Household Statistics from German Population Registers

The Need for Household Statistics

More than fifteen years ago the constant need for better information on the housing market caused researchers both on the national and on the regional level to look for methods that could produce this information without costly special surveys from existing sources. This gave rise to a series of research projects in which the Ministry of Science and the Ministry of Housing supported a small group of municipal statistical bureaus in their attempt to derive household statistics from the municipal population registers. These administrative registers are kept up-to-date by legally required notifications of births and deaths and of all changes of residence. In Germany, nobody can visit a public school, take part in an election, start a job or rent a flat if he or she is not registered. This would seem to be an ideal data base for all kinds of population statistics. However, these registers store the data of individuals necessary for their identification for all kinds of public purposes, but, apart from the address, they lack any information on people's accommodation and on the household they belong to. From the statisticians' point of view, this has always been a serious handicap because housing needs arise from households rather than from individuals.

As the population registers provide numerical links only between married couples and for the basic family of husband and wife and their children under 18, one had to look for additional clues to reveal the real size and structure of households. Some municipalities had built up statistical registers of houses and the number of dwellings in them. Combining these data with the population registered under a certain address it was possible to find out the average size and the average composition of the households per address. Selecting addresses of newly built houses, e.g., this method produced data that could be used for projecting the population size and structure of new settlements. But it could not produce any information on the distribution of households by their size and structure, like single parent households, households of senior citizens, households with adult children, or households of young couples where babies could be expected to create additional housing needs. And there was only a small number of statistical bureaus who, at that time, maintained a statistical housing register as described before.

History of the Project and Institutional Background

Referring to experiments with household simulation models carried out by statisticians in Munich and Hamburg a small group of municipal statisticians started attempts to „create“ or generate households looking for common household characteristics of individual persons living at the same.

Most of the theoretical work was done by Friedrich von Klitzing, at first in his function as a member of a research institute that specialized in planning support for municipalities and now as a private consultant. The first computer programme for „generating“ households developed by that institute was - together with some others - taken over by the KOSIS-Verbund, a confederation of municipalities established by the Union of Municipal Statisticians to provide, on a cooperative basis, computer support for statisticians and planners.

Beginning in the early 1980s the household generating project went through several stages of improvement both of its methods and of the underlying computer programme. It is now available for application on hosts and on Pcs. At the moment more than 30 municipalities are trying to apply it to their own data.
The greatest problem was and still is to transform the data from the registers into the standard form required for generating households. Although the set of data to be kept in population registers is fixed by law, there are at least 30 different computer programmes installed in German municipalities to operate the register, each in its own slightly different way. In order to apply standardised programmes for any kind of population statistics, like the KOSIS programme SIKURS for population projections, the data from the registers have to be transformed into that input format which the specific statistics programme requires. At last, in 1995, municipal statisticians agreed on setting up a new KOSIS-project the object of which is to define and produce one standard base record that can be used as input format for all population statistics. A group of 7 municipal and 2 state statistical bureaus are financing this development project; its main application will be the household generating programme. This programme is presently being tested and further improved. It is hoped that more and more municipalities will want to produce standardised well defined and comparable population and household statistics, no matter which register they come from. If they succeed, up-to-date population statistics will be at hand any time and everywhere, not only for the city as a whole but for any small regional unit, even down to the single building block.

It may surprise those not familiar with the organisation of official statistics in Germany that such a basic statistical project was not initiated and supervised by the Federal Statistical Office. The explanation, however, is simple: Official population statistics is based on the decennial national census, on annual micro-censuses and on current statistics on births, deaths and migration between municipalities. This legally regulated and well established system does not depend on population registers and differs more or less from them in its results. The official reserve towards this municipal project may partly be due to the fact that this project may raise the question, why the next census has to be conducted in the traditional way and cannot be replaced, at least in parts, by statistics from the registers. On the other hand, municipal statisticians are not very eager to lose part of their independance to the state in this central field of their activities. At least, the two state statistical bureaus participating in the project make sure that the information on the project is well distributed.

Apart from methodological problems to be discussed later, there are other central problems, which are by no means solved yet. Critics still doubt the quality of the registers, although the general application of computer techniques has greatly improved them over the last decades. Nowadays, you may get higher population figures from official statistics than from the registers in spite of the fact that some years back registers still contained quite a number of people that had moved away without notification. Apparently, the automated exchange of data between population registers has reduced this problem considerably. And the current statistics project with its built-in plausibility checks may help to improve the registers still further.

**The Household Generating Programme**

The main problems to be solved when generating households from the population register lie in the assumptions to be made with respect to common characteristics of persons to be used as indicators for the existence of a common household. The present version of the computer programme treats these problems as follows:

In a preliminary step the standard statistics database is drawn from the register and sorted by the address. Within each address the names of people are replaced by a common number for all those people who have the same family name. Former names and double names are treated the same way. Husband and wife and parents and their minor children who, in the register, are linked by numerical codes receive household numbers, if they live at the same address, and equally each remaining person without such links. Each person also receives a code to mark his or her position in the „basic household“.

But not every person registered lives in a so called private household. Private households, by definition, are formed by people living together and sharing a common income. People living in hostels, in old peoples homes or in comparable establishments, like training centres etc., must
be excluded from the following steps. This can only be done by selecting them by their address. What is left is the standard statistics data base of people presumably living in private households.

The programme, as it is now in use, generates households by linking, step by step, individual persons with the households already generated, thus gradually approaching the real situation. While in a former version the programme evaluated each indicator seperately, the present version tries to combine all available information to discover „real“ household relationships. Starting with the „basic households“ the programme applies the following indicators:

- gender,
- age,
- marital status,
- origin and date of the move to the present residence,
- identical family name.

Now households are generated by the following steps:

At first, the programme tries to identify and link unmarried couples. It puts together individual persons with different names who moved to the present residence on the same day and from the same former address. In order to minimize misinterpretations of this indicator, unmarried couples must differ in sex but must not differ in age by more than a maximum number of years. For lack of common characteristics it is impossible to discover unmarried couples who have come to their present home from different places and at differing dates.

In the second step adult children living with their parents must be found on the basis of identical family names or maiden names. An adult child must be at least 17 years younger than the youngest parent but no more than 40 years younger than the mother and no more than 50 years younger than the father. The child cannot have a partner but may have a child.

Single parents living in the households of their children are attached to these households using the same indicators as in step 2.

Individual persons not yet linked with another household might be grandchildren living with their grandparents. If they have the same name and are at least 34 years younger than another person, they are considered to be a grandchild and are linked with the household of that person. Unfortunately one cannot discover relations with grandparents on the mothers side because the names don’t give the necessary clue.

In a fifth step the programme tries to link brothers and sisters with each other. Individual persons who have the same name as another person who is not more than 10 years older or younger are linked with the household of that person.

As a rule, minor children cannot live in a household of their own. If they have not found their household up to this step, they are linked with the household of another person, if they have come from the same address or moved to their present address on the same day. If this does not attach them to an existing household, they are linked with any household in which one person is at least 17 years older than the child.

Coming from the same former address or moving in on the same day, this is a strong indicator for people belonging to the same household. If all former steps have not led to a link with an existing household, this final step links individual persons to another person or household who has these same characteristics.

Each step is based on the result of all former steps. The sequence chosen has its influence on the result of the generating process. The theory behind the sequence of steps in the present version of the generating programme is as follows:
(1) The generating process must be based on the authentic relation between husband and wife and between parents and their minor children. This must be observed in all further steps.

(2) Recognizing unmarried couples is a great problem and should not be blurred by other steps of the generating process.

(3) In evaluating relationships between ancestors and their descendants the relationship over one generation must have priority. Therefore steps 2 and 3 precede steps 4 and 5.

(4) Any other link that cannot be derived from „natural” relations must follow after the other steps have been completed.

Some of the criteria for generating households in the present edition of the programme can be modified, like the difference in age; some others can be switched on or off. This way the programme can be adjusted to produce the most satisfactory results. Trying out alternatives also helps to reveal the consequences of such adjustments.

Practical experience and further activities

Applications of the programme over the past decade have produced very satisfactory results if compared, e. g., with data of the last (1987) census, as this was done in Nürnberg. Another comparison will be made with the official micro-census in Hamburg, towards the end of 1996. On the other hand there are some natural limits this method cannot overcome, especially with respect to unmarried couples which are systematically underestimated by the programme. The same is true for single parents, who may live in a household with other adults without sharing with them any of the characteristics observed.

These shortcomings are more than balanced by the advantages: The method produces up-to-date and comparable results any time they are needed for any regional unit required. The comparisons between cities or smaller regional units and the comparison over time are a valuable source of information for researchers, planners and politicians; this is why this method will certainly become an indispensable tool, not only in urban statistics and urban research but almost certainly also on the regional and on the national level.

A precondition for that is a widespread application of the programme. This can only be achieved if the time-consuming preparatory work can be reduced to a minimum. One of the most important next steps, therefore, will be to help produce the standard statistical data base for as many operating systems of population registers as possible. All the other problems to be solved lie in the method itself: For smaller municipalities without a statistical bureau estimates based on the basic households must be developed by applying the experience of other municipalities. It will also be helpful to automatically identify addresses where the inhabitants don’t live in private households, without having to provide the list of these addresses externally. Last but not least, the coincidence of names and dates in big blocks of flats may cause the programme to generate bigger households than it would do in smaller houses, this error must be reduced by suitable correcting factors.

To know which persons make up a particular household is only the starting point for all the population and household statistics required. Types of households and their definitions must be agreed upon in order to produce comparable results. „Single parent households”, „multi-generation households”, types of households by their stage in the life cycle, households of foreigners, all these and many other types must be defined and these definitions must be reproduced in the standard programmes evaluating the data.

Most of the work is still waiting to be done. So perhaps there will be another opportunity to report on its progress.