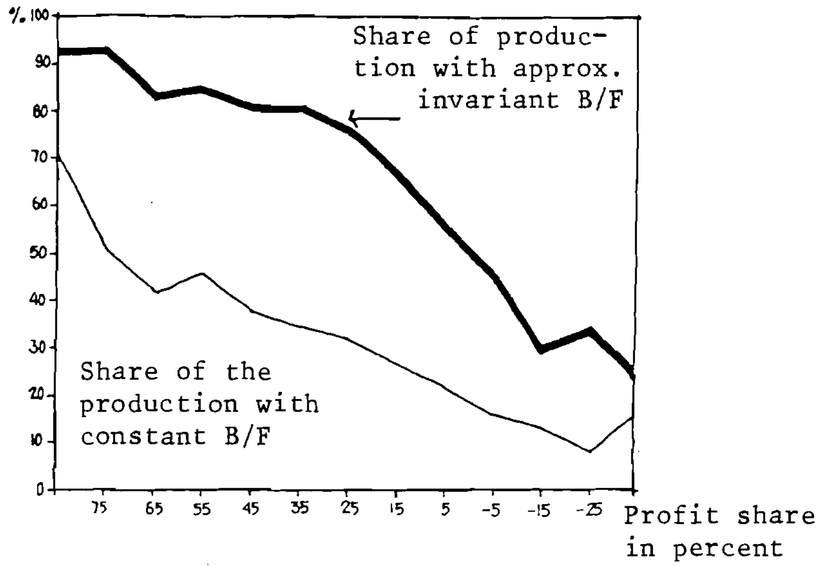


Figure 5b. Proportion of value added in units for which B/F varies less than ±10 percent between years. Manufacturing industry in Sweden 1968-1978. (Source: Strömqvist (1983))

2.5 Interactive and Non-Interactive Parts of the Systems

Both the SIND-data and the SUPPORT systems have (i) an interactive part with a high access dialogue system and (ii) a non-interactive part from which information is obtained on specific request. For this second part there is no dialogue system and the access is low.

The interactive part of each system produces diagrams and tables of the kind described in the preceding sections. The outputs from the interactive systems can be evaluated and interpreted with the help of associated models. Besides the type of invariance measures described earlier, there is information about the intensity of exit and entry processes, the probability of a future loss (negative profit) given the current productivity and profit share. With regard to the probability of exit, i.e., shut down of establishments, there exist estimated exit functions with the following exponential form

$$\xi = \alpha \exp \{\beta W/F\} \tag{3}$$

where ξ denotes the annual probability that a unit with the wage share W/F shall be shut down; α and β are estimated, positive parameters. Observe that according to (3) the frequency of shut down is positive also for units with positive profits.

Figure 6 describes the probability of transition from a positive profit level to a situation of loss. The described curve refers to establishments in the industry as a whole.

Figure 7 gives an overview of the SIND-data system. Most of the features are also valid for the SUPPORT-system which however contains many more options than SIND-data.

The output from the interactive part of the system is for example used

- o to detect and evaluate structural change in regions and sectors,
- o to predict the likelihood of shut down in a 2-5 year perspective,
- o to project the regional impact of changes in price and wage levels.

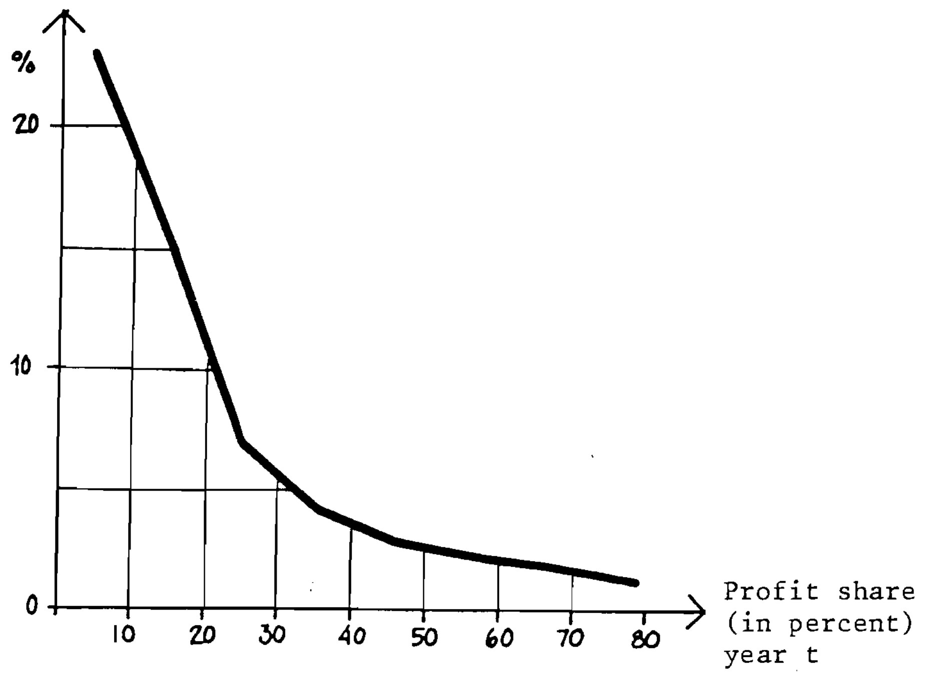


Figure 6. The probability that a person employed in a unit with positive profit year t will be employed in a unit with negative profits year t+1.

This means that tables and diagrams are interpreted with the help of associated models of the kind described earlier.

From the non-interactive part of the Info-system one may describe certain input coefficients (production technique) of establishment groups for each sector in a region. This information is extracted for each group of units from data on (i) labor force and labor costs, (ii) use of oil products (cost and quantity), (iii) use of electricity (cost and quantity), (iv) other input costs. Together with information about production these data are used to assess and forecast the change of production techniques for different segments of a productivity curve.

Some of the models in part A of the system have been described in English. Programming models are described in Johansson, Strömqvist (1981), Johansson (1983). The input-output model system MACROINVEST is described in Persson, Johansson (1982).

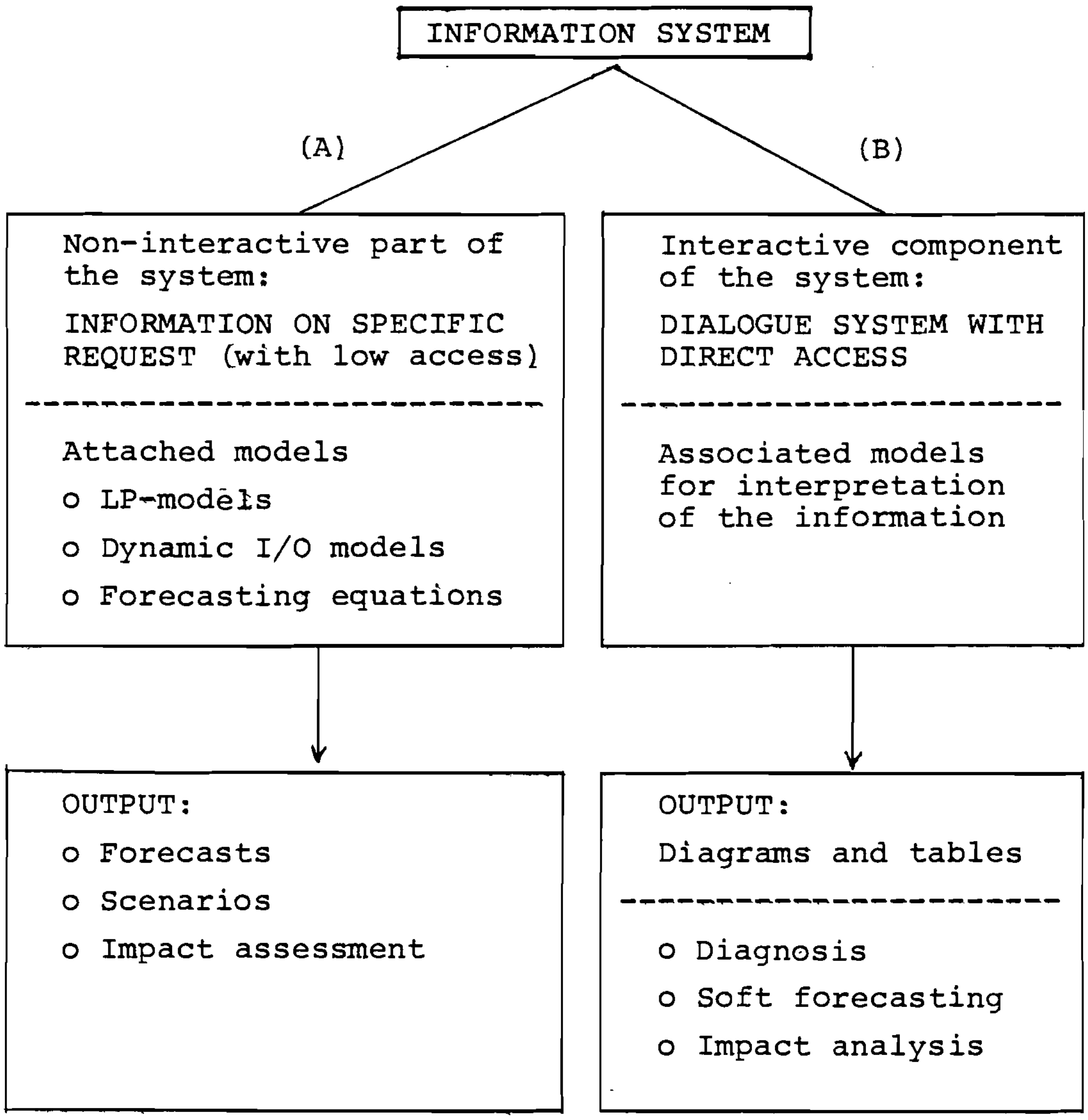


Figure 7. Overview of the SIND-data system.

3. DESIGN CRITERIA AND THE CONFIDENTIALITY RULE

3.1 Confidentiality Rules and Types of Aggregation

Confidentiality is a broad concept that can be defined in various ways depending on the context. For statistical systems, however, it denotes a principle of not exposing any observation unit to the public. The principle is applicable for all features of a unit which during the data collection phase are guaranteed to remain private. In the case of profitability statistics the units of observation are production units (establishments) or alternatively business units (firms).

In Sweden this type of statistic is not allowed to be used in its original form outside the government agency "Statistics of Sweden".[1] If the observation units are "sufficiently aggregated" into observation groups, the resulting statistical system can be used without constraints outside the agency, and one is allowed to process and disseminate information from it. The criterion "sufficiently aggregated" can be deduced from the confidentiality rule which says:

> The accumulative sales value of a permissible group should be more than twice the sales value of any observation unit within that group.

Hence, from this rule one can conclude that a permissible group must contain at least three observation units. This statement follows from the fact that in a group of two units, the largest unit would contribute by at least 50 percent of the sales value in the group, and this makes the group not permissible.

Groups may be formed by means of optimization techniques provided that it is possible to rank different groupings on the same scale. Following this approach, we have introduced an objective which requires that the information lost during the aggregation procedure should be minimal, subject to the confidentiality rule and certain homogeneity constraints on the groups. The information loss concept will be developed in the next section, while the homogeneity constraints are discussed below.

---

1) Official Statistics of Sweden, National Central Bureau of Statistics.

The SIND-data system utilizes a less complicated model of the original data set than the SUPPORT system. Therefore, we shall use the properties of SIND-data to shed light on the general aspects of reshaping the original data set to a non-confidential register. One of the purposes of the SIND-data system is to study profitability. Because of this, it is desirable to make the groups in the register as homogeneous as possible with respect to profitability. This means that the profit share, B/F, of each production unit in a group should be approximately the same.

Alternative principles of group homogeneity could be formulated with regard to the number of persons employed, amount of capital, degree of export orientation etc. Measuring homogeneity in several dimensions makes the process of constructing groups more complex, since then it becomes necessary to evaluate the trade-off between the different aspects. In the SUPPORT system one type of register is constructed for each type of homogeneity dimension. Those registers can be used to process combined information. The SIND-data system utilizes only the profit share variable as the homogeneity dimension.

In SIND-data establishments are observation units, and each such unit has a given location. In such a case it is convenient to form a separate register for each combination of region, economic sector and year. In the sequel we shall describe how a single register of this type is obtained. Hence, the presentation concentrates on an aggregation scheme for the one-variable case.

The aggregation procedure can be visualized as a manufacturing process, producing permissible (non-confidential) groups with observation units as inputs. In order to obtain homogeneous groups with regard to a certain variable, in this case profitability, we sort the observation units along a unit line.

In the assembly phase one unit at a time is moved from the front of the unit line to the end of an assembly line. As soon as the units in the assembly line fulfill the confidentiality rule, they are appended as a new group in a group line.

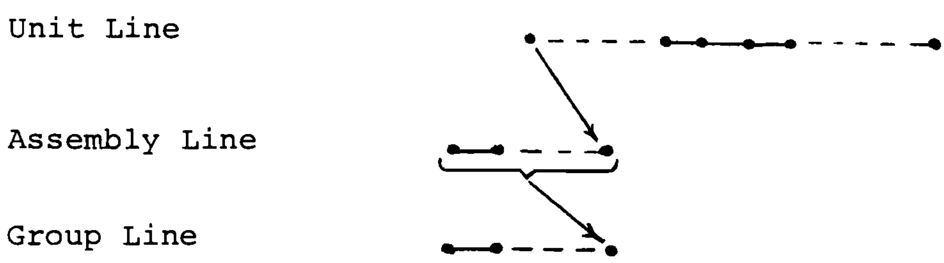


Figure 8. Illustration of the ASSEMBLY PHASE

The process is repeated until the unit line is empty. Figure 8 illustrates the aggregation process.

Suppose now that the assembly phase has ended in the sense that no more groups can be obtained. At this stage it may happen that the assembly line is not empty. The reason for this will then be that the remaining units do not satisfy the confidentiality rule. In such a case a disassembly phase has to be started as described in Figure 9. Then the last group on the group line is moved to the assembly line and is disassembled into its units. This process is repeated step-by-step until the units on the assembly line form a group fulfilling the confidentiality rule.

In the degenerate case the group line becomes empty during the disassembling process. In such a case it is not possible to extract any information about even one single group. In all non-degenerate cases one ends up with a group line which then consists of units aggregated to a sequence of groups.

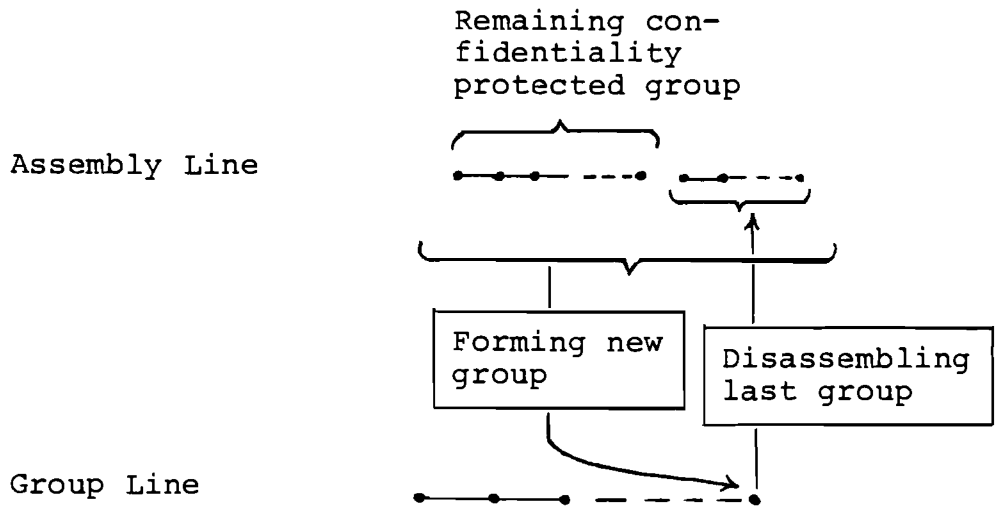


Figure 9. Illustration of the DISASSEMBLY PHASE

3.2 A Minimum Information Loss Criterion

In the filtering process described in the previous section we obtain a set of groups representing the original set of units. We would like this process to have the following characteristics:

- The process should use a *weight* variable such that the group weight equals the sum of weights of the units within the group (value added in the SIND case)
- It should use a *confidentiality* variable for which the largest value within a group is less than half the sum over the group (sales value in the SIND case)
- It should use a *homogeneity* variable which varies as little as possible within each group (wage share in the SIND case)[1)]
- The distribution of weighted homogeneity over the groups should be as close as possible to the same distribution over units (wages/labor costs in the SIND case).

A common underlying principle can be used to promote the last two characteristics, whenever they do not conflict. The principle has its origin in statistical theory. A brief summary is given below.

Suppose that in a statistical experiment n different outcomes are possible. These events are numbered $j=1,2,\ldots,n$ and the probabilities, $p_j>0$, that they occur add up to 1. Let P denote this type of probability distribution. Kullback (1959) has defined a non-symmetrical measure of difference between two probability distributions P and Q. He defines the information gain of P over Q as

$$I(P;Q) = \sum_{j=1}^{n} p_j \ln (p_j/q_j) \tag{4}$$

Equivalently, we choose in this context to interpret $I(P;Q)$ as the information loss from Q to P, where Q is called the *a priori* distribution and P is called the *a posteriori* distribution.

---

1) Observe that if y=wage share, profit share=1-y.

In this application we would like to give the attribute "a priori" to the observation units and "a posteriori" to the groups which are formed during the filtering process.

Let W denote the *weight* variable and H denote the *homogeneity* variable. The number of units is called n and the number of groups is called m. If H is scaled properly, the a priori probabilities Q are given by

$$q_j = h_j w_j \quad \text{for} \quad j = 1,2,\ldots,n \tag{5}$$

From now on we require that h_j and $w_j > 0$ for all j. Let the groups be indexed by J such that J = 1,2,...,m where J, at the same time, denotes a set of unit indices $\{j_1, j_2, j_3 ..\}$, showing to which group a certain unit belongs. We assume that every unit belongs to some group, and that no unit belongs to more than one group. This type of grouping will be called *complete*.

The variables corresponding to H and W, but varying over the groups, are denoted by $\bar{H}$ and $\bar{W}$ respectively. Then by definition:

$$\bar{w}_J = \sum_{j \in J} w_j \quad \text{for } J = 1,2,..,m \tag{6}$$

We prescribe that

$$\bar{h}_J = \bar{h}_j \text{ when } j \in J \text{ for } J = 1,2,..,m \tag{7}$$

Then we can define the a posteriori probabilities P over the units by

$$p_j = \bar{h}_j w_j / \sum_{k=1}^{n} \bar{h}_k w_k \quad \text{for } j = 1,2,\ldots,n \tag{8}$$

Note that $\bar{h}_j \neq h_j$ in the general case.

Now the information loss criterion can be stated as follows:

> Among all *complete* groupings of units, find the one which minimizes I(P;Q) where Q is defined in (5)

and P is expressed in a short-hand form of (8) as

$$p_j = x_J w_j \quad \text{when } j \in J \quad \text{and} \quad J = 1,2,..,m \tag{8'}$$

such that

$$I(P;Q) = \sum_{j=1}^{n} x_J\, w_j \;\; \ln\,(x_J\, w_j / h_j\, w_j) \text{ where } j \in J$$

subject to the following simultaneous conditions

$$\text{(i) } x_J > 0 \text{ for } J = 1,2,..,m$$

$$\text{(ii) } \sum_{j=1}^{n} x_J\, w_j = 1 \text{ where } j \in J = 1,2,..,m$$

$$\text{(iii) } 2 \max_{j \in J} (s_j) < \sum_{j \in J} s_j \quad \text{for } J = 1,2,..,m$$

where S is the *confidentiality* variable.

Given a permissible *complete*grouping the Lagrange function L(x) of the problem is

$$L(x) = \sum_{J=1}^{m} \sum_{j \in J} x_J\, w_j \;\; \ln\,(x_J/h_j) + \lambda\,(1 - \sum_{J=1} \sum_{j \in J} x_J\, w_j) \tag{9}$$

The first order conditions for a minimum are $\partial L/\partial x_J = 0$ for $J = 1,2,..,m$. Differentiation yields

$$\sum_{j \in J} (w_j \;\; \ln\,(x_J/h_j) + w_j - \lambda w_j) = 0$$

Summation gives

$$\bar{w}_J \ln \; x_J + \bar{w}_J(1-\lambda) - \sum_{j \in J} w_j \ln \; h_j = 0$$

Hence

$$x_J = e^{\lambda-1} \prod_{j \in J} h_j^{w_j/\bar{w}_J} \tag{10}$$

We observe that $x_j > 0$ for $J = 1,2,..,m$ since $h_j > 0$ for $j = 1,2,...,n$.

The second other conditions are

$$\frac{\partial^2 L}{\partial x_K \partial x_J} = \frac{\bar{w}_J}{x_J} \delta_{JK} \quad \text{where } \delta_{JK} = 1 \text{ if } J = K$$

$$\text{and} \quad \delta_{JK} = 0 \text{ otherwise}$$

Since $x_J > 0$, a minimum is obtained because $\bar{w}_J > 0$.

APPENDIX 1: Combined Time-Series and Cross-Section Estimates of Productivity Distributions

Estimates of the productivity curves are obtained directly from the SIND-data system. The currently used estimation method differs somewhat from the attempts presented here. In the following table we describe two of the several estimation procedures which initially were utilized on the original data set. Results are presented for the 14 major industry sectors on the national level. In addition we give results for the industry as a whole with regard to the six industry regions.

- o Estimation alternative I:
  Let $\hat{z}$ denote the observed employment cummulative share variable. Observations where $\hat{z} \geq 0.9999$ were eliminated. A weighted least square approach was used minimizing $\Sigma\{h(\hat{z}-\bar{z})\}^2$ where $h = \hat{z}$ for $\hat{z}<0.5$ and $h = 1-\hat{z}$ for $\hat{z} \geq 0.5$.
- o Estimation alternative II:
  This alternative utilizes unweighted least square with elimination of $\hat{z} \geq 0.9999$ and $\mu \geq 300$.

Table 1. Measure of Estimation Accuracy.

Sectors on the National Level	Alternative I F-value	Alternative I R^2	Alternative II F-value	Alternative II R^2
Manufacturing industry (whole)	71387	0.91	35596	0.84
Mining industry	35	0.34	98	0.67
Food industry	3479	0.80	2746	0.77
Textile industry	7572	0.88	3914	0.79
Wood products	12371	0.87	14452	0.88
Paper and pulp industry	409	0.65	727	0.78
Printing and publishing industry	3994	0.87	5624	0.90
Chemical industry	957	0.79	1772	0.89
Stone and clay industry	2860	0.83	2193	0.79
Iron and steel industry	152	0.65	217	0.73
Metal industry	8401	0.85	9218	0.86
Machinery and equipment	4398	0.86	3645	0.83
Electrical industry	365	0.63	583	0.74
Transportation equipment	121	0.44	558	0.78
Manufacturing industry in				
East Sweden	12701	0.90		
Southeast Sweden	12436	0.89		
South Sweden	4383	0.82		
West Sweden	8045	0.84		
Middle Sweden	4470	0.83		
North Sweden	2408	0.82		

Remark: All parameters have significant t-values.

APPENDIX 2: Conversation Menu

The implementation of the database, dialogue access and graphical presentation of results has been made in APL for both the SIND and the SUPPORT system.

The dialogue is especially designed for asynchronous terminals using low speed dial-up connections. The design tries to minimize the number of characters transmitted between the terminal and the computer. Database access is performed via a dialogue session which can encompass several passes through the questioning system.

During each pass a lot of table forms,predefined charts or data variables for user specified analysis can be ordered at the same time. All tables or charts are collected until the end of the session and saved. They are later retrieved in a presentation session, not to be dealt further with in this context.

The dialogue is menu-oriented and is easy to use for the non-experienced user. The expert, on the other hand, can run through a session very fast. At any point in a dialogue pass he can suppress a number of following question texts by typing the corresponding answers at once, separated by a certain delimiter.

Questions on branch, region and time allow for several alternative answers. A pass through the dialogue results in all combinations of the alternatives chosen, thereby reducing the typing effort considerably.

Abbreviations can be made on the response alternatives, which lessens the typing work further. If the user shortens down an answer to ambiguity, then the system only presents the alternatives matching the response and asks for one of them. Single alternatives can consist of several words separated by blanks. Abbreviations are also allowed within single words.

If the user responds with an empty answer, the actual dialogue pass is cancelled and a new one is started. The earlier passes are not affected, however. The empty response to the first question in a dialogue pass will end the session.

The presentation of the data accessed as described above can be flexibly made. For instance it is possible to choose between observed and estimated data. Text could be shown in different languages, and plotting be done on different media.

A short excerpt of a dialogue is shown below, where the computer writes in upper case letters and the user writes in lower case.

```
.
.
.
.
SPECIFY REGION:  any
ANY:  PLEASE CHOOSE BETWEEN:
EAST SWEDEN             MID SWEDEN
NORTHERN SWEDEN         SOUTH EAST SWEDEN
SOUTH SWEDEN            WEST SWEDEN

SPECIFY REGION:  s
s:  PLEASE CHOOSE BETWEEN
SOUTH EAST SWEDEN       SOUTH SWEDEN

SPECIFY REGION:  s e
        (South East Sweden is picked out)
```

A more advanced example is demonstrated below.

.
.
.
.
.

SPECIFY REGION: e, s_w
(First, e picks East Sweden)
(Then, s_w was intended to pick out some South part of Sweden and proceed up to West Sweden)

S: PLEASE CHOOSE BETWEEN:
SOUTH EAST SWEDEN   SOUTH SWEDEN

SPECIFIY REGION: s s   (Refers only to the ambiguous part)
(East, South and West Sweden are picked out as a correct response to the first question)

REFERENCES

Johansson, B. and I. Holmberg (1982), A Regional Study of the Distribution of Vintages and Profits of Industrial Establishments: A Stochastic Transition Model, in Regional Development Modeling: Theory and Practice, M. Albegov, Å.E. Andersson and F. Snickars (eds), North-Holland.

Johansson, B. and U. Strömqvist (1981), Rigidities in the Process of Structural Economic Change, Regional Science and Urban Economics, 11.

Kullback, S. (1959), Information Theory and Statistics, John Wiley, New York.

Marschak, J. (1974), Economic Information, Decision and Prediction, selected Essays vol. 2, D. Reidel Publishing Company.

Nijkamp, P. (1982), Spatial Information Systems - A Research Strategy, IIASA CP-82-35, International Institute for Applied Systems Analysis, Laxenburg, Austria.

Persson, H. and B. Johansson (1982), A Dynamic Multisector Model with Endogenous Formation of Capacities and Equilibrium Prices: An Application to the Swedish Economy, IIASA-PP-82-9, International Institute for Applied Systems Analysis, Laxenburg, Austria.

Strömqvist, U. (1983), An Economic Analysis of Agriculture and Industry in the Skåne Region, IIASA CP-83-15, International Institute for Applied Systems Analysis, Laxenburg, Austria.

Strömqvist, U. (1983), Lönsamhetsstruktur och Investeringsmönster, (Profit Structure and Investment Pattern), SIND 1983:1, Statens Industriverk, Liber, Stockholm, Sweden.