

ESTONIAN FAMILY AND FERTILITY SURVEY:
EXPERIENCE FROM TRANSFORMING STATISTICAL ENVIRONMENT

Kalev Katus Allan Puur Luule Sakkeus

Estonian Interuniversity Population Research Centre
P.O.Box 3012 Tallinn 10504 Estonia
Ph 372 6409451 Fax 372 6409453 E-mail asta@ekdk.estnet.ee

Paper presented to FFS Flagship Conference
Brussels, 29-31 May 2000

Session 3: New Approaches and Methodological Innovations in the
Study of Partnership and Fertility Behaviour
(submitted to Session 5: Research and Policy Agenda for the Future)

Tallinn 2000

The planning of the Estonian FFS was initiated by the Estonian Interuniversity Population Research Centre shortly after the restoration of statehood in 1991. In Estonia, the FFS became the first nationally representative social survey. From the viewpoint of the development of the statistical system, such position represented an extreme case, however, also in several other countries of Central and East Europe, FFS turned to be the first internationally comparable and methodologically advanced representative survey. Estonian FFS, besides its regular tasks encountered the wide range of problems inherited from the previous statistical system. Particularly, the data quality was considered the first priority, while planning and implementing various activities related to the survey. First of all this concerned the female survey (fieldwork 1994), as most of the procedures had to be developed from scratch. The gained experience was applied to the male survey (fieldwork 1997-1998), however, paradoxically demanding no less effort.

1. Characteristics of Survey Environment

The understanding of the preparation and implementation of the Estonian FFS is not complete without considering the general statistical environment, which for half a century was built to serve the Soviet regime. Although the society has rapidly changed, the statistical system inherited from the past is still operating and the principal reform yet needs to be started [Anderson, Katus and Silver 1994; Katus, Puur and Sakkeus 1998]. From the wider perspective of the national statistical system, four major features of the FFS environment deserve particular attention.

The first feature of the survey environment has been the absence of relevant sampling routines which are necessary for any nationally representative survey. Although a series of demographic surveys had been carried out in the Soviet Union, covering Estonia as well, samples were produced outside the country, by the Central Statistical Office in Moscow. These surveys aimed at representativeness for the Soviet Union as a whole, not for regions [Darsky 1986; Volkov 1997]. Though, as far as it is known, the detailed description of the referred sampling procedures had never been published, it is apparent from the sample size that the data cannot be representative for the smaller republics. Due to multistage clustering procedures usually just one-two urban/rural communities were included in the sample from Estonia. In such organisation, neither sample frame on local population nor appropriate skills were developed in Estonia. According to the archive records, the Statistical Office of Estonia never raised the question about the representativeness of survey statistics concerning the country.

Under the conditions of state socialism, scientific community was granted a limited access only to aggregated population data, the use of individual data by academic institutions was legally forbidden. Thus, instead of application of available routines, the sampling and elaboration of respective procedures for the female survey of the Estonian FFS turned into an innovative task [EKDK 1995a; Katus and Puur 1993]. This task included, among others, the development of sample frame and specific address verification routine. The sample frame and established procedures were also applied in the male survey, saving time and resources. However, the task had still to be accomplished by the Estonian FFS Working Group as the statistical

institutions continue to rely on frame/procedures, inappropriate for nationwide representative surveys [ESA 1998a; 1998b; 1998c].

The second feature of the survey environment to be noted was the absence of national interviewer network. The fieldwork of the FFS female survey was conducted by a private survey agency, oriented to market surveys and opinion polls. Consequently, the interviewers lacked appropriate experience for collecting event history data, including on sensitive topics, therefore the training of the network and close supervision of all stages of fieldwork became one of the central concerns in Estonian FFS. The resulting data quality can be considered quite good, but it was achieved with substantial input by scientific community. The Working Group's report to the government on the implementation of female survey gave positive assessment to the performance of contracted survey agency, nevertheless, the establishment of national interviewer network was proposed. Statistical Office got interested and development of the network was started immediately in the autumn 1994.

By the beginning of the male survey of the Estonian FFS, the state interviewer network had been established and shown good results during the Labour Force Survey [Noork^{iv} and Puur 1996]. Unfortunately, the Statistical Office decided to restructure the network just before the male survey, appointing also a new director. Apparently, this appointment interrupted the established close cooperation between the FFS Working Group and the interviewer network. Consequently, various difficulties started to amount, and as a result, the fieldwork period was prolonged about three times. On the whole, the male survey of the Estonian FFS suffered from the lack of professional skills of the interviewer network even more than the female survey, carried out three and a half years earlier.

The third feature of the survey environment concerned funding, more precisely the absence of practices for funding long-term projects from the government budget. The long-term projects need to be approved each year, cumulating the risk of being stopped or cut. This risk is increased by the very strong competition for budgetary resources, as in transition countries many important fields experience insufficient funding. The inclusion of the nationwide survey like FFS in the budget formed the precedent of its kind, demanding extensive clarification efforts by the Working Group with officials responsible for budget formulation. The Estonian FFS as a separate item was several times discussed in the Budget Commission of the Parliament, and voted in the full assembly. Finally, the survey was approved but the actual allocation was cut to one third of the needed resources. This enabled to finish the preparatory stage, which had been started on the voluntary basis, and to complete the fieldwork of the female survey.

For the continuity of funding of the Estonian FFS the support of the Governmental Commission of Population has been essential. Contented with annual reports by the Working Group and intermediate results of the Estonian FFS, the Commission gave its support during budgetary discussions, thus reducing the risk of halting the Project. The activity of Commission subsided in 1997 and securing funds became more complicated. Some tasks, including the completion of the fieldwork of the male survey, have been financed from other sources and/or accomplished on voluntary basis. Also, no funds for analytical work have been allocated. From the current year 2000, the government funding of the FFS project has been stopped.

Fundraising for the preparation of the second stage of the European FFS *Gender and Generations* has to start from scratch.

The fourth feature of the survey environment is related to the efforts to build up an integrated system of survey statistics, as well as to secure the linkage with census and vital statistics in Estonia. Being the first national survey of its kind, it was natural to consider this wider task when preparing and implementing the FFS. All numerous sociological surveys conducted earlier could not set up such task: there was no other alternative but to draw the samples on enterprise/organisation employee' or electors' lists with limited coverage and virtually no personal characteristics. Under such conditions, the individual-level data consistency of various surveys could not be aimed and applying census and vital records for scientific purposes was legally prohibited.

The sampling of the Estonian FFS is based on the recent census records, thus securing the relevant linkage. From 1992, the linkage with birth and death registers has also been secured. Still, it is more important to stress the comparability of FFS with other national surveys, namely Labour Force Survey, Health Survey and National Minority Survey. Besides sampling, the data comparability builds on harmonised definitions applied in all the referred surveys. This aim was considered throughout all the stages of those surveys, starting from the programme formulation. Still, the surveys conducted by Statistical Office are not integrated, because of the prevailing practice to consider each survey as a separate undertaking.

The above-mentioned features of survey environment, in which the Estonian FFS was planned and carried out, added numerous daily tasks to ensure the comparability of the result with other countries participating in the European FFS. Under normal circumstances, such tasks are not addressed in the framework of a particular survey and it is no surprise that their consideration prolonged the duration of the project period. Nevertheless, the work was considered unavoidable, particularly from the perspective of developing a modern statistical system, and in that process, the Estonian FFS as well as other major national surveys had their contribution to make.

2. Extensions towards Integrated Survey Programme

Among social surveys, an essential feature of the FFS which needs to be underlined, is its broad substantive scope. Although the primary focus of the survey lies with fertility and family dynamics, these processes are not considered *per se* but seen in multitudinous interactions with other life domains. Accordingly, aside childbirth and partnerships, the FFS programme includes socialisation environment and departure from parental home, fertility regulation, education and work, residential mobility and housing conditions etc [UNECE 1992]. The coverage of all major individual careers is understandably not accidental but represents a growing recognition of interdependencies between societal processes. In FFS, the information on individual life careers is integrated in a modern event history framework which has become regarded as universal and the most appropriate reflection of reality, embedding a principal extension of analytical possibilities [Blossfeld, Hamerle and Mayer 1989; Courgeau and Lelièvre 1992 and many others].

In case of Estonia, additional reasons strengthened the broader focus of FFS. Most importantly, this refers to the scarcity of resources available for survey statistics. In a small country, undergoing an economic transition, only a limited number of nationally representative surveys can be afforded. To supply essential information for decision-making and research under these circumstances, individual surveys must consider a broader range of tasks than typical to bigger countries with extensive programmes of survey statistics. Regarding the Estonian FFS, the need for greater integration has been visible in two major dimensions of the survey — the programme and target population. In order to provide a more complete insight into the post-war population development and establish an appropriate information basis for population-related policies, on several issues the programme of the Estonian FFS was expanded compared to core questionnaire.

The first module with extended programme concerned pregnancy/abortion history. The immediate necessity for corresponding information was especially pronounced because of very high level of induced abortions and limited knowledge on related behavioural patterns [Anderson *et al* 1993; EMSB 1996]. First of all, the Estonian FFS was expected to cast light on social circumstances which led to different pregnancy outcomes. Also, the family formation was paid attention to in this context. In Estonia, like in other countries with established post-transitional pattern of population reproduction, generational replacement is to a large extent determined by the prevalence of third births. They have been rather frequent among native-born population in 1968-1990 but decreased substantially during the recent decade [Katus 1997]. The parity-specific approach to matrimonial, sexual, abortion and fertility histories was expected, among others, to outline the pathways which have led to the three-child family.

Regarding partnerships, extension concerned the definition of events. In particular, both in case of formation and dissolution of partnerships, additional to usual *de jure* / *de facto* dimension (registered marriage/consensual union) two other aspects were distinguished. One of them referred to sexuality which defined regular and/or long-term dyadic relations, being considered neither marriage nor consensual union by partners involved. The second of them referred to habitual dimension which defines marriages and consensual unions with partners living apart for extended periods. The entry into and exit from unions was addressed separately across this dimension. Separate specification of habitual dimension was considered particularly important in view of administrative allocation of housing under centrally planned economy and following principal transformation in the 1990s. Additionally, the range of partner characteristics was expanded.

The second optional module included in the programme concerned the migration and residential mobility which have strongly shaped the post-war population development in Estonia but are rather inadequately covered by vital and/or census statistics [Katus, Puur and Sakkeus 1998; Katus *et al* 1998]. Being part of the former Soviet Union, Estonia had been exposed to very high in- and out-migration volumes of international migration. Estonia has also been characterised by the administratively directed internal migration, which has resulted in noticeable regionalisation of population development. The FFS programme also considered the need to apply internationally comparable definitions of long- and short-term as well as long- and short-distance migration which have not been possible on the basis of other data

sources. The programme of the Estonian FFS was somewhat extended also with respect to parental home and socialisation environment.

On the other hand, some attitudinal questions included in the core questionnaire appeared inapplicable, particularly due to the rapid transformation of society. For example, even questions on childbearing plans, rather standard in demographic surveys, yielded clearly misleading picture. Applying the most conservative definition of expected births, childbearing plans predicted the increase of Estonian fertility trend. However, at the same time in reality the opposing trend has been strongly prevailing and the period fertility has undergone the steepest decline over the century. In other words, the female cohorts in fertile ages have behaved rather differently compared to their plans, revealed by the survey. By the present time the divergence between actual fertility trend and the attitudes measured by the survey has continued to increase.

Aside the programme of the survey, more integrated approach was applied also in the definition of target population. Firstly, the Estonian FFS extended the cohort range of the target population beyond the fertile age, differently from the standard recommendation. The upper age limit of the target population was increased for twenty years, i.e. up to the birth cohort of 1924. The extension of the cohort range was motivated, among others, by the existing gaps in population information for the immediate post-war decades. This period of 15-20 years has been very poorly covered by vital statistics, neither have there been any census nor representative surveys which could supply the data on respective cohorts. At the same time, these decades were marked by extensive discontinuity in demographic and social development of Estonia such as sovietisation, absence of post-war baby boom etc. And last but not least, leaving to future the data collection on older cohorts, who by the time of the survey had already reached age 70, would have probably led to eventual loss of information.

The second extension of target population concerned the inclusion of population of foreign origin, i.e. immigrants and their second generation. The necessity of such an extension stems from post-war immigration. According to the last census, the proportion of foreign-borns in total population has exceeded one quarter, together with the second generation, the share of population of foreign origin comes close to two fifths of the total population [Sakkeus 1991; 1996]. In the formation of foreign-born population, two major immigration waves can be distinguished, reflecting the principal expansion of Estonia's migration hinterland, and implying the growth in the heterogeneity of foreign-born population. In case of Estonia, the omission of the referred population would have clearly prevented the FFS being representative to total of the country.

Apart from its size, in case of Estonia the population of foreign origin has systematically been distinguished from the native-borns in terms of demographic and social behaviour, the difference dating back to the timing of demographic transition [Katus and Sakkeus 1993; Puur 1998; Viikberg 1999]. This aspect has been particularly important for analytical stages of the survey and the presentation of results: because of the diverse, sometimes even opposing trends between the populations of native and foreign origin, consideration of total population without disaggregation would have yielded rather misleading picture of population development [UNECE 2000].

3. Potential Risks and Quality-Supportive Procedures

The extension of survey programme and target population in the mentioned directions implied several risks to the data quality. The inclusion of optional modules and questions additional to the core questionnaire implied prolonged length of interview. In the stage of questionnaire development and pretests, the average length of interview was expected to be around hour and a half, in reality it took 87.5 minutes in female survey and 107 minutes on average in male survey. The proportion of interviews longer than 2.5 hours accounted for 13.7 per cent and 32.6 per cent respectively. One distinguished foreign expert of the FFS Working Group had expressed particular concern in this respect, pointing to potential overburdening of respondents, decline in the accuracy of responses, and a threat, that substantial number of respondents could quit interviews or refuse to answer. Specific concerns were also expressed about the inclusion of modules related to fertility regulation and abortion. Related questions were considered too sensitive for respondents, provoking part of them to interrupt the interviews and refuse, or to yield unreliable answers.

Regarding the extension of target population, the concerns were also expressed. In case of older cohorts, the recall problems were expected because of weakened memory and long temporal distance between the events and interview. Considering the event history methodology, the critical question was whether the older respondents were able to provide consistent answers on exact year and month of events, particularly those which concerned not the respondent her/himself but related persons. Regarding the population of foreign origin, concerns addressed primarily the problem of non-response. Dissolution of the Soviet Union in early 1990s had triggered return migration, covered poorly by official statistics [ESA 1998d]. This hampered evaluation of its impact in advance, confounded by the political tension in connection with restoration of Estonian Republic.

The risks related to the extension of survey programme and target population were addressed by the FFS Working Group. The position of FFS as the first nationally representative survey after the restoration of independence meant that there was no earlier experience on which the Working Group could rely. To safeguard the data quality of the survey, careful attention was paid to the elaboration of survey procedures, starting from the development of data collection instruments. The survey was preceded by three pretests. The inclusion of the population of foreign origin required the questionnaire and other survey instruments to be developed in two languages, as well as splitting the interviewer training and field procedures. In order to inform the public, including the respondents, the Working Group launched a media campaign, holding press conferences, preparing newspaper articles and TV/radio broadcasts.

In both female and male survey, the training of interviewers was considered of key importance. The FFS Working Group conducted a series of training seminars, focusing primarily on specific requirements of event-history survey (supporting the recall of events, checking the consistency between parallel careers etc). At the end of the seminar each interviewer had to pass a special test, after which (s)he was approved for the task. The test foresaw an interview with one of the Working Group members taking the role of respondent. Such "respondent"

had a complex life history, and the task of the interviewer was to detect deliberately inserted inconsistencies. Similar quality-supportive procedures were devised also for subsequent stages of the survey. To ensure control over the progress of fieldwork, the Working Group employed its own coordinator for field operations, parallel to supervisors in contracted network.

Coding and data entry, which were combined with primary consistency control, were detached from survey organisation and operated by a separate unit under close supervision of the Working Group. Each case an inconsistency was encountered, life history of the respondent was analysed. The Working Group summarised the typical errors after the receipt of about first hundred questionnaires and provided interviewers operative feedback. This measure evidently had its effect, as in the course of the survey, the need for clarification decreased by nearly one third.

After the completion of fieldwork, the evaluation of data quality proceeded across two major directions: (1) representativeness of the sample and (2) completeness and consistency of responses. Followingly, both directions are examined.

3.1. Representativeness of the FFS Data

A survey sample must be representative to target population, regarding its demographic, social and regional composition. Although the complete identity of the two can never be reached, the lack of systematic differences, exceeding random variation, must be secured. In the following, the assessment of representativeness of the Estonian FFS is based on the comparison of the composition of target, respondent and non-respondent population. The comparison builds on a set of personal characteristics available from the sampling frame (the last population census) for both respondents and non-respondents.

Considering the extension of target population, neither older female nor male cohorts demonstrated increase in non-response. The reality was the opposite, and combining different reasons of non-response, older cohorts displayed lower non-response rates, accounting for one third of fertile age cohorts among males (Table 1). The decomposition of non-response revealed two major causes, rather equally important, refusals and non-locations. Due to poor accuracy of residence registration, the latter reason appears particularly high in Estonia. Non-response due to ill health had been expectedly higher in older cohorts but it did not determine the general pattern.

The actual problem of non-response related to older cohorts stems from the selectivity. The birth cohorts defining the target population are represented by survivors, those who have died, are necessarily been omitted. While in younger cohorts potential selectivity is relatively weak, it keeps strengthening towards older cohorts, reflecting the age-pattern of mortality. It is evident that survivors and non-survivors of each cohort are selective with respect to life histories, introducing the bias in data. Due to prolonged mortality stagnation, extensive gender gap in life expectancy, the impact of the referred selectivity is likely stronger in the Estonian FFS than on average among the European participant countries. The importance of the referred problem should not be ignored, however, alternative designs (prospective studies) will hardly present a viable option for FFS.

Regarding the distinction between native- and foreign-origin population, the non-response appeared higher among the latter, in line with the experience of many immigration countries. Similarly to cohort dimension, heterogeneity in non-response turned to be significantly higher among males. In addition to more frequent temporary absence of foreign-origin population, higher non-response rate of foreign-origin population was attributable to less accurate address information. Due to heavy concentration of foreign-origin population in urban areas, the differential pattern of non-response has likely influenced also the urban-rural differentiation. At the same time it is interesting to note that refusal rate has been fairly equal between native- and foreign-origin population, suggesting no major discrepancy in the attitude towards the survey.

Table 1. Representativeness, Estonian FFS

	Female survey				Male survey			
	Target population	Respondents	Non-respondents	Non-response rate	Target population ¹	Respondents	Non-respondents	Non-response rate
1969-73	9.8	9.2	10.2	13.8	11.8	11.5	15.2	27.1
1964-68	9.4	9.9	10.5	13.5	10.7	10.7	11.9	23.7
1959-63	11.0	10.1	9.3	13.6	12.1	11.5	13.8	25.0
1954-58	11.4	9.9	10.4	14.6	12.0	11.5	13.1	22.9
1949-53	10.9	9.6	7.4	10.4	11.2	11.1	13.1	22.1
1944-48	9.1	9.6	11.6	15.3	9.0	9.3	9.3	20.6
1939-43	9.4	10.9	9.1	12.4	9.2	9.3	9.5	20.9
1934-38	9.8	10.6	10.5	13.5	9.1	10.0	6.4	13.1
1929-33	9.7	10.0	10.2	10.6	8.3	8.8	4.3	8.7
1924-28	9.6	10.1	10.7	10.2	6.6	6.3	3.3	9.9
Native	59.8	62.7	52.9	12.5	58.8	63.9	51.7	17.1
Foreign	40.2	37.3	47.1	13.2	41.2	36.1	48.3	25.9
Married	62.2	63.1	59.1	12.2	64.7	68.6	55.5	17.4
N-married	19.9	18.9	22.9	14.3	26.7	24.9	32.8	25.7
Widowed	7.0	7.1	6.8	11.2	1.3	1.0	1.7	27.0
Divorced	10.8	10.8	11.1	14.1	7.0	5.4	10.0	30.0
Childless	24.2	22.6	30.8	16.3	—	—	—	—
1 child	27.1	26.6	30.0	14.2	—	—	—	—
2 children	34.9	35.0	29.3	11.3	—	—	—	—
3+children	13.8	15.8	9.9	15.6	—	—	—	—

¹ In the sample frame (the latest census) the information on number of children was collected only on females and the comparison between the target and survey population is not possible across this characteristic.

Across marital status, the highest level of non-response have been observed among never-married. Closely correlated with their younger age distribution, never-married individuals appeared more mobile and therefore their non-response was primarily related to frequent temporary absence from the place of residence. Regarding other groups, the pattern appeared somewhat different among males and females. Among females, observed differences between married, divorced and widowed were very small. Among males, those married had

significantly higher probability of being interviewed than widowers and divorcees. Due to the availability of information from the sampling frame, for females it was possible to assess the differentiation of response also by parity. The data reveal a slight positive association between the participation in the survey and the number of children ever-born. The information on reasons on non-response suggests that the pattern may have partly resulted from lesser mobility of higher parity women, and partly from their increased interest to participate in the survey. The parity-specific analysis revealed acceptable consistency of Estonian FFS with vital registration which was not the case with all FFS countries [Barkalov and Dörbritz 1998].

The structure of target population, respondents and non-respondents is compared in Table 1. Across cohorts, the data reveal quite close similarity between the target population and respondents, although slight underrepresentation of the younger and overrepresentation of the older cohorts can be observed. Regarding the distinction between native- and foreign-origin population, the FFS respondents are characterised by somewhat lower proportion of immigrants and their descendants than represented in the sampling frame. This discrepancy could, to a considerable extent, reflect the impact of migration flows that have occurred in Estonia after the 1989 census to which the structure of target population refers. Whichever the reason, potential underrepresentation of foreign-origin population should not be exaggerated as the development and behavioural patterns of these populations necessitate their separate analysis. Returning to other principal characteristics (marital status, regional distribution, educational attainment) the data of the Estonian FFS can be considered also representative and require no weighting. The Estonian FFS census-based sample frame supports the evaluation of representativeness against more than 50 characteristics.

3.2. Completeness and Consistency of the FFS Data

The second major dimension of survey data quality refers to completeness of responses, collected from sampled individuals who were eventually interviewed. This kind of non-response should not be underestimated as even low frequency of omissions at individual items tends to cumulate and limit sharply the amount of information which can be effectively used in the analyses. Such omissions may occur for different reasons, however, for the purposes of classification two broad categories should be distinguished. The first category of item-specific non-response stems from unwillingness of respondents to provide answers for specific questions. Given the intimacy of spheres addressed by the survey, particularly extended modules on fertility regulation, pregnancy/abortion and sexuality, these concerns required careful attention.

Analysis of responses by individual questions and modules revealed that item-specific refusals were encountered only in exceptional cases. Regarding even most delicate issues, in female survey there were only three respondents who refused to provide information on the number of sexual intercourse during the last four weeks preceding the interview. Regarding induced abortions, only one person explicitly refused. In male survey, altogether eight respondents were not willing to provide information on the number of sexual intercourse during the last four weeks and three respondents refused to talk about their experience with commercial sex or

male partner. Understandably, the referred numbers are too small for any generalisation across respondent characteristics.

Being an indication of fairly low schisma, negligible level of item-specific refusals, however, does not necessarily imply the completeness of reporting. Sensitive and/or uncomfortable issues can be easily evaded less explicitly, for example, leaving certain events or circumstances just unrecorded. From the viewpoint of data quality, the extent of kind of hidden evasion is much more difficult to assess. To assess the extent of such concern, a special validation study was conducted in connection to the Estonian FFS [Anderson *et al* 1994]. In that framework, the truthfulness of survey responses on abortion was checked against objective information from external source. Differently from prevailing practice in such studies, the verification of responses was accomplished on individual level. The sample for the study was drawn from patient records of women who had an abortion which served as an external standard for assessing the completeness and accuracy of responses.

The second type of item-specific non-response stems from inability of respondents to provide answers for specific questions. From the viewpoint of event history analysis, the most critical was the issue of completeness of event dates (exact month and year). As noted above, in case of Estonian FFS the task was complicated by the extension of survey programme and the inclusion of older cohorts. Nevertheless, the analysis performed after data collection reveals generally high quality of date information. To measure the accuracy of responses, the number of incomplete dates was related to the total number of questions on event-dates (altogether close to 250 thousand). In female survey, respondents could not provide complete information to slightly less than 0.4 per cent of questions addressing the year of event. Regarding the month of event, the proportion of incomplete answers was understandably higher, accounting to 1.6 per cent. In male survey, the year of the event was missed in 0.9 per cent of relevant questions and month in 2.3 per cent of questions. At least partly, the difference between male and female survey should be ascribed to varying performance of contracted survey organisations [Pungas 1999].

To assess the impact of date omissions to FFS analyses, date-specific non-response has been decomposed according to modules of the questionnaire, on one hand, and characteristics of respondent, on the other hand. To better demonstrate the variation, an alternative measure of data completeness has been introduced: the proportion of questionnaires with at least one year/month missing. It should be noted that the referred measure appears very conservative and tends to amplify the limitations of the data: even one missing piece of information from the average of 30-40 event-dates provided by the respondent is enough to drop the case from the count. Table 2 presents the variation of date-specific non-response by modules of the questionnaire in female survey, applying parallelly both measures.

The data reveal that the lowest level of non-response has been characteristic to partnership and birth history. Similarly low levels, particularly considering the year of event, appeared in migration and residence history, education and work, and household history. In all these main careers, the proportion of answers with missing year was lower than 0.1 per cent. Regarding the month of event the proportion of missing information varied between 0.1 and 0.6 per cent in these modules. On the other extreme, the highest level of date-specific non-response concentrated in parental home module. Typically, the respondents failed to recall the exact

birth and death dates of their siblings which had occurred in infancy or childhood. In case of male survey, the concentration of omissions in parental home module appeared even more expressed [EKDK 1999a].

The module of pregnancy/abortion history positions between parental home and other modules. Considering the proportion of events with missing information, data quality of the pregnancy/abortion module stands closer to parental home module with 0.9 per cent of all recorded years and 4.1 per cent of months not specified. Applying a more conservative measure, the proportion of respondents with at least one missing response in respective module, the prevalence of date-specific non-response drops to 0.4 per cent (year of event) and 2.1 per cent (month of event). According to these figures, particularly regarding the year of event, data quality of the pregnancy/abortion module appears closer to other major life career modules.

This exceptional and at first glance rather paradoxical result — more conservative measure yielding a better estimate of data quality — can be explained by specific parity distribution of abortions together with other pregnancies not resulting in live birth, compared to other life careers. In particular, the proportion of women with no abortions tends to be manifold higher compared to those having no event in partnership, birth, household, migration, education and work history. The inclusion/exclusion of zero parity group from the denominator of alternative item-specific response rates explains much of the referred anomaly.

Table 2. Date-Specific Non-Response Rate by Questionnaire Modules, Female Survey of the Estonian FFS

	Proportion of date-questions with missing		Proportion of respondents with at least one missing	
	Year	Month	Year	Month
Partnership history	0.0	0.2	0.2	0.8
Birth history	0.0	0.2	0.1	0.5
Pregnancy history	0.9	4.1	0.4	2.1
Household history	0.0	0.6	0.1	0.8
Migration history	0.0	0.4	0.0	0.9
Education/work history	0.0	0.1	0.1	0.6
Parental home	1.2	4.9	4.3	17.1

To assess the impact of the extension of target population and present the inter-group variation of date-specific non-response, more sensitive measure was applied. Judging upon the proportion of respondents with at least one missing event-date, date-specific non-response expectedly increased towards older cohorts. The omission of event-dates reached maximum in the 1924-1928 birth cohort, exceeding the average level in currently fertile age-span more than two times (Table 3).

The observed cohort pattern date-specific non-response should be explained by two different factors. Firstly, the increase in the proportion of respondents with at least one event-date missing stems from the accumulation in the number of life events experienced by the

respondent. The referred impact concentrates clearly in fertile age, for example between cohorts 1969-1973 and 1949-1953 the proportion of respondents with date-specific non-response increases 3-4 times. Women in the 1949-1953 birth cohort were only 40-44 years old during the interview, and therefore, it would be inappropriate to relate the pattern to the difficulties of recall. Beyond age 50 the number of life events stabilises, and in these age groups, recall difficulties enter the scene. Although the levels presented in Table 3 look rather high, the extent of the problem should not be exaggerated.

Firstly, both in absolute, and particularly in relative terms, the increase in date-specific non-response occurring in older cohorts appears much smaller than that observed in fertile age-span. Secondly and more importantly, the examination of the pattern by modules reveals that across cohorts, the overwhelming majority of date omissions is limited to parental home [EKDK 1995b]. The frequency of omissions in major life careers of the respondent herself remains moderate also in older cohorts. For example, in the 1924-1928 female birth cohort, in modules other than parental home, the proportion of respondents with at least one event-year missing ranged between 0.2-1.2 per cent. Regarding the month of event, corresponding percentage ranged between 1.0-2.7 per cent.

Table 3. Date-Specific Non-Response, Female Survey of the Estonian FFS

	Proportion of respondents with at least one missing	
	Year	Month
1969-73	0.1	6.0
1964-68	0.9	6.3
1959-63	3.4	10.3
1954-58	3.5	14.7
1949-53	4.5	18.1
1944-48	4.8	19.9
1939-43	6.2	26.0
1934-38	7.4	30.8
1929-33	7.9	31.3
1924-28	11.8	33.9
Native	3.9	16.4
Foreign	7.6	26.1

Good quality of information provided by older respondents is revealed also by other characteristics. Regarding the consistency of life histories, there were only slightly more interviews which required clarification in older cohorts. Similar conclusion can be reached relying on the number of edits introduced during data cleaning and verification of survey responses against individual-level census records of the respondents. It is interesting to note that differently from some expectations, the same analysis revealed no significant deterioration of data accuracy parallel to the increase in interview duration [EKDK 1995a]. In other words, lengthy interviews do not necessarily represent interviewing difficulties, but could, on the contrary,

indicate particularly good cooperation by the respondents and willingness to provide the interviewer with various details. Notably, none of the lengthy interviews had been quitted in the middle, although two per cent were interrupted and completed at some later date.

Regarding the population of foreign origin, some deterioration in the completeness of information can be observed. In view of previous discussion and younger age structure of immigrants, this finding may look somewhat surprising. Evidently, it should be explained with diverse historical experience of the two populations, dating back to the XIX century. Already during that period, native population featured virtually complete literacy which facilitated the transfer of information from parents to children. Among immigrant population, the

heterogeneity of educational attainment has been much greater and illiteracy has been not exceptional in the cohorts born in the early XX century. Additionally, considerably later demographic transition in the home regions of immigrants has meant larger families as well as lower survival of siblings and parents which has a direct impact on completeness of the data.

To sum up, the extension of survey programme and target population did not imply any significant compromise in data quality of the Estonian FFS. It can be assumed that the extension of survey programme has, on the contrary, contributed to the increase in consistency of life histories and data quality in general. From analytical viewpoint, the discussed omissions impose no major restrictions on event history analysis. Aside the referred conclusion, the analysis revealed an important quality impact of network and interviewer performance, however, these results go beyond the present paper and are reported elsewhere [EKDK 1999a].

4. FFS in the System of National Surveys

The Estonian FFS Working Group considered the survey as a constituent of national statistical system. Therefore, scientific community proposed the FFS to be prepared and implemented with wider aim of building up the survey statistics in Estonia. The proposal was acknowledged by the Governmental Commission of Population and the FFS Working Group followed this task throughout all the stages. Due to scarcity of funding, the excellent cooperation among scientific community was mostly accomplished on voluntary basis. Among others, the contribution of FFS to the elaboration of national definitions and concepts for various societal fields, compatible with international standards, should be underlined. Those concepts considered the applications in vital and census statistics, as well as the requirements of policy-oriented research.

In 1993 the Governmental Commission of Population adopted the programme for integrated system of national surveys, which foresaw in the first round the completion of eight national surveys in the 1990s. The system of national surveys had to secure, together with census and vital statistics, the basic information on the most essential population and social processes [Katus *et al* 1993]. The female survey of the Estonian FFS was followed by the Labour Force Survey 1995 [Noorkõiv and Puur 1996], Health Survey [Leinsalu *et al* 1998], National Minority Survey [EKDK 1999b] and the male survey of the Estonian FFS.

The surveys included in the programme were chosen to cover the major societal fields. Each survey was given a dual task: on one hand, the development of definitions, concepts and appropriate measurement tools in its field, and on the other hand, systematic application of definitions, already elaborated by other national surveys. In the case of Estonian FFS, the contribution of the first kind comprises the definitions related to family, household and fertility, including reproductive health. Regarding the other task, FFS had to consider in advance the basic requirements of the following surveys, in order to secure the comparability within the system. In view of that, the FFS Working Group had invited the key researchers, responsible for the following surveys to participate in its activities.

For the needs of national survey statistics, FFS Working Group developed the sample frame and sampling procedures, principles of interviewer training, organisation of fieldwork, coding, data entry and editing routines. The following national surveys draw on the experience of FFS and extended it in their specific directions, contributing to the formation of integrated survey statistics. In that process, the essential role of working groups for planning and implementation in each national survey should be underlined. The working groups brought together the efforts of all interested institutions and individual researchers. The membership of such working groups changed from one survey to another with some core-members participating in all groups and securing the comparability across individual surveys. Generalising the experience of the 1990s, only those surveys which had applied the institution of working group, have been integrated, leaving others carried out as separate undertakings, including those by Statistical Office, outside the system.

As a contribution to general statistical system, equally important has been the linkage with vital and census statistics. The individual-level linkage of survey responses to census records,

safeguarded by sampling frame, has allowed to analyse the content of definitions. Starting from 1992, the organisation of vital statistics records allows the linkage with survey data on the individual level, supporting the follow-up. The possibilities of individual level linkage, in addition to comparability on aggregate level, have been particularly appreciable in case of Estonia, experiencing the transition from Soviet-type statistical system to the international one, without losing the consistency of time-series on major population and social processes.

5. Comparability of Major Definitions

In the following, the definitions and concepts, elaboration of which started in the framework of Estonian FFS and continued in other national surveys, are compared. The comparison addresses the consistency of major definitions and provides a basis for integrated use of national surveys. Considering the gender composition of target population in different national surveys, male survey of the Estonian FFS is compared to male data from Labour Force Survey and Health Survey. The female survey of Estonian FFS is compared to female data from Labour Force Survey and Health Survey, and National Minority Survey, which covered the Russian and Ingerian national minority populations in Estonia.

The analysis of comparability covers the definitions of various events and statuses, giving preference to standard characteristics applied also in census and vital statistics. Aside the latter, some comparisons involve innovative characteristics which in case of high consistency across surveys, could lead to the extension of a set of standard characteristics. Understandably, reflecting the underlying regularities of demographic and social processes, all comparisons apply disaggregation by cohort/age. In the context of the present paper, the definitions of partnership- and fertility-related characteristics are paid particular attention.

Family/household status is represented by the proportion of never-married, cumulative number of partnerships (including consensual unions) over lifetime, average household size and proportion of one-person households. For female samples of the surveys, average number of registered marriages has been additionally presented. The comparison of fertility data builds on cumulative number of ever-born children, representing completed total fertility rate for cohorts born before 1950. Being one of the core subjects of the FFS, consistency of fertility data should be evaluated in a broader context of pregnancy outcome. The set of figures includes also the cumulative number of abortions. All the referred comparisons (five first graphs of Figure 1 and seven first graphs Figure 2) reveal consistency of main patterns across the surveys. At the same time, it should be noted that compared to FFS the Health Survey has indicated, for instance, lower prevalence of induced abortions, due to inferior performance of interviewer network as discussed elsewhere [EKDK 1999a].

Concerning life careers related to family and fertility, comparisons involve employment and educational attainment, represented by employment rate and the proportion of higher education population. In case of males, occupational distribution has also been added. Reflecting the underlying dichotomy of modern Estonian population, for both male and female population, the comparison involves the distinction between the popu-

Figure 1. COMPARISON OF DEFINITIONS IN MALE SURVEYS

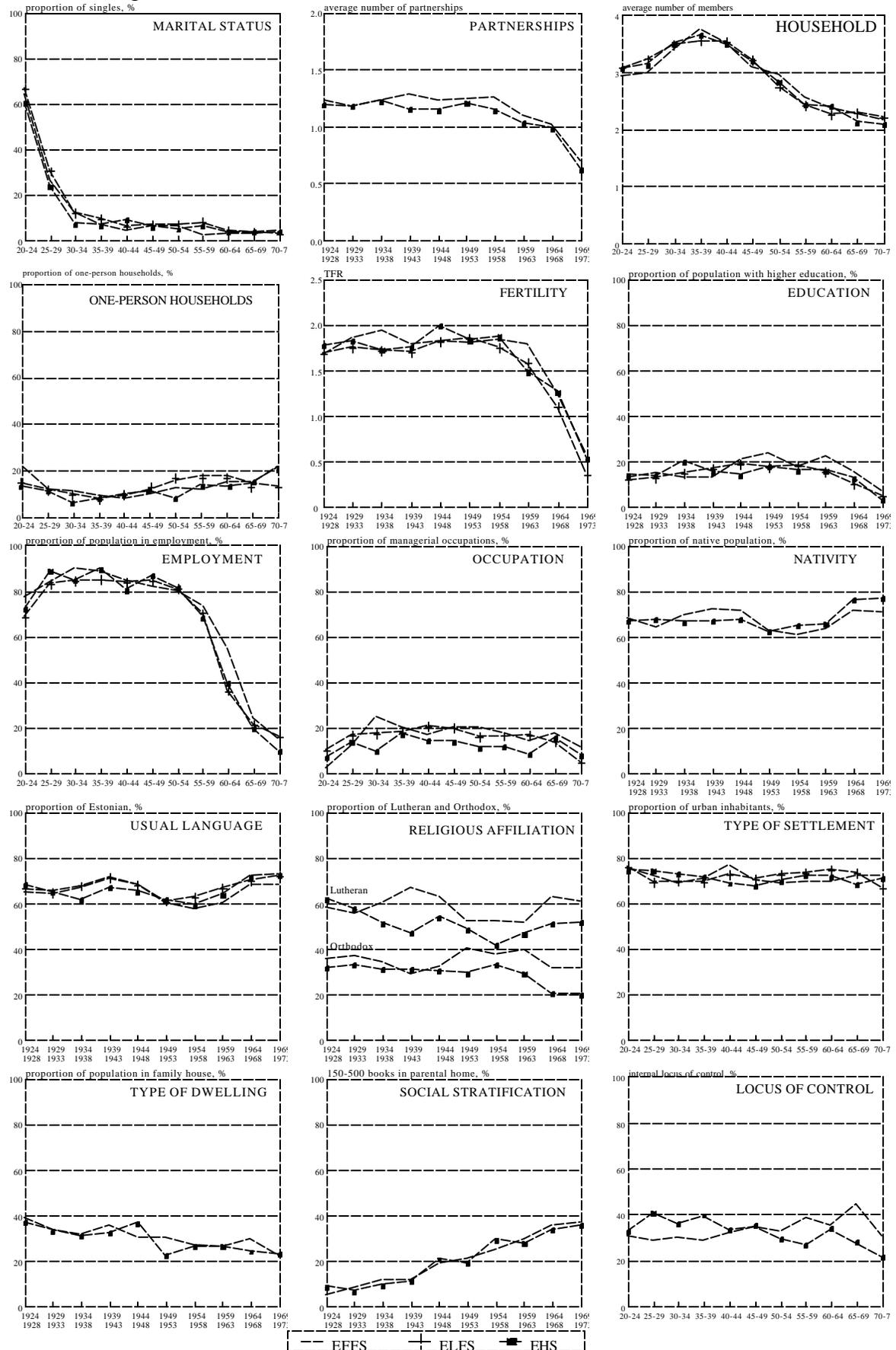
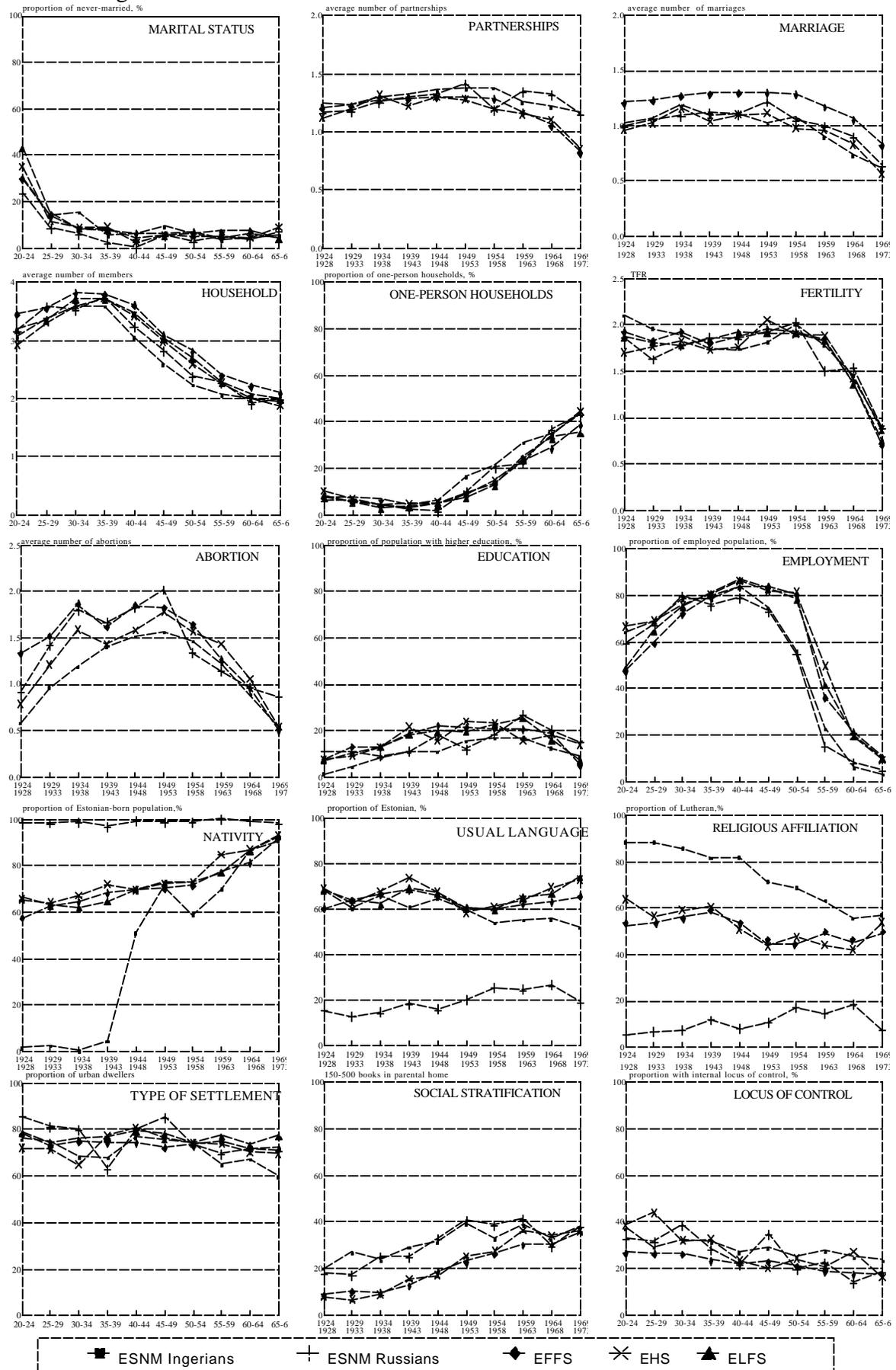


Figure 2. COMPARISON OF DEFINITIONS IN FEMALE SURVEYS



lation of native and foreign origin, supplemented by data on usual language and religious affiliation. Quite expectedly, in respect to these characteristics, National Minority Survey stands out for its specific target population which needs to be considered also across all the other comparisons. Regarding other standard characteristics, comparisons are presented for urban/rural distribution, type of dwelling, social stratification and locus of control. Alike family and fertility characteristics, consistency of other social characteristics is fairly high across surveys. Regardless of presentation by five-year cohorts/age groups, no systematic differences are observed.

6. Summary and Discussion on the Importance of the FFS

The European FFS has proved to be the innovative exercise in most if not all the participating countries. This has been particularly emphasised in the countries of Central and Eastern Europe which are in the process of reforming and/or building up their national statistical systems. Among several innovative contributions of the FFS three aspects should especially be outlined in case of Estonia, summarised below. Very likely, those aspects are playing similar role in other countries of the region.

The European FFS builds methodologically on the event history approach. In spite of being recently developed, it has been applied to a variety of processes ranging from family planning to deviant behaviour and development of business entities. Nevertheless, the European FFS has been the first attempt to apply that methodology on a broad cross-national scale. The undertaking has been quantitatively impressive with a quarter of hundred countries in the European region covered. Much more important, however, has been methodological challenge to compare the dynamic development of European populations. Event history approach, following the life events starting from parental home, unwillingly requires the consideration of diverse historical experience of nations and countries for at least half a century. Even collecting internationally comparable cross-sectional data proves to be a complex task, in case of event history approach difficulties are manifold increased.

The European FFS has been initiated and developed with active participation of major population institutes in the region. This, among others, has secured the dissemination of the best practices for the success of the survey, and comparability of collected event-history data. The countries of Central and Eastern Europe with less extensive experience with such surveys have undoubtedly been the major beneficiaries from the pan-European cooperation. For example, in case of Estonia, only one survey with event history approach had been carried out previously, focusing on urban population [Puur and Vikat 1990; Katus *et al* 1994].

Event history approach is most appropriate tool to study the cumulative processes, particularly such as fertility. Fertility and other similar processes, when studying cause-specific relations, are dependent on the chain of previous events, and less on the momentary circumstances. In this respect, the most important innovative aspect of the life course methodology applied in the European FFS is the possibility, or even demand, to integrate different life careers. In comparison to cross-sectional information the event history data substantially increases the value when adding parallel histories. Firstly, the reliability of data is improved because of the possibility to apply much more complex consistency checks. Secondly, and most importantly, the explanatory capacity expands when increasing the variety of life histories. The European

FFS has undoubtedly been the first event-history survey at wide international scale with all major life careers combined into one. In case of Estonia, the extension of survey programme and inclusion of additional modules has deliberately maximised the advantage of the referred innovation.

The second important feature of the European FFS is its comparative pan-European perspective. This aspect has often been stressed by the coordinators and participants and accepted by policy-making community [UNECE 1999]. Nevertheless, it is once again important to stress this aspect from point of view of Central and East European countries, and particularly for the new countries or those having re-established their independence. Many of the referred countries have been cut off from internationally comparative perspective, with very limited possibilities to be regarded in the context of developments in Europe.

After the collapse of communism, the Central and East European countries have found their economies and social organisation lagging seriously behind, depreciating the value of comparisons with the rest of Europe. The population trends and various life careers, however, have been much more in correspondence with long-term development, making the comparative perspective appropriate. Moreover, this perspective enables generalisations about the impact of diverse social and political environments. In that sense, for Central and East European countries the role of FFS goes far beyond family and fertility development. In case of Estonia, the extension of target population to older cohorts, accumulating the long-term impact of social discontinuity following the WW II, has particularly been motivated by this perspective.

The third feature of the FFS, in the context of Estonia and other countries with social discontinuity, is its contribution to the development of national statistical system. Due to the specific societal and statistical environment, the preparation and implementation of the FFS had to go beyond an individual data collection exercise and aim at much wider range of activities than considered under the normal circumstances. In case of Estonia, the FFS became the first nation-wide undertaking of its kind, aimed at building up survey statistics and statistical organisation in general. From the organisational point of view, the integration between researchers and governmental officials was established. Independent from medium-level administrative bureaucracy, the FFS Working Group was delivered the responsibility for all stages of the survey planning and implementation. In the conditions of underdeveloped official structures of transitional society, such an organisation proved to be the most efficient.

Being a successful exercise, the FFS in several ways contributed to the emergence of survey statistics in Estonia. However, from the several years distance it is also necessary to acknowledge major shortcomings and even failures in that direction. The principal difference of the FFS from previous practices proved to be the methodological work. It has been fully accomplished by scientific community joined in the Working Group i.e. outside regular structure of statistical institution. By today, several surveys carried out by Statistical Office have shown resistance to absorb the gained methodological experience and preference to much simpler, less time- and work-consuming procedures. In other words, the real cooperation between scientific community and statistical administrators has not been established and each new national survey needs to start from the very beginning.

Secondly, data quality issues are still widely and wildly underestimated. In that context, it has been common to refer to the experience of the Estonian FFS as something irregular which does not deserve to be followed. Such practices as forming a working group uniting the available intellectual capacity, providing it with the responsibility for all stages of the survey, decision-making authority, including the allocation of relevant funds etc which were essential in securing the data quality, are regarded as useless, or even harmful parallelisms to the nomenclatural structure of statistical institution.

To sum up, the experience of the Estonian FFS as system-creative and innovative exercise has been confronted with old working habits, and has lost. As a result, survey statistics, although newly established, requires to be principally reformed and the broader aim of the Estonian FFS to build up national survey statistics, still remains to be achieved.

References

- Anderson, Barbara, Kalev Katus, Allan Puur, Brian Silver (1993). Characteristics of Women Having Abortions in Estonia. *International Population Conference*. Vol.1. Montreal, IUSSP, pp.215-234.
- Anderson, Barbara, Kalev Katus, Allan Puur, Brian Silver (1994). The Validity of Survey Responses of Abortion. *Demography*. Vol.31, No.1, pp.115-132.
- Anderson, Barbara, Kalev Katus, Brian Silver (1994). Developments and Prospects for Population Statistics in the Countries of the Former Soviet Union. *Population Index*. Vol.60, No.1, pp.4-20.
- Barkalov, Nicholas and Jürgen Dörbritz (1998a). *One or Many Patterns? A Comparative Study of Period Parity Progression Ratios in FFS Countries*.
- Barkalov, Nicholas, Jürgen Dörbritz, Sharon Kirmeyer (1998). *Fertility Characteristics of the Baltic Populations: a Parity-Progression View*. Paper presented to the International Conference Estonia and the European Union: Estonia on its Way to a Changing Europe, November 4-6. Tallinn.
- Blossfeld, Hans-Peter, Alfred Hamerle, Karl Ulrich Mayer (1989). *Event History Analysis. Statistical Theory and Application in the Social Sciences*. Hillsdale, New Jersey, Lawrence Erlbaum Associates Publishers.
- Courgeau, Daniel, Eva Lelièvre (1992). *Event History Analysis in Demography*. Oxford, Clarendon.
- Darsky, Leonid (1986). *Evolucija rozdaemosti v SSSR po dannym obschesojuznyh vyborochnyh obsledovanii*. Paper presented to Deuxieme Seminaire Franco-Sovetique de Demographie, September, 15-19. Suzdal.

- EKDK (1995a). *Estonian Family and Fertility Survey. Methodological Report*. RU, Series A, No.39. Tallinn, EKDK.
- EKDK (1995b). *Estonian Family and Fertility Survey. Standard tabulations*. RU Series C, No.6. Tallinn, EKDK.
- EKDK (1999a). *Estonian Family and Fertility Survey. Methodological Report of Male Survey*. RU Series A, No.40. Tallinn, EKDK.
- EKDK (1999b). *Estonian National Minority Survey. Methodological Report*. RU Series A, No.41. Tallinn, EKDK.
- EMSB (Estonian Medical Statistics Bureau) (1996). *Estonian Health Care 1990-1995. Some Facts about Health Care in Estonia*. Tallinn, EMSB.
- ESA (Statistical Office of Estonia) (1998a). *Estonian Labour Force Survey 1997. Methodological Report*. Tallinn-Viljandi, Eesti Statistikaamet.
- ESA (Statistical Office of Estonia) (1998b). *Household's Income and Expenditure Survey 1997*. Tallinn, Eesti Statistikaamet.
- ESA (Statistical Office of Estonia) (1998c). *Täiskasvanute koolituse uuring. Metoodiline ülevaade*. Tallinn, Eesti Statistikaamet.
- ESA (Statistical Office of Estonia) (1998d). *Social Trends*. Tallinn, ESA.
- Katus, Kalev (1997). Post-Transitional Fertility: Case of Estonia. *K.Katus, V.Stankuniene and A.Vikat (Eds). Demographic Development in Baltic Countries*. Special issue of Revue Baltique. Vilnius, pp.49-63.
- Katus, Kalev, Luule Sakkeus (1993). *Foreign-Born Population in Estonia*. RU, Series B, No.19. Tallinn, EKDK.
- Katus, Kalev, Allan Puur (1993). *The 1989 Estonian Population and Housing Census: Data Description*. RU, Series A, No.33. Tallinn, EKDK.
- Katus, Kalev, Jüri Kõre, Marje Pavelson, Allan Puur, Luule Sakkeus (1993). *Individaalstatistika ümberkorraldus Eestis*. Isikuandmenõukogu aruanne. Tallinn.
- Katus, Kalev, Allan Puur, Luule Sakkeus, Andres Vikat (1994). *Data of the Tallinn'88 Survey*. RU, Series B, No.22. Tallinn, EKDK.
- Katus, Kalev, Allan Puur, Luule Sakkeus (1998). Data Quality in the Former Soviet Union. *Migration: European Journal of International Migration and Ethnic Relations*, No.29/30/31, pp.119-133.
- Katus, Kalev, Anne Herm, Marek Kupiszewski, David Powell, Philip Rees, Luule Sakkeus (1998). *Internal Migration and Regional Population Dynamics in Europe: Estonian Case Study*. RU, Series B, No.37. Tallinn, EKDK.

- Leinsalu, Mall, Marina Grintshak, Rivo Noorkõiv, Brian Silver (1998). *Eesti Terviseuuring. Metodoloogiaülevaade*. Tallinn, EKMI.
- Pungas, Enel (1999). *The Impact of Interviewer Characteristics on Data Quality: Experience from Estonian Family and Fertility Survey*. Paper presented to the European Population Conference. The Hague, August 30-September 3 1999.
- Puur, Allan, Andres Vikat (1990). Elutsüklilised erisused Tallinna perekondade demograafilises koostises ja ainelises kindlustatuses (ankeedi Tallinn'88 andmetel). RU Sari A, No.26. Tallinn, EKDK.
- Noorkõiv, Rivo, Allan Puur (1996). *Estonian Labour Force Survey 1995: Experience from Retrospective Data Collection*. Paper presented to the European Conference of Statisticians. Paris.
- Puur, Allan (1998). Eesti Põlissuuring: lähtekohad ja põhijooned. *Etnilised vähemused Eestis, nende tänased probleemid ja tulevik. Mustvee II konverents*. Jõgeva Maakonnavalitsus, lk.85-99.
- Sakkeus, Luule (1991). *Post-war Migration Trends in Estonia*. RU Series B, No.15. Tallinn, EKDK.
- Sakkeus, Luule (1996). Estonia. T.Frejka (Ed). *International Migration in Central and Eastern Europe and the Commonwealth of Independent States*. UN, New York and Geneva, pp.57-65.
- Tuma, Nancy, Michael Hannan (1984). *Social Dynamics. Models and Methods*. New York, Academic Press.
- UN ECE (1992). *Fertility and Family Studies in the Countries of the ECE Region. Questionnaire and Codebook*. New York and Geneva, United Nations.
- UN ECE (1999). *Population in Europe and North America on the Eve of the New Millenium: Dynamics and Policy Responses*. Regional Population Meeting. 7-9 December Budapest. New York, United Nations.
- UN ECE (2000). Fertility and Family Surveys in the Countries of the ECE Region. Standard Country Report. Estonia. *UN ECE Economic Studies*. New York and Geneva. Forthcoming.
- Viikberg, Jüri (1999) (Ed). *Eesti rahvuste raamat*. Tallinn, Eesti Entsüklopeedia Kirjastus.
- Volkov, Andrei (1997). Fertility Surveys in the Baltic countries in the second half of the 20th century. K.Katus, V.Stankuniene and A.Vikat (Eds). *Demographic Development in Baltic Countries*. Special issue of Revue Baltique. Vilnius, pp.49-63.