

AFRICAN HOUSEHOLDS: AN EXPLORATION OF CENSUS DATA

Discussion (Véronique Hertrich)

The subject of the conference, the range of the papers

When this conference was announced, according its title, I supposed that it will be focused on the use of classical, standardised national census data, with the idea to question their ability to analyse family¹ structures or family-linked topics, but also to explore the potentialities of these available data, generally largely under-analysed (and mainly through rough tabulations), and finally, as a result, to "boost" an enlarged use of census data.

The papers cover a much larger range of topics and even of data. They give rise to very rich and stimulating questions and exchange, which largely exceed the basic subjects I had in mind. Before trying to organise them in some cross-sectional issues, I would like to give some observations about the data in use and some specific features of census which were not exploit.

- *Diversity of the data analysed.* Among the 11 papers, 3 are based on census data collected in the frame of a demographic surveillance system (Hosegood and Timaeus, Townsend et al., Garenne), with original concepts and innovative methodologies. The other papers are based on national census data but one is restricted to one sub-population (Luke) and another (van de Walle and Gaye) questioned an unusual methodology adopted by the census to register household & relationships. To sum up, there is hardly more than one half of the papers (6/11) which analyse, at the national level, usual standardised national census data.
- *Trends.* In theory, censuses are periodically done. While it is far from being always the case in Africa, most of the countries have now realised more than one census. However, if I am right, there is only one paper (Vimard and Fassasi, the paper of Townsend et al. could be added but it is not based on national census) which uses the potentiality of comparison in time and then the possibility of questioning trends using national censuses. (Well I know that it is difficult to obtain micro-data for older censuses ... but isn't it one of the challenges of the ACAP project?)
- *Exhaustivity.* The main characteristic and surely the main value of national censuses is that they are exhaustive. They allow to produce, and to compare, indicators at very small-geographical levels (village, district in a town). It seems that there are now very powerful softwares to draw maps with such detailed data. I am surprised that it was made no use of this "main feature" of census in any paper.

Following this last remark, from one point of view, the question, which arises, is: what is the additional contribution of national census data, comparing to national survey? What analysis (based on national census) would not have been achievable with survey data, as DHS? (Eric: I check that relationships with the household head are registered, even in the recent DHS)

¹ I use in my comments the word « family » in a wide, generic sense.

From another point of view, the purpose can be reversed to say that the papers proved that it is possible to analyse very specific topics (which usually are in the frame of specific surveys) on the basis of the "rough" census data. So convincing analysis were proposed concerning heterogamy (Bartiaux), children education (Kabore and Pilon, Townsend et al.), socio-economic differentials (Vimard and Fassassi), socio-economic position of widows (Luke), the elderly (Noumbissi and Tukufu), the living arrangements of children (Sibadan).

These papers, which were gathered in the "substance" part of the conference, give lessons on the ways usual national census data can be called up to analyse family-related issue. I will focus my purpose on the combination of different levels of analysis, which seems to me a very promising way to optimise census data on our topic.

Optimisation of national census data analysis concerning family issues: combining different levels of analysis

All the papers are based on original analysis of census micro-data, that means wide and free possibilities of outputs. In particular, the available data make it possible to work at the "family" level (household, compound, nucleus) and at the individual level and *to merge* these data. 5 papers show that there are a lot of potentialities in these matching-processes. Townsend and Kabore and Pilon follow such a reasoning in order to introduce the children's household feature among the determinants of schooling. For instance, both papers show that school attainment is highly linked with the sex of the household head (higher schooling level among female-headed children). But other features of the household are taken into account (or are announced to be introduced in analysis) : the language in use in the household, nb of children in the households, relationship of the child with the head of household, having a father in migration ... Sibanda also matches the children individual record with different features of their households (type of household -single/two parents/extended -, household size, head's employment status, sex and education) in order to test the effect of different individual and household characteristics on the living arrangements of children, with a special focus on ethnic variations. Bartiaux and Luke also use merging-processes, to optimise census-data, in order to explore specific subjects. Françoise Bartiaux creates a new level of analysis, the couple, by merging the record of a wife with the record of her husband, in order to analyse homogamy (gap in education level, age gap). Though analysis is restricted, due to the limitations of the census data, to the head-couples and to coresident couples, and though selection effects surely exist, her analysis is really interesting and promising. The question of heterogamy, under study for a long time in industrialised countries, is still now rarely analysed in Africa and she really found stimulating results, for instance that the large age gap pattern seems much more sound than the difference in education level in the reproduction of, said roughly, the male-to-female domination system. Hopefully her work will encourage other studies on this topic in other countries and different censuses in a same country (it would be nice to have results on trends) and also to take into account other characteristics of heterogamy (as ethnicity or place of birth, occupation...). Concerning Nancy Luke, as others, I am really fascinated by the algorithms she developed to try to capture the ever-widow from the census data, using the questions on the marital status of the women, on the marital status of the head, on the orphanhood status of resident children and on the relationship in the household. Though she concludes rather negatively to the possibility of catching the ever-widow with the census data, this work seems to me really exemplary on the imaginative capacities we have to develop to get the best of the censuses.

The possibility of merging different levels of analysis was more rarely used in the other papers. In general, each level (individual, nucleus, household, compound) was analysed separately.

More precisely, when 2 levels were taken into account, it was generally for a descriptive purpose (ex: distribution of population by the type of household) or with attention paid only to the relationship with the head of the higher unit (head of nucleus or household for individual, head of compound for households...), which is in fact registered as an "individual" variable. Given the fact that the data used in several papers recorded different level of "family unit", it would probably be fruitful to extend the "merging" process between these units. For instance: is the household size correlated to the compound size? Are the nucleus features (monogamic/polygamic couples, nb of children living with their own parents) different according the nb of households in the compound, according the presence of old people (old father, mother in law) or, in other words, does the old generations "control" the structuration of young nucleus? Is the survival of one's father (individual question on the survival of parents) linked with the structuration of his household, with the fact that he is (or not) head ? ...

Data registration on family issues. Dilemmas.

It seems to me very difficult – in fact impossible - to summarise or organise the methodological points which were treated in the papers and largely debated during the conference. This is partly due to the fact that the papers used various types of data (using different definitions of family units, different ways of collecting and coding family links, local or national census) and that discussion focus more (from my point of view) on the scientific relevance of family related concepts than on the practical way to improve data collection in national censuses. (*This is also largely due to the fact that I need more time to re-read, understand and “digest” all the comments!*)

In these debates, it seems to me that there are several underlying dilemmas.

- *Exogenous/standardised items versus local/culturally accurate items.*

It's an old debate and, I'm afraid, an insoluble cause of research frustration. We would like to catch the more and the more accurately the features of the populations we study. And for this purpose, exogenous/standardised categories appear usually in the best case very poor, and in the worse, completely inadequate. The 3 papers based on DSS all shows strong efforts to adapt the field tools to the characteristics of the population: registration of oblique relationships in the bilinear society of Sereer (Garenne), very detailed and flexible registration of relationships in Agincourt (Townsend et al.), inclusion of non-residents among the household members of KwaZulu where migrations are very important (Hosegood and Timaeus). In addition, all these studies take into account different level of family units. The approach of the Senegalese census analysed by van de Walle and Gaye has some common points with the Niakhar census. Using categories given by the population, identified by vernacular terms, can be very useful but has also limits, even in a small-scale study. Indeed if, as often concerning family items, they are associated with social and moral values (and thus subject to stake), they are all the more subject to “reshaping”. This “fluidity” is part of the efficiency of the social system. For instance, a distant relative can be reclassified as a cross-cousin to conform to the model of the preferential marriage, or considered as ‘a husband’s brother’ to valid a levirate. If such fluidity wouldn't exist, according demographic constraints (size of the population, aleas in the numbers of males and females), it would be impossible for systems of preferential marriages (as other features of the social system) to function in many populations. Family relationships are specially submit to such reinterpretation, for social purpose or individual interest (Caroline Bledsoe wrote interesting things on how people use this flexibility to appropriate, according individual strategy, as “close relative” a distant one or to reject as remote parent a close one). We have not to underestimate people's ability to conform reality to the pattern they want to

follow (neither to the surveyor's expectations). And finally we have to question if we want to register reality or the picture people want to give of this reality. By the way, such speculations have little to do with national census. I completely agree with Eric Udjo: registration of detailed relationships or different levels of family units is unrealistic in a national census, which has first of all to be operational.

Factual/objectives definitions versus subjective appreciation

In the same vein, the necessity to give precise definitions on the notions in use or to let the people decide by themselves was a point of discussion. Among the papers, Hosegood and Timaeus and Townsend et al., make room to subjective appreciation on some subjects: identification of the members of the household (whatever they are resident or not) and of the head in the first case, identification of the household head (declared by the older women in the household) in the second case. While, as it was argued, some situations may be obvious for the people but difficult to define for the researcher, the principle of "priority to the perceptions" is not without risk. It is probably a valuable approach to catch situations on which we nearly know nothing (with a restricted definitions, a large part of the phenomena might escape). On the other side, the risk is that the «non-definition» results in a wide heterogeneity and that we do not know about what we are speaking. Maybe there is a possible mid-place between the "researcher-imposed definition" and the "empty definition of the researcher"? For instance, an open-questioning in a pre-survey and more controlled questionnaire in the census?

Registration of the household structure around fixed pivotal point (head), variable pivotal points (nucleus heads) or no pivotal point

It's another point, which deals with the precedent dilemmas and the difficulty to resign ourselves to describe family unit around one pivotal point while we would like to catch its complexity. The definition of this pivotal point is in itself very problematic, as it appears in several papers. From an ideal point of you, yes, it would be nice to be able to register only the structure without pre-defined fixed points (that's genealogical description) so that the relationship between each person to the other is caught and that, according the question, we can look at the family structure through one prism or another. But how should we code (each relationship between 2 household members?) and analyse such data? Even if postponed to the analytical step, the choice of the way to deliver a lecture of family structures has to be assumed one time...

Family units versus family networks

While "families" are registered as bounded, exclusive groups, the DSS papers propose to rethink the questioning on more open, extensive way. Indeed, people are not fixed points, they are involved in different relationships, which are as many networks. Family organisation, welfare etc. in one place is not independent of the links with people who are outside. Though it seems obvious everywhere and in every time, the overwhelming development of migrations (and more widely globalisation), challenges increasingly the ability of the "family unit" approach to catch family structuration and organisation. The network paradigm is very exciting and surely very promising. But, once again, how to take into account out of DSS or specific survey?

In addition to these dilemmas, it seems to me that there are 2 topics, explicitly or implicitly evocated, which are recurrent in the conference and need to be examined more deeply. These are *migration and gender* issues. They are problematic in the definitions of census concepts, in the field management of the concepts, in the analysis, in the interpretation, in comparative work... They need to come explicitly examined.

Improving and questioning the quality of national census data

As already said, it seems to me difficult to infer constructive lessons from these dilemmas, concerning national census data. Eric Udjo clearly outlined the constraints and the aims of censuses. I completely agree with him and I have nothing to add. It seems unrealistic to think about a complete and developed reframe of family-related questions in censuses. However I think that we still need a questioning on the quality of census data and that realistic proposals to improve census questions are not out of purpose. DSS, local surveys, and the use of other existing data have surely a lot to bring in this perspective.

The questioning on the quality of census data for themselves (i.e. did they captured what they are supposed to record?) was finally rather absent from the conference (except in Luke's paper). Here, the question is not that of the relevance of the census questions, but on the accuracy of the records according the given question. We have to deal with this question of quality if we want to use national census data.

Some proposals:

- testing the internal consistency of census records gives easily first insights on the quality of data. For instance: proportion of inconsistent answers concerning relationships and other variables as marital status (mother/father/grand-parents who are single), age (children older than their parents), orphanhood questions (different records for brother/sisters of the household head)...
- comparison between local and national census data at the aggregate level. While they are much richer than the national censuses data, local in-depth data are generally also able to produce similar indicators to the national census. Such recalculations can help to discuss, by comparison, the quality of census and to see, for further interpretation of national data, where the biases are.
- matching different sources at the individual level. Consistency in the data recorded for the same person by different operations is probably the most fruitful way to test the quality of data. Post-census surveys probably exist in some countries and should be used. Such job can also be developed in the frame of DSS or local survey if it is possible to access to the nominative census data for the area of the survey. It was possible in the frame of different surveys in Western Africa (Burkina, Togo, Mali), it might be possible elsewhere?

In-depth projects might also give the opportunity of a better understanding of the features of family structures the census is able to capture and on the features that are out of its scope. For this purpose, once again, it would be nice to recalculate usual indicators available with census data, and to compare the picture of family life they give to the most in depth results of the small-scale study.

Finally, in-depth projects have surely a role to play to improve census data questions, but in a realistic perspective: identify simple but powerful questions, identify useless questions or items that could be thrown away...