Fertility and partnership change : FFS contributions and requirements for the future.

Ron Lesthaeghe
Interface Demography
Vrije Universiteit Brussel
Pleinlaan 2
B-1050 Brussels.

(Summary prepared for the FFS-Flagship Conference, Brussels, May 29-31 2000 : session 5, 2.00-2.30 pm.)

1. Major contributions.

The FFS round of surveys is beyond any doubt one of the main sources of documentation of recent changes in fertility, household formation and dissolution, and residential patterns of Europeans.
The dominant feature of the survey was the retrospective data collection of major life cycle events pertaining to these topics. As a consequence, a set of standard country reports could be produced, with a common lay out for most participating countries. These country reports permit a cross-national comparison of demographic data that are otherwise not available. An example of this is the common typology of individual household positions by sex and age group. Another example is the reconstruction of cohort patterns of household formation and the stages that were followed during the last quarter of a century by each successive generation.

It is, however, quite clear that the standard country reports were produced for a larger public and did not proceed to the stages of more detailed multivariate analyses. The papers currently presented at this conference partially fill this gap, together with earlier research contributions made by individual authors. This was facilitated by the existence of a standard recode data tape and a programme for international comparison. I understand that the ECE will distribute a complete bibliography of research papers that have used the FFS-data so far, and a more detailed meta-analysis of this output is required before arriving at a final assessment. Moreover, these FFS-data will remain a major source for future comparative work, e.g. when new data need to be compared with those collected for the earlier 1990s.

The present paper was not commissioned to give an overview of the FFS archievements and possibilities (for that see the contributions by Macura and Cliquet). Rather, it is supposed to provide a "constructive" critique and an exploration of what can be done next. It is also written before I had seen the research contributions to this conference, and hence some of my points may have become obsolete.

The following issues will be addressed :

- The need for greater documentation of sampling procedures and for the validation of results against other sources.
- The classic issue of the weakness or unavailability of independent variables needed as covariates in the analysis of theoretically predicted differentials.
- The possibility of using the present round as a first wave of a panel study.
- Some additional research topics of increased policy relevance.

2. Sampling and validation.

One of the topics not addressed in the country reports pertains to the sampling procedures used in the individual countries. This is not a problem that particularly affects the FFS but virtually all international survey work. In my experience, heterogeneity in sampling procedures and even more so in the execution of the field work is one of the main sources of incomparability. It is therefore regretable that so little attention has gone to the publishing of more detailed information concerning such crucial items as sampling procedure, drawing of repondents (lists, replacement of persons not found), non-cooperating respondents and their possible characteristics etc. Moreover, if the actual field work has been executed by commercial agencies, one often remains in the dark with respect to the nature of biases induced by random walk or non-probability sampling designs. In many countries today it is considered a success if a face-to-face interviewing procedure in a representative probability sample yields more that 60 percent response. Who, then, are the 30 to 40 percent missing respondents ? We all know that the non-response is highly selective and leads commonly to the underrepresentation of singles, younger persons, persons living in urban areas and particularly in flats. What corrective procedures have then been adopted ? Was there a replacement list, or was there some correction via differential weights, which themselves stem from other sources ? Or was there simply an extra sample for the hard-to-get categories which was left to the interviewers themselves (partial quota sampling) ? Or just none of the above ? So far, I have not come across any document that addresses these issues in detail for either a national FFS-survey or in a comparative way for a set of them.

The matter is further complicated or aggravated by a lack of published validation checks of FFS-results against results from other major sources such as censuses, micro-censuses or other large scale surveys. Validations by age and marital status, and very often also according to education categories, are commonly possible in many countries. The FFS-surveys typically have about 200 to 900 respondents in each five year age category (female sample) with a median size of about 450. Any further breakdown by 3 or 4 education classes quickly leads to a sample fragmentation leaving us with 50-150 respondents for each of these schooling levels. Fertility rates are furthermore often calculated for each of these cells for successive 5-year age intervals up to the present age. Hence, it is quite hazardous to trust these results at face value. Yet, many countries have had a 1990-round of (micro-) censuses with questions on parity and level of education against which the FFS data could be validated. But even if censuses do not contain information on parity, at least two or three-way distributions according to age by education, employment, type of housing or residence could have been performed.

Here I am not stating that these validations would have been absent across the board. It could well be that these were done in the national institutions responsible for the field work. But, the results of these validations are not readily accessible, at least not to the common user like the present author. The outcome of it all is that we may recur to significance testing to bring out differences between categories, but that we are often in the dark about the *exact orders of magnitudes* of such significant differences.

My suggestion for the future is that not only the questionaire design, but also the sampling design and protocols, together with a battery of possible validation tests should be given attention during the preparatory meetings of participating countries. This would considerably enhance the trustworthiness of the raw materials on which large statistical edifices are to be built subsequently. Furthermore, it may be a good idea to insert two or three questions (e.g. on education, type of housing) in the next FFS questionaire that are direct copies of questions

used in the (micro-) censuses. These questions would of course be country specific, but they would be of great help for the subsequent and much needed validation.

3. Covariates, covariates ….

Another recurrent problem in much international research is that it takes a long time to produce a commonly acceptable core questionaire. The FFS is no exception : typically the experts are chosen on the basis of their knowledge of the dependent variables, and rarily because of their expertise of the *independent* variables. This leads to lopsided instruments, with large portions being dedicated to detailed recording of events that will need to be explained, but also with small sections pertaining to the variables that will serve later on as covariates. For instance, my co-authors and I, using the standard recode tape for internationally comparative purposes, found no variable pertaining to the partners' education and masses of missing values for the partner's occupation. Yet, one of the best known theses explaining low fertility is Easterlin's theory of relative deprivation : in the lower income categories, women work to supplement the income of partners in order to meet the consumption standards that they all developed during the socialization phase, and therefore postpone childbearing or stop at parity 1 or 2, thereby contributing to sub-replacement fertility. Also, an adapted version of such a deprivation thesis seems remarkably suitable to explain the rapid drop in Eastern European fertility. We found that there was no way, at least not with the FFS recode tape then available, to check this possibility, simply because we could not set up covariate categories *combining* female and male education or female and male employment status as proxies for the gender specific contributions to overall household income.

Similarly, every country seemed to have picked different attitude batteries in the set that was made available at the onset, and hence cross-national comparisons were seriously hampered by the lack of comparable questions on these issues. Furthermore, attitude batteries are not little instruments that stand on their own and can be used in a "pick and run" fashion. Instead, there are always used as *sets*, and from *all* of them dimensions are then being produced (often through factor analysis techniques) that *only subsequently* serve as covariates. This classic avenue for the analysis of ideational dimensions was obviously not in the mind of the questionaire designers nor in that of the national users. The result is the not very meaningful use that was made of these chopped up batteries in the national country reports.

To sum up, the net outcome is that subsequent analyses of the FFS data will be heavily restricted to *describing* the demographics of the life cycle events associated with household formation and to *describing* the national diversity in these variables (which is of course highly valued !). But, the full *testing of theoretical propositions*, i.e the fundamental research goal, will be hampered by this lack of covariates. It is to be feared that researchers will have to resort instead to proxies that are far from being the appropriate explanatory variable needed for the proper operationalisation of the concepts used in theories. Event history analyses and sophisticated path analytic devices will not be a substitute for the information on relevant covariates that was missed in the first place.

4. Causality and timing : conversion possibilities to a panel survey.

Any cross-sectional survey is only suited for the measurement of associations or correlations. If dual or triple calendars of events pertaining to the timing of one dependent variable and two major time dependent covariates are retrospectively recorded, one may be in a better shape to tease out the two one-way causality effects that produce the overall correlation. It is obvious that this latter design was taken as the ideal for the FFS, with the draw back of course that other objectives had to be dropped (inter alia, the measurement of more covariates). So far, not much research output along the lines of joint histories analysis have been performed on the FFS, but this is likely to be corrected in the future.

For the next round, we suggest that a panel design may be a viable alternative to such detailed retrospective reporting (at the exclusion of much else). More specifically, the persons in the age groups 18-22 and/or 22-26 at the time of the last FFS-round could be considered as members of the first wave that now experience all crucial events. If it is possible to trace these individuals in sufficient numbers (a requirement !) they could be interviewed in the second wave, and the events occurring in the window between these waves can be retrospectively recorded. If X1 and V1 would be the sets of socio-economic positions and values orientations during the first wave, and if X2 and V2 would respectively be the same sets of measurements during the second wave, one could predict the occurrence of event E (or the non-occurrence) in the window on the basis of X1 and V1. This amounts to the measurement of one causal path, i.e. that of the *selection effect.* The causal path in the opposite direction, i.e the *adjustment effect*, can be measured by relating the changes in positions (or the lack thereof) X2-X1 or V2-V1 to the occurrence (or non-occurrence) of specific events E taking place in the window. In this fashion a dynamic rather than descriptive picture would emerge, and, furthermore, changes in positions and/or values could also be used as *outcomes* (dependent variables) of demographic events. *At present, such adjustment effects showing the consequences of demographic choices are hardly addressed, and this gap will continue to prevail in the absence of a panel design.*

If such a panel design would be feasible for at least the younger cohorts at the time of the latest FFS round, there remains the issue that the window would be of the order of 10 years. Many events and changes in position can occur in such a large window. This would still call for the dating of events and position changes of relevant covariates within this slice of 10 years or so. Hence, the panel design, as it can be applied to the case of the FFS at least, does not solve all problems. But the countries which collected more covariates at the onset in wave 1 would be rewarded for their effort, because they would be in a better position to document the selection and the adjustment effects with respect to a larger set of X and V variables. Particularly countries which did better on the V-set would be rewarded by now having a rare panel study in which values orientations would be measured at two moments in time. They would not only be capable of teasing out the selection effect of values orientations (controlling for the X-set) *prospectively* on both the *occurrence and timing* of event E, but they can also measure the adjustment effect (V2-V1) as a consequence of the occurrence and the timing (remoteness) of the earlier event E.

Ideally, a panel design works best if three conditions are simultaneously met : (a) the windows between waves are relatively short, (b) the drop out rate remains small, and (c) all variables concerning relevant positions are measured during each wave. The FFS scores poorly on condition (a), but the retrospective dating of events in the long window would be of major help. For countries with only restricted measurement of covariates to start with, the idea

of adding a panel component can be abandonned, and next time they could concentrate on an enlargement of measurements of explanatory variables. Also, the construction of a panel may run counter to the "protection of privacy" laws in a number of countries. Those who had the foresight of asking respondents for their permission to be contacted again will also be in a better position to add the panel component since this is the major requirement specified by this legislation. In the next FFS round, such a permission may be sought from all respondents as *a matter of routine*, and one would then be in the panel business again with the possibility of using a shorter window as well. In other words, even if the partial panel extension is not feasible at present, one may contemplate such a possibility for the future. Life does not end with the Brussels Flagship conference.

5. Some research topics of increased policy relevance.

One of the salient features of demographic changes in the 1990s in both Western and Eastern European countries is the further postponement of marriage and childbearing. This has now led to universal sub-replacement fertility, with some countries having TFRs below 1.5 children for more than a decade. The issues of differential postponement and recuperation by country, and within countries by socio-economic position or values orientations is bound to stay on the agenda. So far, there seems to be a substantial body of research that addresses differential postponement, but much less attention has been given to *fertility recuperation* at older ages (say, after age 30). Yet, from cohort fertility patterns it is clear that especially Western European countries exhibit very different patterns of recuperation and that later age fertility is a major component of the evolution of TFRs. Countries that have maintained TFRs above 1.70 have done fairly well in recuperating for the fertility loss prior to age 30, whereas those with TFRs below 1.40 are typically exhibiting weak or hardly any recuperation effects. Such differences observed at the macro-level have so far rarely led to a further scrutiny using micro data such as those provided by the FFS. Hence, questions arise as to which social/cultural groups in each country are more likely to have higher versus lower fertility recuperation given their earlier differentiated postponement behaviour. The coupling of differential recuperation to differential postponement according to sets of covariates in a single analysis using cohorts older than 35 in Western European countries therefore ranks high on the agenda. This also leads to issues pertaining to social diffusion mechanisms. For instance, to what extent have lower educated strata followed the postponement behaviour of higher educated ones, and to what degree are such education-related (or other) tempo differentials responsible for the smaller or larger drop in the TFR by country ?

The research agenda for Eastern Europe is likely to differ substantially. Many of these countries now have tempo shifts in union formation and fertility that are linked to the events of 1989. For instance, all cohorts irrespective of age seem to react by lowering fertility during the entire 1990s, leading to period fertility levels below 1.5 children. But in tandem with this, several Eastern European populations have rising teenage fertility, rising proportions of single parent households and still very high induced abortion rates. All of these issues require more in depth scrutiny (where the FFS data could be of assistance), but even more than that, also more focussed policy research concerning ways of improving the contraceptive profile and several other aspects of reproductive health (including STD prevention !). Many of these issues go far beyond the scope of the classic academic analyses, but the FFS-type data may be of help in identifying particular groups at risk. We therefore suggest that the use of the FFS data would be more focussed on such more urgent policy issues.

6. Conclusions

At the technical level a future round of FFS-surveys definitely has to pay greater attention to (a) sampling and validation procedures, and (b) restore the balance in the questionaire between the time used for the measurement of dependent and independent variables respectively. So far, too much attention has been given to the former and too little to the latter.

In terms of study design, some countries may consider using the latest FFS round, at least partially, as a first wave of a panel study. Others may consider this option for the future, but would have to make preparations now to this end. The major bonus of a panel design is that demographic events are not solely given the status of an outcome variable, but can also be considered as a cause of shifts in socio-economic positions and of values orientations. In this way the true dynamics of two-way causation is better adhered to.

Finally, a greater use of the FFS could be made in more policy oriented research. We recognize that academic journals need to be filled, but this cannot be the only or the primary goal. To give the FFS a greater utility in policy research, more and better socio-economic status variables need to be incorporated (indicators of income with levels and sources, housing conditions, job security etc). Also the panel design could be helpful in this respect given the increased opportunity of adding ad hoc questions during specific waves.