SYSTEMATIZING SOCIAL
INFORMATION IN ARGENTINA

LUIS BECCARIA
National Institute of Statistics
and Censuses, Argentina

INTRODUCTION

It is possible to recognize that, since the sixties, welfare
analysis has tended to show an increasing preoccupation with
social aspects and to realize that economic development is a
necessary but not a sufficient condition for improving well
being and for alleviating social problems. This view implies that
the welfare situation of a country cannot be properly evaluated
by resorting exclusively to economic information.

This view is particularly relevant in developing countries,
where economic and social inequality is a common feature and
where a major portion of the total population is unable to
satisfy its basic needs. In such cases, it has been largely
recognized that it is inadequate to analyze development
performance by considering the aggregate results concerning the
evolution of domestic production. It seems necessary, therefore,
to study how the achievements were distributed among the
different population groups and how such dimensions as health
and education are affected. This broader view was deemed
necessary in order to evaluate not only the overall development
strategy but also particular investment projects.

The availability of social information -both in quantity and
in quality- does not usually meet the requirements implied by
this orientation. Those interested in analyzing social questions
and in designing social policies are likewise affected by lack of
adequate information. Resources from the public sector and
efforts by the international statistical community were mostly
directed towards improving economic data. The inherently more
difficult questions posed by the production of social information
are another factor explaining this relative lagging behind of
social data.

Against this background -and spurred by these reasons- there
is a noticeably increasing preoccupation of the national
statistical systems in many countries and of various
international agencies with the production of social information.

1. THE NEED FOR INTEGRATING SOCIAL STATISTICS

1.1.The state of social statistics

In a number of countries social data are of poor quality and
cover only a few areas, specially those which are by-products of
administrative activities. In general, they are seriously flawed
by deficient coverage and inadequate accuracy and,moreover they
are usually available with important delays.

In some natinal statistical systems,it is the supply of
data, rather than their demand, that determines the work pattern.
In many instances, moreover, efforts to improve social data (and
this also applies to economic ones) reflect a certain

independence of the statistical agencies from the policy makers.

Sometimes, when specific demands for social data arise, they are usually restricted to a relatively narrow range of variables. Moreover, these demands are satisfied mostly by means of ad hoc, short-lived, mechanisms, unrelated to the official statistical machinery.

In many cases, the analysis of basically interrelated questions -such as social ones- is carried out by using information which refers just to only a given dimension. When the need to take into account several aspects is obvious, specific and costly statistical endeavours are required in order to asses the different varibles influencing a given phenomenon. Isolation and inefficiency are the usual consequences of such situations.

1.2. Alternative proposals

Given the increasing need for social data and their deficient state in many countries, several proposals have been advanced in order to tackle such lack of correspondece between supply and demand. Two different approaches can be identified in this work on social statistics: one group of researchers tries to respond to the restrictive nature of the economic aggregates by considering other dimensions unaccounted therein. In this first group two streams can be distinguished, one aimed at defining a single synthetic indicator capable of measuring the social welfare of a society more adequately than the GNP or related national accounts aggregates; the other stream includes those approaches tending to define a 'battery' of social indicators complementary to the economic data.

The first of these two streams tries to tackle the issue in the most straightforward manner: the GNP is 'incomplete' hence it is necessary to add other elements. In general, the proposed strategy aims at valuing the time spent in non-earning activities and adding it to the GNP; sometimes, certain elements which are viewed as 'costs' to the society are deducted from the national income. Within this group it is possible to include those who attempt to derive an aggregate by combining several social indicators (as in the UNRISD's "level of living index"). The second stream of the first group comprises those who propose a set of indicators covering various areas of social concern, such as the OECD programme on social indicators. In many cases, the indicators are selected by previously defining a set of goals or by deciding which series can be of help in monitoring the 'quality of life'.

However, the type of data resulting from both of these approaches cannot be satisfactorily used to explain either the social structure or its changes. Its chief drawback is that none of the proposals are presented within a framework highlighting the relations between the different indicators. Thus, not only the usefulness of these approaches for analytical purposes (as opposed to 'descriptive' ones) is limited but they can also

originate a somewhat ad hoc selection of indicators.

The second group of proposals which may be distinguished in the research on social data is based on the notion that it is necessary not only to define social indicators but also to discern how they interact, both among themselves and with other variables which are not usually regarded as proper 'social indicators'. In other words, in this set of proposals a framework is envisaged which provides information suitable not only to the description of a phenomenon but also to its explanation. The obvious difficulty met with by those attempting this approach is that, even if theoretical hypotheses about particular areas do exist, there is no general 'macrotheory' which can be readily used as a basis for integrating those different fields.

Problably, the most interesting concept is that of the United Nations System of Social and Demographic Statistics (SSDS) (UNSO,1975). It is not based on any specific theoretical premise or unifying notion, but its specific objective is to provide relevant guidelines for gathering information that can be integrated. Basically, these guidelines aim at harmonizing the data collected by many agencies and at providing links between different areas of social interest.

1.3. Systematizing social and demographic statistics

From the point of view of designing a nation's statistical policy, the proposal of integrating the various efforts devoted to the compilation of social statistics seems to be a convenient guideline. It is based on the idea that the statistical apparatus should produce the information necessary not only to describe a given situation but also to analyze and to explain it.

The view taken here is that the objective of organizing social statistics can be satified by using a blueprint that: i) does not impose make any a priori restriction concerning the type of information and seeks data needed to 'explain' besides those necessary to 'describe'; ii) asks for unified classifications, concepts and definitions; and iii) is able to connect the different areas of social conern. These characteristics are largely the same as those of the UN's SSDS.

The first requirement clearly stems from the restrictive character of certain approaches discussed earlier; the second introduces the necessary requirements of systematic data collection. The third implies that the usual practice of considering data on different fields in isolated fashion will have to be abandoned. These three requirements are interconnected: links between areas would be very difficult to establish if there were no intention to harmonize the concepts and definitions used in each. Also, to explain a situation in a given field would be very difficult if variables pertaining to other areas were not considered.

Consequently, as it was pointed out by Zagorski (1975), "the

systmem of socio-demographic statistics (is...) a set of statistical data elaborated in such a form and by means of such methods as to present actual state or process integrated as mutual interrelations between different influences of the total number of individuals, groups or institutions."

The lack of a 'conceptual core' in the SSDS might be criticized because it is a handicap against fulfilling the first above mentioned requirement and, to a certain extent, the third one. Indeed, the lack of conceptual basis seems to prevent any meaningful selection of variables and areas to be connected. However, even if no general,or macro, theory is available, there is a large volume of accumulated knowledge concerning the different areas. In fact, it is asserted here that the type of data required in each field, and hence the way in which they are related, must be the result of analyzing the theoretical hypotheses that have been advanced in order to understand the different areas. The identification of the necessary information should derive from the analysis of existing theories. This selection should thus take into account alternative views of the different phenomena. Again to quote Zagorski (1975), "systematization of statistics should, therefore, be based on systematization and structuralization of the object of research (i.e. society) by investigators". Summing up, the SSDS should be seen not as a list of variables, such as the System of National Accounts, but as a set of recommendations on an efficient organization for the collection and production of information. It is now commonplace to quote Moser and to refer to "systematization" instead of "system".

It goes without saying that not every need can be satisfied, but the present approach offers a strategy for data gathering which is flexible enough to satisfy most of the demands that can be expected. This flexibility is important from another point of view: the requirements of information arising from the analysis of theories will not always suffice to define the set of actual variables to be measured; it should also take into account government demands. But since the objectives of a government may change relatively quickly, its requirements cannot be the sole guide to data collection. The ideal course to be followed when organizing data collection starts from a comprehensive, and ambitious, set of requirements. They should not be fixed but they should undergo constant revisions, constituting a sort of permanent yet mobile target with every effort striving towards its fulfillment. Such a target is not an absolute 'ideal', for it must take into account restrictions likely to be faced by any data collection operation, whether they be financial or administrative. Government requirements should provide the order of priorities and should add precision to the definition of certain variables.

2. THE STRATEGY FOLLOWED IN ARGENTINA

2.1. The state of social information in Argentina

4

Social data in Argentina share many of the traits described
in section 1.1. As in many other countries, an important portion
of basic data is obtained from administrative records; hence, the
usefulness of the information which could be derived from them
depends -to a certain extent- upon the peculiarities of
administrative procedures. These are often deficient and
originate data lacking in accuracy and inadequate both in
coverage and timeliness.

Moreover, the production of social data is the
responsibility of different agencies, working separately. This
situation generally implies effort duplication since a given
phenomenon can be independently measured by more than one agency.
For example, there are several different wages series compiled
for the manufacturing -and other- sectors. The Ministry of
Labour, the Social Security Department as well as the Statistical
Institute provide, not always comparable data about the same
variable.

There is, however, another factor compounding the
inefficiency of the statistical efforts,i.e., the lack of
harmonization and adequacy of definitions, coverage and
classifications used in administrative procedures and in various
statistical activities, when they are developed in an isolated
manner. The usefulness of the information derived from these
sources is substantially reduced because data collecting
instruments, e.g. questionnaires, records, etc.- are
unsystematic and because the definitions and classifications
employed by different agencies are not homogeneous even within
the sphere of the Federal Government.

The production of data by isolated units or agencies may
also help to explain why those working in a given sector -say
education- have no contact with -and consequently do not use-
information that could be helpful to them but is produced by an
agency in a different sector -say health.

The fact that Argentina is a federal republic and the
political and administrative implications of this type of
constitutional organization should be taken into account when
working out a framework for systematizing social information. In
fact, Provincial governments provide, fully or partially, a
variety of social services and play a significant role in the
design and/or execution of certain social policies, for instance,
in the fields of education and health. The Provincial governments
have their own statistical agencies. Although in many cases
operative agreements have been formally established between the
provinces and the Federal Government in order to coordinate their
policies and activities, the existence of twenty-four different
political units and authorities tends to complicate the efforts
aimed at coordinating the production of statistical data.

Fortunately, there is in use in Argentina a statistical
instrument contributing to an integrated analysis of several

social problems -the permanent household survey established precisely with that stated objective in 1974. It is carried out twice a year in the main urban centres of the country. Basically, it inquires about employment, income, education, housing and household composition. Obviously, the regular population censuses are also important sources for the integrated analysis of different aspects.

2.2. The work programme

Against this background of isolation that marks the production of social data by different agencies, a programme aimed at improving their usefulness was designed, based on the ideas of :(i) integrating the information from different areas of interest through the use of common definitions and classifications; (ii) improving the administrative procedures (where personnel training becomes critically important); and, (iii) adequating statistical activities (censuses, surveys and registres) to make them functional to the overall strategy.

Actually, this is not a "once and forever" project in the sense that it will establish a definitive situation diagnosis, will then submit a set of final recommendations that it will then proceed to implement without further questions. On the contrary, the idea is to set forth a scheme allowing continuous adjustment efforts towards the systematization of social data. In fact, such a scheme is now in formal existence because every data producer has been integrated into the National Statistical System so, that their activities should be coordinated by INDEC, the National Institute of Statistics and Censuses. In practice, however, this goal has not yet been fully attained and additional efforts are required to that end. The programme has two chief objectives – on the one hand, to implement a set of recommendations based on a diagnosis of the current situation; on the other hand, to strengthen the coordination mechanisms in order to cope with new demands and to ensures the smooth functioning of the system.

Three stages can be distinguished in the development of this programme, namely

> (a) identification of data produced in different areas of the public sector.

> (b) situation diagnosis concerning the availability and quality of the existing information and the demand for data.

> (c) designing the system

To achieve stage (a), a survey of different agencies was carried out to inquire about data produced (either published or not), data collection instruments, definitions and classifications used. This information on information -which is being orderly maintained (see next section)- constitutes the

6

basic element for designing the programme; it points out operation and effort duplications within the system and allows the detection of variables not previously measured and of discrepancies in definitions and classifications.

Furthermore, the completion of this stage makes it possible to gather available data and present them in a systematic way. This is an important feature in the process of ending the isolation which has marked the work of many potential data users. It has also been considered that providing them with better information will promote the widespread use of statistics throughout the administration and improve their contribution to defining information needs. Moreover, this could help some users realize that it is necessary to relate some of the variables in their areas of work to those in other fields.

A major feature in the proposed strategy is to identify as clearly as possible the demand for social data from both the public and private sectors. As it was stated above, the type of data produced reflects in many cases either the views and/or criteria of the statistical agencies or the peculiarities of the supply of information derived from administrative procedures weighted with the inertia from the past. The user's views and needs are not always properly considered. Besides, no formal mechanism exists through which such requirements might be introduced in the formulation of statistical policy. By contrasting the demand for information to its present supply (from phase (a) above) a diagnosis of the situation can be made in which not only the availability of data but also their quality (including adequacy of definitions and classifications), coverage and timeliness should be taken into account.

Based on this diagnosis, a set of proposals will be put forth consisting essentially of a number of articulated recommendations aimed at reducing the gap between data supply and demand. The proposals should define the most efficient means of obtaining the information needed in the different areas of concern. This has to be done taking into account not only the activities -either statistical or administrative- of the agencies in a given area but also those performed by units in other fields. In particular, it should be evaluated to what extent administrative procedures (currently in practice or new ones to be introduced) and statistical activities (censuses, surveys , registres) will be the sources of basic data. The latter may be specific to a particular sector (such as school or hospital surveys) or they may be of a general nature. In this last case, information related to one field represents only a portion of the total data requirement (as in population and household surveys).

A major aspect of this programme will be the setting of priorities, which in many countries has proved to be a source of conflict, since the statistical system is hardly able to satisfy every demand at the same time. Regarding this point, the situation is particularly difficult in a country like Argentina, with its decentralized statistical system and with Provincial

governments following their specific interests and having their
own statistical projects.

One must bear in mind that the dimensions considered when
designing the programme were not only those typical of the so-
called "areas of social concern" in the SSDS, such as health and
education. Actually, the situation analysis of certain population
groups throughout all such areas can also be relevant to many
users. Therefore, when establishing priorities both a sector
perspective and a population group outlook should be considered.
The latter is particularly relevant in Argentina, where social
policy has identified (and this trend will probably increase in
the future) certain target groups. In some cases, a well designed
statistical system can satisfy both perspectives. For instance,
if young people or women are selected as target groups, it might
suffice to classify the principal variables according to age and
sex. In other cases, when target groups are not so easily
defined, the strategy should perhaps strive to identify specific
data sources for those groups. The poor are a good case in point
since the situation of underprivileged sectors has been a source
of increasing concern in Argentina over the last few years.

In fact, the need for information on the poor usually
originates in a policy aimed at reducing economic and social
inequalities. In that framework, an essential classification
scheme is based on socioeconomic and/or income strata, since the
diagnosis of welfare distribution and the evaluation of the
differential effects of certain policies are of primary concern
to governments.

Hence, a basic point in the strategy is to provide
information enabling us to assess the varying incidences of a
given problem -say morbidity, school dropouts or housing
deficiencies- among different population groups. Ascertaining the
unequal access to certain services -e.g., waer supply, health
care, etc- by those different population groups is another
objective to be pursued. A similar importance should be assigned
to those data required to evaluate the amount and type of
resources that society as a whole -both public and private
sectors- should devote to different areas.

Of similar importance should be those data required to
evaluate the amount and type of resources the society -both,the
public and the private sector- devote to different areas.

It is almost needless to emphasize that a main feature of
the programme will be the use, to the widest possible extent, of
common classifications and definitions. In certain cases, a
specific classification becomes "common" because of analytical
reasons; in others, specific policy-oriented reasons are the
determining factors e.g., when reducing inequalities is a major
objective.

The basic notion underlying the proposals is to resort as
much as possible to current administrative procedures. Although

admittedly these are beset by a number of problems, to invest in their improvement will probably be a more efficient alternative than to undertake new ad **hoc** statistical inquiries.

As regards this data source, two different steps should be considered, even if sometimes they can be carried out simultaneously. The first one is to improve the administrative procedures. It has been repeatedly asserted that many of the problems presented by such data derive from the inadequacy of the administrative machinery and not necessarily from the type of instrument used to gather the information. Many examples of this kind have been found in Argentina. In the field of education, for instance, data from the primary and secondary school levels are obtained with considerable delay because of inadequate communications between different agencies. The basic data coming from each school are summarized at the provincial level by the statistical division of each provincial educational authority, and these summaries are forwarded to the statistical service of the Federal Ministry of Education. However, there are "noises" in the information flow between these three levels, mostly processing and/or transmission delays as well as errors in producing the information.

In order to solve this sort of problems and other similar ones affecting data originated in the administrative apparatus, three main paths should be followed: improving personnel training, introducing modern technology and increasing economic resources. With respect to training, it should be kept in mind that such personnel does not necessarily belong to the statistical system. Lack of adequate trainning programmes for personnel filling administrative returns partially explains the shortcomings in the data they provide. For instance, intensive training courses should be offered to "school secretaries", who are responsible for producing the statistical reports discussed above. Similar courses should also be offered to civil registres officers.

In regard to the introduction of modern technology, neither sophisticated nor grandiose plans are envisaged; suffice it to say that microcomputing is slowly being adopted in the Argentinian public sector. The impact of this relatively unexpensive technology on the improvement of a variety of administrative procedures is well known.

Since the best way to integrate the analysis of different areas of concern is at the level of microdata, i.e.,to measure variables in several of these fields for the same persons or other relevant analytical units (such as households), an important point in the programme will be to enlarge certain administrative registres concerning those units by introducing information about at least a (necessarily reduced) number of basic characteristics. This proposal should be assessed against the actual possibilities of obtaining adequate data. In this respect, a thorough knowledge of the administrative procedures is required and, particularly, of the abilities of those responsible

for filling out the returns.

2.3. Description of activities already developed

The bulk of the work initiated so far is the systematic gathering of information on data produced by different agencies in the public sector. Metadata and the information itself are being systematized by means of METIS, which allows for the orderly preservation of information about different data collection instruments (designated as "basic information collection instruments") such as definitions and classifications used, coverage (geographical or otherwise), reference periods, periodicity, type of registres where the information on each basic unit is kept, etc. Data on many of these dimensions from different instruments are filed together thus making it possible to compare , for example, the definitions of a given variable, the classifications of a given atribute, the geographical coverage, and so on, used in different statistical inquiries or administrative procedures. For the sake of illustration, by resorting to the file on "definitions" the various definitions used for "household" can be reviewed. Similarly, by referring to the file on "classifications" it is possible also to find out where the persons employed are classified according to the type of industry of the establishments where they work. Furthermore, information is provided about the type of classification being used in each case. The operation of this system has confirmed that several classifications are currently in use in Argentina: some agencies use the ISIC first revision, others use the second revisions and a number of them resort to ad hoc classifications.

In building up this metainformation stock, an important aspect was to identify all the agencies producing data which are being —or could be- used for statistical purposes. As a result, it is now clear that there is more information than had been expected, albeit much of it of poor quality and/or irregularly produced. A disturbing factor is the number of specific statistical inquiries, which not only demand considerable human and financial resources but are also directed to units (households or others) which are already being surveyed by other agencies. The duplication of efforts, leading to overall inefficiency, and the negative impact on many respondents are potentially significant drawbacks. In many cases, neither the National Institute of Statistics and Censuses nor the statistical departments of each involved sector (or province) knew that those agencies were carrying out statistical activities.

Also, METIS allows the information itself to be filed, thus assuring easy access to it. This will facilitate its dissemination with the beneficial effects mentioned in the previous section. To increase the usefulness of those data a set of indicators was measured. Hence, the files contain raw information and indicators. In fact, the number of indicators

defined is larger than that of those actually measured; in other words, certain indicators cannot yet be measured with the information available. This proved to be a helpful instrument for the discussion of data characteristics and priorities with both users and producers of information.

Social indicators and data will also be disseminated through a forthcoming regular publication .

Initially, the organization of metadata was undertaken for information of Federal agencies. Eventually, a new project was designed -which received some assistance from UNICEF- in order to analyze the production of social data at the provincial level. Even though this project is restricted to information about children and young people, it paved the way to an important exercise: the meeting of data users and producers from every Argentininan province 1/. As a result of the work previous tothose meetings as well as that carried out thereafter, information coming from provincial agencies is now being incorporated to METIS.

In regard to the formulation of a diagnosis and a set of proposals, work is about to begin especially in three fields of concern -health, education and family (which includes such matters as the situation of children). The choice of areas stems from government priorities and also from the availability of resources to these sectors.

A redefinition of the data collecting instruments used by the health and educational systems is sought, as well as a definition of variables that should be included in the permanent household survey and/or the population censuses. Since the project includes the implementation stage, mechanisms must be designed and tested in order to improve the administrative machinery and the capacity of the sectorial statistical departments to cope with the operation of the new scheme. For the reasons already stated a major feature of the project points towards improving the relationships between the Federal and Provincial agencies in charge of gathering, processing, analyzing and disseminating information.

Another important trait of the overall programme is the redefinition of the permanent household survey, a project which is being carried out jointly with the preparation of the 1990 population census. An even closer relationship between both statistical sources is sought, as well as the enlargement of the survey's geographical coverage. In regard to the subject matter dealt with in this instrument, it appears that in theory an agreement will be reached only after the subsystems of the different areas of social concern are defined. It is not easy, however, to comply with this approach because: (1) the household survey is already the obvious source for important variables in
_____
1/ Several regional meetings were held,with three to five provinces attending such meetings.

11

Argentina (such as unemployment and non-wages incomes); (2) a
redefinition of the survey would be needed before any specific
programme for data requirements could by agreed upon inmany
areas. Still, it is not an unsurmountable problem because the
survey is quite a flexible instrument which can readily
incorporate new items to the basic questionnaire.

3.CONCLUDING REMARKS

It should become clear that current social data are
insufficient to satisfy both the demand of a coherent public
sector programme aimed at improving the community's welfare and
the requirement of social scientists.

There is scarce knowledge about the situation and the
problems i such major areas as health, housing or education. Even
more disturbing is the lack of information that could throw light
on those vital social questions; one of the main stumbling blocks
is the difficulty in linking information related to a specific
social concern with variables in other fields. In particular, it
is usually difficult to assess the propensity or risk of
different population groups (as defined by a specific set of
attributes) of falling pay to certain social ills. Isolation in
the production of data, the lack of a programme to articulate and
strenghthen the efforts of different agencies, is in the last
analysis the main reason for such a situation.

An easily identified cause is the lack of data on resources
-human, financial and others- applied to different fields and/or
population sectors. Its incidence prevents an adequate
cost/benefit analysis and makes it difficult to establish sound
basis for the judicious distribution of public expenditures.

A final -althourgh perhpas obvious- point it seems
neccessary to make is that we do not intend to suggest that
"social" problems can only be analyzed by means of better
"social" data. For it is not only that the limits between
questions labelled as social or economic or demographic,although
reflecting an obvious division of labour, are drawn with a
certain measure of arbitrariness, but also that aspects "typical"
of each of those "separate" fields interact closely and
intensively. As a consequence, economic variables must certainly
be taken into account when designing a system of social
statistics. Furthermore, even though the establishment of a
system of either social or economic statistics is necessary as a
first approximation, it is likely to prove impossible in more
advanced stages of statistical development, when only an overall
structure of socioeconomic data seems necessary.

REFERENCES

Zagorski, .(1975)

UNSO(1975) TOWARDS A SYSTEM OF SOCIO-DEMOGRAPHIC STATISTICS ,
          New York, United Nations.