COMPARABILITY OF POPULATION DATA IN PREVIOUS USSR: CASE OF ESTONIA

Kalev Katus Allan Puur Luule Sakkeus

RU Series B No 30

Tallinn 1997

ISBN 9985-820-30-4

EESTI KÕRGKOOLIDEVAHELINE DEMOUURINGUTE KESKUS
ESTONIAN INTERUNIVERSITY POPULATION RESEARCH CENTRE
Postkast 3012, Tallinn EE0090, Eesti

avaldada või jätta avaldamiseks infovõrgus, ümber kirjutada mistahes viisil või vahendiga elektrooniliselt, mehhaaniliselt, fotokopeerimise, salvestamise või muul teel.

The comparability of demographic data has always been the major prerequisite to its use for the purposes of research and decision-making. However, the importance of comparability, both in time and space, has tremendously increased with the acceleration of social change and integration of societies during the 20th century. The development of research methodology, the elaboration of sophisticated analytical tools enabling the measurement of social processes in much greater detail, have also contributed to the growing importance of the issue. As the appropriate time unit to characterise the relevant changes in population is a generation, the data required for valid conclusions on population development is expected to cover much longer time intervals, but also regions than typical to some other disciplines.

Those mentioned are only a few of the motivations underlying the activities to promote the comparability of population data on national and international level. The experience gained in Estonia over the past few years suggests that comparability as well as any other major improvement of post-soviet population and social data comparability cannot be achieved without a profound change of the entire statistical system. The main indicator of having overcome the Soviet-specific data management is the degree of integration of statistical institutions with research institutions and scientific community.

1. AVAILABILITY AND COMPARABILITY OF POPULATION DATA IN THE FORMER USSR: GENERAL FEATURES

In the former USSR the availability of population data has been traditionally poor. As it is widely known, there were two parallel series of statistical data, those for open use and those for official use only, i.e. restricted for public dissemination and scientific analysis. Such conditions have supported the genesis of a widely spread impression among Western scholars that the USSR authorities had more or less complete population statistics of relatively good quality, particularly to cover the needs of central planning. Despite some aspects of the deficiencies of the Soviet population statistics (open as well as classified) were discussed by some Western scholars [Anderson, Silver 1985a; 1989; Blum, Chesnais 1986; Feshbach, Friendly 1992; Kingkade 1985; 1989], it was believed that the major problem was related to data access.

Due to the political transformation, the restrictions concerning the access to population data vanished more or less in an instant. However, this principal change has not been accompanied by the immediate involvement of the existing population databases into the scientific circulation. One of the reasons - inadequate documentation of the datasets - is discussed below. Another reason has proved to be the disordered statistical archives.

In Estonia, after seven years of free access to data, there is still no complete overview of existing sources of population data for the Soviet period. The Statistical Office is very little or not at all interested in bringing the existing materials for the past decades into systematic order. The Office considers itself to be responsible primarily for the statistics on present moment. From this position even the latest census data are regarded as "old" and out of interest. Due to that kind of neglect, some archive materials have been recently destroyed in Estonian Statistical Office. The disorder has been strengthened by periodical restructuring of the Office, and corresponding re-distribution of archives between subdepartments. As a result, the statistical materials deposited in central archives (census tabulations, annual reports on vital statistics of earlier years etc) have been preserved and are now in relatively better condition than those kept in Statistical Office.

Due to similar system of population accounting across the countries of the former Soviet Union, the situation and the problems with population statistics are not unique for Estonia, challenges and prospects of which are discussed elsewhere [Anderson, Katus, Silver 1994]. Among the statistical institutions of successor states of the former Soviet Union, only the Statistics Russia seems to be relatively better off because of its earlier responsibility for maintaining the all-Soviet statistical system which foresaw the coordination and supervision of different governmental authorities, and making statistical calculations for all former republics. Statistics Russia has obviously benefited from the Research Institute at the Office which served as a regular basis for cooperation between statisticians and population scientists. As discussed below, judging upon the experience of Estonia, the real availability of population data for earlier decades in countries of previous USSR seems to depend much more on activities and potential of

scientific community than on the official statistical institutions of the corresponding country.
The latter statement seems even more appropriate regarding the quality of data. From the scientific point of view, a poor quality of the existing Soviet population statistics means primarily the lack of comparability. Being a complex issue, comparability involves several aspects considered below. Though discussed from the Estonian viewpoint, these dimensions are common to other countries with similar statistical context.

1.1. Population Data Are Poorly Comparable with Other Countries

In many cases internationally recommended concepts and definitions, and particularly data collection procedures have not been followed in Soviet statistics to a full extent. However, only in few cases like social structure and household/family concepts, the specificity of the Soviet statistics have been obvious. More often the differences in concepts and definitions have been less apparent, and under deficient documentation they could easily be thought to be consistent with internationally recommended practices. Such situations involve hidden incomparability which, in general, creates much more misinterpretations than outright definitional differences could do. In the following, some specific conceptual features in regard to the Soviet population statistics are presented.

In population statistics, the number of population which provides the denominator to the calculation of majority of demographic and other population-based indicators, usually refers to permanent population. In its coverage, the permanent population is close to *de jure* concept, however, having certain modulations in the Soviet application [Anderson, Silver 1985b]. On the procedural level, the definition of permanent population relies on the system of *propiska*: permissions issued by authorities providing a citizen with legal rights to live at the specified address (dwelling). Apart from regular procedures, there were temporary *propiska* and special *propiska* for some categories of population, particularly military personnel, which could make a remarkable difference between permanent and *de jure* population for some regions. As an example, in Estonia the army recruits in frontier forces (operating under KGB command) were included in the permanent population by place of service while recruits serving in regular army units were counted during a census by their residence prior to service [Katus, Puur 1993]. Besides having the implications on age-sex structures of small regions, the principles of recording military personnel must also be considered in case of other statistics. For instance, on the background of very small numbers of international migration during the Soviet period, Estonia was characterised by a considerable international migration exchange with Mongolia [Sakkeus 1994]. The explanation to this peculiar pattern comes from the inclusion of movements of military personnel into civil migration statistics.

Regarding live births and infant deaths, the definitions introduced in the Soviet Union as well as the reporting procedures did not follow the World Health Organisation's recommendations. Whereas its impact on fertility statistics has been minor, the level of

infant mortality has been considerably underestimated [Anderson, Silver 1986; Dmitrieva, Andreev 1987]. In Estonia calculations for the years 1992-1993, applying parallely both the Soviet and WHO definitions revealed the difference resulting in 16.6 percent lower infant mortality, and 25.6 per cent lower stillbirth rates [EKDK 1994d]. In cause-specific mortality statistics, the causes of death were aggregated to an extent which restricts the comparability to the level of broad ICD groups.

Compared to births and deaths, the comparability of marriage and divorce records is more complex. In Estonia, the prevalence of non-marital cohabitation as the form of conjugal unions is close to Scandinavian levels [Vikat 1994], however, neither the census nor vital statistics do pay the phenomenon necessary attention resulting in distorted estimates of marital composition of population. The biases became evident when matching individual level records from census and the Estonian Family and Fertility Survey [Katus, Puur, Sakkeus, Silver 1995]. Regarding divorces, the events were counted in statistics not in relation to the legal dissolution by court but later, at their registration in Civil Registration Office. In many cases registrations have taken place after official dissolution with significant time lag; in considerable number of cases at least one of the divorcees has not visited the Civil Registration Office at all. These procedures have violated the comparability of divorce and family statistics, but also the consistency between census and vital registration.

In the Soviet Union, the concept of household has not been applied, however, the definition of family combined some aspects of household definition. For example, in census the special category of "family members living apart" was introduced which has no parallels in international statistics. Several other specifities of Soviet family concept should also be considered [Anderson 1986; Bondarskaya, Darsky 1990; Volkov 1986; 1991].

As to migration, the registration of moves has been carried out in the Soviet Union by applying rather peculiar procedures (discussed briefly above), and correspondingly the migration data should be regarded with special care. However, it must be noted that in migration statistics in general considerably less comparability can be found between countries: for internal migration even no international recommendations do exist. In these circumstances the comparability matters of Soviet migration statistics have not been extensively discussed in demographic literature. Due to its impact on population numbers, family/household composition and other characterisitcs of population, comparability of migration statistics should not be disregarded.

Compared to registration of events, recording the individual's characteristics of those involved into the population events is creating more problems with the data comparability. Relative to the concepts and definitions of vital events, problems of comparability of personal characteristics stem mainly from the development of specific classifications which did not aim at international comparability at all. For example, the official social structure was classified into three main categories - workers, collective farmers and the intelligentsia - which was useless not only for international comparison but also for whatever scientific analysis. Several important characteristics like place of birth (everybody was born in the Soviet Union) or religion (everybody had to be an atheist) were fully omitted from the official statistics. For economic characteristics, the

occupation and industry were not recorded in vital statistics; classifications of occupations which were applied in the census cannot be mapped into ISCO. Also one must be cautious to use the data on ethnicities, particularly on minority nations because the ethnic characteristics was given the political meaning, and in times, people were subjugated upon these characteristics.

1.2. Population Data Are not Comparable in Time

Following the time horizon of central planning, statistical institutions like other institutions were oriented to work with one-year or maximum five-year perspective. When receiving an order from the Communist Party or the Council of Ministers to prepare specific materials, usually short deadlines to accomplish the task were considered. Those, requesting the data were interested in quick and clear answers and had little understanding or concern for data quality.

Data on more distant periods was usually considered as "old", and of little interest. Under such circumstances statistics were compiled restrictedly according to concepts, definitions and methods valid at the moment, with no efforts to achieve the comparability over time. Emphasis by the statistical institutions on data quality, including the comparison in time could only cause trouble for the statisticians themselves. When the access to data was provided in the end of the 1980s, it was to some extent a surprise even to local researchers, how little comparability in time the population statistics possesses.

In the conditions of restricted data access, there was no way to develop skills and habits among statisticians to consider data quality. Statistical institutions were not expected to publish or make otherwise available any documentation addressing the reliability and accuracy issues. Neither were they interested in it by themselves; besides extra work, preparing such materials could have created only criticism towards the institution. Because of overcentralisation, even calculations of the most basic indicators were performed in Moscow, and the staff of regional statistical institutions in republics had no need to be aware of the methods applied.

Being dependent on the good will of the Statistical Office on access to data and documentation (if available) it was hard if not possible at all for scientific community to evaluate data quality. At the most, there have been some conferences to discuss the concepts and definitions applied at censuses, however the practice has not spread to republics [TsSU 1958; 1969; 1979]. In short, the described practice suggests that it is naive if impossible to expect good quality data in the circumstances of restricted access. It is even hard to discover simple mistakes once they have been made. For example, it has recently been discovered that the Estonian population figures of 1970 census by concept of permanent population, available in local publications, proved to be incorrect [TsSU ESSR 1975]. Once taken from an irrelevant tables and published, they went into circulation from one publication to another. Only more than 20 years later, after comparing the total numbers, archived age structures and documentation, the mistake was recovered [EKDK 1994b].

In the following, some specific continuity problems of the Soviet population statistics over time are discussed. Regarding the population numbers, up to the 1979 census, the postcensal estimates were produced on the basis of the present (*de facto*) population, later on they were performed on the basis of permanent population. From the methodological point of view, producing the estimates by updating census counts, based on present population with vital events, referring to the permanent population, is inconsistent. What concerns the crude rates for the country as a whole such an inconsistency arises minor differences, however, for more sophisticated indicators the differences could be of statistical significance, particularly in regard to republics and smaller regions of the previous USSR.

Despite after each census the population estimates were on republic/oblast level recalculated, the time series of demographic indicators were not updated. From the comparability perspective, the undoubtedly progressive introduction of new analytical methods and techniques should be accompanied with careful evaluation of effects for the consistency in time series. In Soviet official statistics such evaluations were hardly done and the extent of potential problems remains to a large extent unknown [Mesle, Shkolnikov, Vallin 1992; Andreev, Scherbov, Willekens 1993]. Recent recalculation of Estonian life tables using primary data and the same method for all census years displayed that the methods used by Central Statistical Office to compute the official life tables in different years, had had differential effect on the results. Life table indicators for the 1970 census year deviated considerably more than the corresponding indicators for other census years [Katus, Puur 1991].

Concerning the statistics of illegitimate births, marriage and divorce, considerable discontinuity occurred in 1968 when changes in recording practices were introduced by the new Marriage Law. Again, no efforts were taken to secure the consistency with earlier years. Migration statistics had undergone repeated changes, therefore without recalculation no any migration indices can be compared over time [Katus 1989; Sakkeus 1993]. Relative to vital statistics, despite in several points not fully consistent with international recommendations, the stability of census methodology has secured greater comparability of census data.

Presenting here a short listing of the problems connected with comparability over time, the purpose is by no means to give the negative evaluation to the introduction of concepts, definitions and calculation methods applied in population statistics during the Soviet period. As concerns the calculation methods, the problem arises from the same overcentralisation phenomenon, described above. Only in a few cases the selection of calculation methods may have been dictated primarily by manipulation interests.

As the comparability with other countries may be thought to be generally, though with different pace and with decline, improving, the case with comparability over time is more complex. Namely, the harmonisation of statistical definitions with international recommendations may bring about new discontinuity in time series, and the incomparability in time may even decrease. Therefore the innovations must certainly consider the need to maintain comparable time series.

1.3. Regional Population Data are Poor and Often non-Comparable with Regions of Higher Order

One of the specific features of the Soviet population statistics has been its hierarchic nature, particularly regarding vital statistics: little, if any, statistics was produced for smaller administrative units compared to larger ones. Only the number of major demographic events without sex/age distribution, and respectively only crude rates are available below oblast level. In relation to smaller countries of the former Soviet Union, including Estonia, this resulted practically in the absence of regional vital statistics. Consequently, research on regional heterogeneity/homogeneity of demographic processes was virtually impossible. Two main reasons for that feature can be outlined.

First, the traditions of working with individual level data are poor, statistical institutions have been used to operate with aggregated data. This split dates back to the period when data processing started to be done on mainframe computers, and when computing centres were formed as separate units within statistical institutions. According to the division of functions, the statistical staff gave commands to the computing centres, who performed the processing, and supervised the results. It resulted in a non-interest in data processing and the lack of technical skills among the statistical staff. For example, as late as in 1991 Estonian Statistical Office showed no initiative to start work with the individual level census files because having the tabulations produced in Moscow; when introducing the new birth registration system in 1992, the Statistics Estonia expressed views that the individual records are not statistics as the latter begins only starting from the aggregated tabulations. Regarding the regional level, the incapacity to work with individual-level data restricted the statistical institutions from tabulating the data for smaller regions, particularly under periodical boundary changes between these regions.

Secondly, under the Soviet regime the statistical information was not meant to support decision-making, based mostly on ideological grounds and preferences of Communist party leaders. Decisions were made primarily on the top-level and to such scheme, no disaggregated data on regions below the republic/oblast level was necessary. And naturally it was not produced in the framework of usual routines of statistical calculations.

Besides the poor data availability at regional level, the kind of system had inherent consistency problems. By Soviet routines the results, whatever the indicators, were first produced for the whole country, and only later for the regions. Often, the summary of the data of individual regions did not sum up to the numbers produced directly for the higher hierarchical level. In case inconsistencies were detected, the regional data was "adjusted", usually introducing inaccuracies. For example, using this kind of procedure, in the 1959 census the number of population of one small Estonian town, Loksa, was estimated ca 200 per cent larger than its real number [TsSU ESSR 1960, p.44]. Even as late as in the 1989 census, the number of population for the whole Estonia was adjusted

by adding 7000 individuals to permanent population. The Central Statistical Office in Moscow motivated it with the fact that a certain number of permanent residents of Estonia were counted in other territories. However, actually 7000 temporary residents, enumerated at the time of census in Estonia, were recoded into permanent, supposingly to match present and permanent population estimates for the Soviet Union as a whole. Similar kind of adjustments were usual in compiling the vital statistics.

Another factor introducing non-comparability into regional population statistics is connected with the pattern of producing estimates of population after new round of census, for the previous intercensal period. If the numbers and age structure of the USSR and the republics were recalculated, numbers for smaller units like *raions* and *selsovets* were not. This kind of approach eliminated the possibility of matching the total numbers of regional units at the higher hierarchical level with lower ones, actually comprising these regions. For Estonia, the population numbers and particularly the age structure, when recalculated at the regional level and balanced with national figures, differ considerably from old official figures proceeding from only the national level [EKDK 1994a; 1994b; 1995c; 1995d; 1995e; ESA 1990].

The comparability of regional data is further complicated by continuous administrative transfers and boundary changes. Lacking any autonomy in decision-making, the boundaries of smaller administrative units were repeatedly redrawn without any real need. Therefore in the Soviet Union, the smaller the regional unit, the harder efforts are needed to build comparable time series for demographic indicators.

2. REORGANISATION PLAN OF THE POPULATION AND SOCIAL STATISTICS IN ESTONIA

After cessation of the last Soviet regulations limiting the availability and publication of population and social statistics, it became possible for the scientific community and other interested organisations to make a comprehensive review and evaluation of the existing data. Covering different domains, relevant reports were prepared and repeatedly discussed at the sessions of Estonian Demographic Association (EDA) [Katus, Puur, Sakkeus 1992].

As a result of these investigations and meetings, an unanimous understanding emerged that Estonian population data available for the Soviet period is of unsatisfactory quality, particularly due to the presence of non-comparability discussed above. For the same reason, it was concluded that most of the social and population statistics in existing quality is not consistent for making valid scientific conclusions. Besides research, the current state of the data was regarded a hindrance for qualified decision-making which is especially true when heterogeneity aspects and behavioural mechanisms are concerned. Recalculation of the population data from the Soviet period, initiated already in the 1980s by the EDA and EKDK, was brought into the long-term programme to achieve time- and region-consistent series of basic demographic indicators for Estonia. Some accomplishments of the programme are discussed in the third section of the present paper.

Regarding the data on present and future population, it was concluded by scientific community that the incomparability and low quality of statistics does not result from a few specific reasons but rather from the entire statistical system, lacking particularly the overall responsibility and interest in data integrity. Under the Soviet system, the population data, particularly vital events were registered by different authorities. The following stages of the statistical system, namely data coding, entry, cleaning and documentation, production of standard tabulations, computation of basic indicators, archiving etc involved numerous additional institutions into the system, according to the Soviet practice to have separate institutions for nearly each function. Such kind of disaggregation also served for the aims of a totalitarian society to control the information. To overcome the deficiencies imposed by such a system, the main attention had to be focused on the redefinition of responsibilities and cooperation of governmental authorities involved in data collection, processing, analysis and dissemination, not just improving the performance of existing structures. In other words, it became clear that the task which at first glance was supposed to address specific problems, proved to be extensive reorganisation of the entire statistical system comparable with the currency reform in economy. Apart from general problems resulting from the Soviet-type statistical context, in case of Estonia, the reorganisation has to overcome the non-existent country-level coordination, which earlier had been performed on the level of central authorities in Moscow.

Based on the initiative of the EDA, to reorganise the statistical system the Governmental Commission on Social and Population Statistics was established in 1993. The principal aim of the Commission was to integrate all stages of statistical system: starting from data collection up to the dissemination and archiving into the system, where the institutions responsible for the different stages would be acting in a coordinated manner, supporting the activities of each other. Thus, the purpose was to overcome the Soviet practice under which these interrelated functions were only loosely integrated with some of them significantly underdeveloped. In operational terms, all relevant governmental and non-governmental agencies (more than 30 altogther) as well as scientific institutions were involved in work of the Commision, at the executive level the relevant Ministers were nominated to the Commission.

As the result of the systematic review and negotiations, the plan of statistical reorganisation was prepared and presented to the Government for adoption [Katus, Kõre, Pavelson, Puur, Sakkeus 1993]. The plan covered all stages of statistical system and was secured with cooperative agreements between agencies as well as with finances. The plan was evaluated by international experts and coordinated with the statistical authorities in neighbouring Latvia and Lithuania (Training Seminar of Baltic Population Statisticians, Laulasmaa, July 27-August 1,1993). Contrary to the expectations, the plan was not implemented. Despite the goal of the undertaking remains unachieved, the work in the Commission clearly augmented the challenges and opportunities of the post-Soviet statistical environment.

From the positive side, the excessive majority of governmental and non-governmental institutions in Estonia proved to be cooperative and supportive to the reorganisation of social and population statistics. For the first time in the post-war history, different

Estonian institutions provided their time to discuss systematically the issues of statistical organisation. For most of these organisations the data collection and/or participation at other stages of statistical system was not the main task, however, they willingly agreed to go along with the reorganisation to improve the performance of their role in data management. It must be positively underlined in particular because most of these institutions had until recently worked under separate regulations (mediated only by central authorities in Moscow), and had little, if any, direct contacts between themselves.

From the negative side, one group of administrators systematically hindered the reorganisation plan, and finally managed to prevent its implementation. In a transitional society, one could indeed expect some social groups to oppose new developments, however, in case of reorganisation of social and population statistics in Estonia it happened to be a rather special case that deserves explanation.

In Estonia, the opposition to the reorganisation was formed by the computing staff as a professional interest group. All over the Soviet Union the computing staff was overproduced during the mainframe era, and huge computing centres were established. Being part of the official structures, these centres achieved the status of monopoly in data processing. Because of techological lag, these centres as well as their commanding position in statistical system survived until the beginning of transitional period. In this period, computing centres developed their for-profit orientation while the technological innovation made their existence redundant. To survive, taking an advantage of their old privileged position and connections, computing staff is trying to maintain their status, which in Estonia has resulted in the claims for the central role in the statistical system. Having administratively gained the central role and lacking interest in ensuring the integrity of statistical system, computing staff mainly concentrates on increasing the volume of data processing, i.e taking the narrow approach. Proceeding from the technological perspective, they neglect international concepts and definitions and underestimate the harmonisation of data collection procedures. Being mainly for-profit oriented, the dissemination stage and making the data publicly available is not in their interest as well. As a result, even despite their good will, the data quality in the society cannot be secured. In such circumstances, the reorganisation of statistical system would have clearly harmed their position: computing staff would have lost their commanding role or, in better case, it would have given them much more responsibility than now.

In short, because of strong opposition of computing personnel the social and population data reorganisation has not succeeded. However, such a result in a society, driving for democracy, would have not been possible, if the Statistical Office would not have supported their narrow approach. The present leadership of Statistical Office belongs to the same interest group, having the computing rather than the statistical background. The relatively poor understanding of the importance of information for decision-making process and data quality issues among the members of the Government was another feature contributing to the failure of the reorganisation.

Currently, Estonia is allocating relatively large funding in data managament (computing and communication facilities etc). Unfortunately, giving the funds to institutions which are neither responsible for nor interested in the comparability of concepts and

definitions, data accuracy and availability, the overall results are depressive. The data quality, including comparability in time and space, in spheres like migration, marriages and divorces, calculation of population indices etc has even worsened compared to Soviet period. For example, in vital statistics the ISCO-based coding of occupations for persons undergoing certain demographic events was ceased in 1994 and replaced with simple text-entry without slightest classification. The preparations for the new 2000-year-round census have not begun from the methodological point of view and are thought to be solved solely through administration decisions. Statistical Office addressed once the scientific community through EDA, where the latter expressed the view for the need to adapt the internationally recommended standard methodology to country-specific circumstances, which would secure comparability not only with other countries but also earlier censuses taken in Estonia. As previous censuses have used different approaches, the adaptation work required is rather extensive. Statistical Office was took by an unpleasant surprise by the unanimous understanding among the scientists that the real preparation work for census in Estonia has not begun in 1977. Despite readiness expressed by the scientists to secure the time for this large-scale methodological effort, Statistical Office has not considered the involvement of the local scientific community worthwhile.

In some ways, Estonian Statistical Office seems to be approaching a reality of which Frans Willekens has warned in connection of the development of computer technology: just as much helping to produce lot of sense it could be used to produce a lot of nonsense [see also Willekens 1992].

3. REORGANISATION FAILED, WHAT CAN STILL BE DONE

After the reorganisation of social and population statistics has failed in Estonia, the willingness of data management institutions and the Statistical Office, run by computing staff, to cooperate with the scientific community has noticeably decreased. This tendency has strengthened after the last Parliament elections and formation of the new Government. As mentioned above, some of the scanty achievements in reorganisation like introduction in the records of all vital events of personal characteristics based on comparable definitions, introduction of internationally recommended procedures in the calculation of indices and principles of archiving have returned to the old practice.

In spite of being unable to reform the statistical system, and regardless of the worsening perspectives towards it, the scientific community and other institutions/persons interested in population data availability and quality are continuing the efforts to achieve the aims set up. During the recent period, the advancement has been possible in three directions.

First, in contrast to the deterioration of vital statistics the improvement in the infrastructure of survey-based data collection has been noticeable. In 1993 the Governmental Commission on Population and Social Statistics adopted the plan of nation-wide surveys to be laucnhed in Estonia up to the next census. Altogether the

plan included 8 surveys. In 1994-1995 two of these large-scale sample surveys, female part of Estonian Family and Fertility Survey (FFS) and the Labour Force Survey (LFS) have been conducted, which are the first surveys to be based on nationwide probability samples in Estonia [Katus, Puur, Sakkeus, Silver 1995; Katus, Noorkõiv, Puur 1997].

Characteristic to the Soviet-type statistical system, out of three sources of population data - the census, vital statistics and surveys, the latter was obviously underdeveloped. As an organisational innovation, to carry out these surveys relevant Working Groups as project-based bodies outside the regular administrative structures, were established. Working Groups were delivered all responsibilities for the survey starting from planning and instrument development up to data release and scientific analyses. Such solution proved to be an efficient way to unite relevant experts from scientific institutions as well as represent the interest of governmental institutions. Because of opposition from Statistical Office and computing staff to new arrangements, Working Groups for both surveys were registered at the Governmental Commission.

In Estonia, given the lack of experienced personnel and skills within Statistical Office, the newly-established commercial survey agencies had received certain competitive advantages in conducting nationwide social surveys. For that reason, the procedures for a large-scale social survey were elaborated and tested by the FFS Working Group. On the basis of that experience, the LFS Working Group managed to establish a regular interviewer network based on the infrastructure of county statistical bureaus, which showed already good performance compared to the commercial ones. Both surveys relied on the only available sampling frame with nationwide coverage - 1989 census microdata - which proved not only to support the strategies but also for, the first time in Estonia, individual-level linkages with surveys. However, in the development of the interviewer network the Statistical Office has not emphasised the data quality and intergration with other data sources as ultimate goals. Despite the advice provided by scientific community, network has not established a core of permanent full-time interviewers and instead, is preferring greater number of less experienced interviewers. Another serious deficiency relates to working language of interviewers, too small number of Russian-speaking interviewers threatens the principle according to which respondents should be entitled to the interviewer, sharing the same mother tongue. These circumstances are placing considerable threat on quality of the results and timely completion of the two ongoing surveys: male part of FFS and Survey on National Minorities.

In terms of budgetary and planning arrangements, the Statistical Office working under yearly budgeting cycle had never practiced project-type planning for statistical undertakings longer than one year. However, the surveys of this scale demand usually several years to be prepared and implemented, and even more if to include the analyses stage into the project. Therefore the budget arrangements became an important part of the Working Group activities. As yet, the Estonian Statistical Office is still opposing this type of financing, in principle opposing the cost-efficiency inherent in the proposed scheme. In these circumstances, although the Working Groups were able to introduce the elements of project-based financing, securing funds on multilateral basis, however, the principles of such kind of funding have not yet been rooted. Therefore, the continuity of the similar undertakings is not secured.

Besides conducting research, the Working Groups for FFS and LFS have also committed themselves to provide data for decision-making. Volumes of standard tabulations, in close cooperation with interested governmental agencies are published or planned to be published [EKDK 1995a; EKDK 1997], and relevant training sessions for administrators will be organised. In short, the FFS and LFS Working Groups aim at building the infrastructure for large-scale surveys, starting from budgeting routines to applications in public administration.

The failure of systematic reorganisation of statistical system is especially regretful from data comparability perspective, as it has slowed down the pressure for Statistical Office to move towards international concepts and definitions. As the second positive direction, methodologically FFS and LFS are contributing to introduction of internationally comparable concepts and definitions of demographic events and statuses. As a part of the analysis of the two mentioned survey data, the definitions of household, place of residence, economic activity etc are elaborated to be introduced into official statistics. The introduction of these concepts depends, of course, on the statistical institutions, and, in case of Estonia, on the understanding of such need by computing staff in charge of it.

As regards to official structures, the Medical Statistics Bureau operating as a separate unit under Ministry of Social Affairs has been responsive and much more cooperative with scientific community, achieving better results in its field of responsibility. As an example, in 1992 the WHO definition of live births was introduced together with giving a start to the medical birth register. Together with characteristics concerning the medical details of delivery, social characteristics of a mother and a father, which are the same comparable to characteristics in the civil birth certificate, were included in this register. The parallel activities of medical and civil registration systems provides a unique basis for evaluation of the quality of birth registration, particularly the recording of personal characteristics. Since 1994, Medical Statistics has also started the abortion register. The Institute of Experimental and Clinical Medicine together with Medical Statistics Bureau have introduced international classifications in morbidity and mortality statistics. [Bogovski, Laan 1994].

The third direction of activity, that can be mentioned is the recalculation of the retrospective population data. Actually, this is the realisation of an initiative of the EKDK and Demographic Association already from the middle of the 1980s. Despite it was quite obvious already by that time that in existing quality the population data could not support valid scientific conlusions, the initial task has expanded tremendously after gaining the full access to population data. Until now the funds have been very scarce, limiting the publication possibilities to just a few volumes of statistical recalculations [EKDK 1994a; 1994b; 1994c; 1994d]. The recalculation plan proceeds with the series of new tabulations from 1979 and 1989 censuses at municipality level, and provides recalculated vital rates on the basis of recalculated age structures. Three separate volumes mainly concentrating on the period 1965-1989, also including data from previous periods as well as current population data, have recently been published [EKDK 1996a; 1996b; 1996c]. The retrospective computerisation of birth records from the 1960s has also been started from 1995.

The above-mentioned direction foresees the documenting and rearranging of existing population databases, in order to bring them into the scientific circulation. The major individual-level dataset that has been documented until now is the 1989 census microdata [Katus, Puur 1993; Puur 1994]. The data from the 1989 census has been already used for drawing the database on the elderly for the *Dynamics of Population Ageing Project*, coordinated by UN ECE PAU, as well as for FFS and LFS sampling [EKDK 1995a; EKDK 1995b; EKDK 1996]. The similar work is going on with the data of the 1979 census and the 1985 microcensus. The aggregated population data has been organised into the Estonian Population Database, managed by EKDK [Katus, Puur 1992; Katus 1992]; all of the recalculated data will be added to this database.

Rearrangement and documentation of databases goes in parallel with seeking possibilities to make data available. In addition to traditional means of data dissemination, the INTERNET is in consideration. In that sense, the CIESIN country node which attempts to mediate the main Estonian databases into international circulation, may be one of the ways [Human Dimensions Quarterly 1994].

4. INSTEAD OF CONCLUSION

As a major conclusion, the experience gained in Estonia during the past few years suggests that comparability as well as any other major improvement of post-soviet population and social statistics cannot be reached without the profound change of the entire statistical system. This matter implies that to recover the situation, the main attention has to be focused on the redefinition of responsibilities and advancement of cooperation between governmental authorities involved in data collection, processing, analysis and dissemination, not just on improving the performance of existing structures.

From the perspective of international cooperation and particularly assistance programs, this involves a delicate situation, indeed. The assistance programs for East European countries which normally operate through official structures may unwillingly appear in the position of working with conservative Soviet-type structures. In a better case this means inefficient use of resources, in worst ones, it may result in the strengthening of these structures, i.e. decreasing the data quality and comparability. For Estonia, it has meant that assistance programs have provided the Statistical Office with excuses to ignore the local scientific community and reject its willingness to cooperate.

Unfortunately, the leadership of Statistical Offices of Baltic countries has rather found easier to fully rely on direct coordination from Luxembourg than to cooperate with their own scientific community. Seldomly have international gatherings focused on bringing together statisticians and relevant scientific communities of these countries, in order to improve the methodological basis in all fields of statistics. One of the few examples have been the Training Seminar of Baltic Population Statisticians, Laulasmaa, July 27-August 1,1993, which addressed the common problems inherited from the previous USSR and tried to secure further coordinated steps. The other such an example was the

Seminar on Migration Data, organised together with IOM, OECD and Eurostat and relevant Baltic scientists in February 1996. Just the opposite seems to become true with respect to preparations for the year 2000-round census in Estonia, where the scientific community had to declare that until 1997 the methodological preparations for that undertaking have not yet begun.

It seems that in the former Soviet Union countries, one of the most important criteria of adequate reactiveness, regarding statistical institutions, is their degree of integration with research institutions and scientific community. The latter is considered an indicator of the extent to which the statistical institutions have managed to overcome their Soviet-specific role and started serving the needs of society in providing accurate information.

REFERENCES

Anderson, Barbara (1986). Family, Marriage, and Fertility in Russian and Soviet Censuses. *Ralph Clem (Ed). Research Guide to Russian and Soviet Censuses*. Ithaca, Cornell University Press, pp.131-154.

Anderson, Barbara and Brian Silver (1985a). Estimating Census Undercount from School-Enrollment Data: An Application to the Soviet Censuses of 1959 and 1970. *Demography*, vol.22, No.2, pp.289-308.

Anderson, Barbara and Brian Silver (1985b). 'Permanent' and 'Present' Populations in Soviet Statistics. *Soviet Studies*, vol.37, No.3, pp.386-402.

Anderson, Barbara and Brian Silver (1986). Infant Mortality in the Soviet Union: regional differences and measurement issues. *Population and Development Review*, vol.12, No.4, pp.705-738.

Anderson, Barbara and Brian Silver (1989). The Changing Shape of Soviet Mortality: An Evaluation of Old and new Evidence. *Population Studies*, vol.42, No.2, pp.303-320.

Anderson, Barbara and Brian Silver (1990). Growth and Diversity of the Population of the Soviet Union. *Samuel Preston (Ed). World Population: Approaching the Year 2000. The Annals of the American Academy of Political and Social Science*, No.510, pp.155-177.

Anderson, Barbara, Kalev Katus and Brian Silver (1994). Developments and Prospects for Population Statistics in Countries of the Former Soviet Union. *Population Index*, vol.60, No.1, pp.4-20.

Andreev Evgeni, Sergei Scherbov and Frans Willekens (1993). *Mortality in the former Soviet Union: Past and Future*. IIASA Working Papers, WP-93-13. Laxenburg, IIASA.

Blum, Alain and Jean-Claude Chesnais (1986). La pyramide des ages de l'Union Sovietique au recensement de 1979. *Population*, vol.41, No.6, pp.1043-1058.

Bogovski, Pavel ja Ilmar Laan (Toim) (1994). *Rahvusvaheline haiguste klassifikatsioon. RHK-10. I köide.* [International Classification of Diseases. ICD-10]. Vol.I. Tallinn, EKMI.

Bondarskaya, Galina and Leonid Darsky (1990). Marital status of women and fertility. *Andrei Volkov (Ed). Demographic Processes in the USSR*. Moscow, Nauka.

Dmitrieva, Rimma and Jevgeni Andreev (1987). O srednei prodolzhitel'nosti zhizhni naselenia SSSR. [On the Life Expectancy of the Population of USSR]. *Vestnik Statistiki*, No.12, s.31-39.

EKDK (1994a). *Population age structure: Counties 1979-1989*. RU, Series C, No.2. Tallinn, EKDK/EKMI.

EKDK (1994b). *Population age structure: Counties 1970-1979*. RU, Series C, No.3. Tallinn, EKDK/EKMI.

EKDK (1994c). *Life tables: Counties 1986-1991*. RU, Series C, No.4. Tallinn, EKDK.

EKDK (1994d). *Infant Mortality: Counties 1965-1993*. RU, Series C, No.5. Tallinn, EKDK.

EKDK (1995a). *Estonian Family and Fertility Survey. Standard Tabulations*. RU, Series C, No.6. Tallinn, EKDK, .

EKDK (1995b). *Estonian Population Ageing Project. Standard Tabulations*. RU, Series C, No.7. Tallinn, EKDK.

EKDK (1996a). *Reviewed Population Census and Vital Statistics. Viljandimaa.* RU, Series C, No.8. Tallinn, EKDK.

EKDK (1996b). *Reviewed Population Census and Vital Statistics. Lääne-Virumaa.* RU, Series C, No.9. Tallinn, EKDK.

EKDK (1996c). *Reviewed Population Census and Vital Statistics. Järvamaa.* RU, Series C, No.10. Tallinn, EKDK, forthcoming.

EKDK (1997). *Estonian Labour Force Survey. Standard Tabulations*. RU, Series C, No.11. Tallinn, EKDK/ETU, forthcoming.

ESA (1990). *Eesti vabariigi külanõukogude ja maa-asulate rahvastik. 1970 aasta, 1979 aasta ja 1989 aasta rahvaloenduste andmeil*. [Population of Rural Administrative Units. On the basis of Data of 1970, 1979, 1989 censuses]. Tallinn, ESA.

Feshbach Muriel and A Friendly (1992). *Ecocide in the USSR: Health and Nature under Siege*. New York, BasicBooks, xvii.

Human Dimensions Quarterly (1994*).* Vol.1. Saginaw, CIESIN.

Katus, Kalev (Ed) (1989). *Migration Development: Towards the Regulation*. *Conference Papers*. Tallinn, EKDK.

Katus, Kalev (1992). *Eesti Regionaalne Rahvastikuandmete Pank II*. [Estonian Regional Population Databank II]. RU, Series A, No.31. Tallinn, EKDK.

Katus, Kalev, Jüri Kõre, Marje Pavelson, Allan Puur ja Luule Sakkeus (1993). *Indviduaalstatistika ümberkorraldus Eestis.* [Reorganisation of Indvidual Statistics in Estonia]. Report for the Governmental Commission on Population and Social Statistics. Tallinn.

Katus, Kalev ja Allan Puur (1991). Eesti rahvastiku suremustrend elutabelite analüüsi põhjal. [Estonian Mortality trends Based on Analysis of Life Tables]. *Akadeemia*, No.12, lk.2516-2549.

Katus, Kalev ja Allan Puur (1992). *Eesti Regionaalne Rahvastikuandmete Pank I*. [Estonian Regional Population Databank I]. RU, Series A, No.29. Tallinn, EKDK.

Katus, Kalev and Allan Puur (1993). *The 1989 Estonian Population and Housing Census: Data Description*. RU, Series A, No.33. Tallinn, EKDK.

Katus, Kalev, Allan Puur and Luule Sakkeus (1992). *Riiklik rahvastikustatistika: terviksüsteem ja osised*. [National Population Statistics: the Integral System and its Constituents]. RU, Series A, No.32. Tallinn, EKDK.

Katus, Kalev, Allan Puur, Luule Sakkeus and Brian Silver (1995). *Estonian Family and Fertility Survey. Design and Methodology.* RU, Series A, No.39. Tallinn, EKDK.

Katus, Kalev, Rivo Noorkõiv and Allan Puur (1995). *Estonian Labour Force Survey. Methodological Report.* RU, Series A, No.40. Tallin, EKDK/ETU, forthcoming.

Kingkade, Ward (1985). *Evaluation of Selected Soviet Population Statistics.* CIR Staff Report, No.9. Washington, US Bureau of the Census.

Kingkade, Ward (1989). Content, Organization, and Methodology in Recent Soviet Population Censuses. *Population and Development Review*, vol.15, No.1, pp.123-138.

Puur, Allan (1994). *1989 Estonian Population and Housing Census. Microdata Description: Variable Definitions and Coding Schemes II.* RU, Series A, No.36. Tallinn, EKDK.

Sakkeus, Luule (1993). *Post-war Migration Trends in the Baltic States*. RU, Series B, No.20. Tallinn, EKDK.

Sakkeus, Luule (1994). The Baltic States. *Solon Ardittis (Ed). The Politics of East-West Migration.* London, MacMillan Press, pp.68-85.

TsSU USSR (1958). *Vsesoiuznoe soveshchanie statistikov 4-8 Iunia 1957 g.* [All-Union conference of statisticians of June 4-8, 1957]. Moscow.

TsSU USSR (1969). *Vsesoiuznoe soveshchanie statistikov 22-26 Aprelia 1968 g.: stenograficheskii otchet* [All-Union conference of statisticians of April 22-26, 1968: stenographic record]. Moscow.

TsSU USSR (1979). *Sovershenstvovanie gosudarstvennoi statistiki na sovremennom etape: materialy vsesoiuznoi soveshchaniia statistikov.* [Improvement of state statistics at the contemporary stage: materials of the all-Union conference of statisticians]. Moscow.

TsSU ESSR (1960). *Tschislennost naselenia i ego terriorialnoie razmeshchenie.Nacional'nyi sostav naselenia. Vseosojuznaia perepis' naselenia 1959 g.*[Number of Population and its regional Distribution. Ethnical Structure of Population. All-Union census of 1959]. Vol.1., Part 1. Tallinn, TsSU ESSR.

TsSU ESSR (1975). *Naselenie Estonskoi SSR. Dannye perepisi naselenia 1970 g*. [Population of Estonian SSR. Population Data of 1970 census]. Tallinn, TsSU ESSR.

Mesle, France, Vladimir Shkolnikov and Jacques Vallin (1992). Mortality by Cause in the USSR in 1970-1987: the reconstruction of time series. *European Journal of Population*, vol.8, No.4, pp.281-308.

Vikat, Andres (1994). *Family Formation in Estonia.* Helsinki, Finnish Demographic Society, No.15.

Volkov, Andrei (1986). *Semja - objekt demografii*. [The family - object of demography]. Moscow, Mysl.

Volkov, Andrei (1991). Family and household. *Vestnik Statistiki*, No.7, pp.40-46.

Willekens, Frans (1992). The IUSSP Software Project. *Expert meeting on Demographic Software and Microcomputing*, vol.1-2. Hague, IUSSP/NIDI.