

14 Sweden

The Demographic Data Base at Umeå University: A Resource for Historical Studies

Sören Edvinsson

Introduction

The Demographic Data Base (DDB) is a unit at the University of Umeå, Sweden. Its principal responsibility is to register and prepare social and demographic historical data, making this material available for research, educational and archival use. The DDB operates from three places in northern Sweden. Registration, some system development, and research services are carried out in Haparanda, with additional registration and genealogical services in Jörn. The main aspects of system development are performed in Umeå and the staff there has the main responsibility for providing service to researchers.

In contrast to many other databases, the DDB source material is not collected for specific research projects. The aim is rather to serve as a general resource for different sorts of studies. Historical sources, mainly taken from ministerial records, have been digitalised to the benefit of researchers from different fields, making it easier for researchers to use these sources in their research and carry out more from other Swedish universities as

Dr. Sören Edvinsson is an Associate Professor at Umeå University. His work on the Demographic Data Base is focused on increasing accessibility to digitized sources of historical demographic data. His research interests include 19th century labor history and demographic behavior, and social class and mortality.

For more information on the Demographic Data Base, see the DDB site at:
<http://www.ddb.umu.se/DDB/eng/index.htm>.

well. In fact, any academic researchers, regardless of nationality, can contact DDB to investigate qualified analyses. Even though a unit at Umeå University, the DDB also has a national responsibility and is therefore open to researchers the possibility of using the digitalised material. It is estimated that about a third of the researchers receiving material come from Umeå, another third from other Swedish universities, and a third from other countries.

A Short Historical Background

The Demographic Data Base started in 1973. It originated partly from the pioneering work of the pedagogy professor, Egil Johansson, who studied literacy in Sweden from the 17th century onwards. His results were primarily based on information from the Swedish parish registers. Realizing the potential of the church records, he managed to convince others about the value of digitalising these archival sources. His proposal aroused great interest among historians, social scientists and ethnologists, among others. What made the large enterprise possible, was an increasing focus on quantitative social and economic history research at that time. Originally, the work was organized as part of an employment policy, offering employment in parts of northern Sweden where opportunities for female labor were scarce. After five years, the value of making the project permanent was recognized, with Umeå University assuming the organizational responsibility.

During the 1980s and 1990s, DDB increased the amount of digitalised material while at the same time making it more easily accessible for researchers. The organization has become well established in the research community, and its material has found widespread use in different disciplines and for very different research topics. Our (unfortunately not complete) bibliography lists more than 500 scientific reports in which material from DDB has been used.

The Centre for Population Studies (CPS) was formed on July 1, 1990, as an integral part of the Demographic Data Base. Its main objective is to develop research methodologies for the DDB source material, stimulating national and international research on the DDB material, whenever possible within a multi-disciplinary context. The CPS also promotes the creation of a number of visiting research positions and scholarships in order to encourage wider use of the DDB research material.

The Digitized Source Materials

The materials registered by the DDB originate mainly from eighteenth and nineteenth century church archives. Two main groups of material complement each other. The first is parish registers consisting of microdata on individuals from a selection of Swedish parishes. This material constitutes the population database known as "POPUM." The other large group of material, the so-called "Tabellverket" ("Table Authorities"), covers all Swedish parishes for the period 1749–1859. It contains aggregate parish statistics where central demographic components such as population structure, mortality, fertility, and nuptiality can be studied. Currently, this large quantity of material is under preparation to be made available for research.

Apart from these main materials, other sources have also been digitalised. One example is material that was collected for a large research project in the 1930s, containing parish level statistics for the period 1895–1930, however with less information than in Tabellverket. There is also the poll tax registers for Sundsvall. These, as well as various other minor sources, will not be commented on any further in this presentation.

Some Notes on Swedish Parish Registers

Understanding the content and possibilities of the material from the DDB requires a short presentation of the source materials. The law regulating the keeping of parish registers came into effect in 1686, giving Sweden sources with a wealth of information that serve as a good foundation for studies in historical demography. The background to the keeping of parish registers was not exclusively, or perhaps not even primarily, clerical. One important purpose was, of course, to enable the clergy to control the Christian belief of their parishioners, together with their literacy and moral behaviour. The registers were, however, also used to keep track of the population. The state used them for tax purposes and to recruit soldiers. Thus, the clergy became state employees, administering many duties demanded by the central authorities. And when the government wanted national population statistics from the middle of the eighteenth century, the parish registers were naturally used. The Authorities expected the clergy to calculate these statistics from

the parish registers and then deliver them to “Tabellverket” in Stockholm. In the 17th and 18th centuries, membership in other religious congregations was forbidden, and even when free churches were permitted, the clergy in the Swedish state church still had the responsibility for the population registers. Therefore, the parish registers ideally covered the whole population.

The clergy kept registers on births, deaths, marriages, and migration between parishes, but migration registers were generally not kept before the nineteenth century.

In addition to the usual sources registering the demographic events mentioned above, the clergy also kept records about the complete population in the parish in the so-called “husförhörslängder” (catechetical registers).¹ This is a longitudinal source where individuals and families can be followed as long as they were present in the parish. Each volume covers a period of about 5– to 10 years, where the information is continuously updated, making the catechetical register a dynamic source. If someone died, if a child was born, if a person migrated during the time-period of the volume, the information was entered into the register. Since most central vital events, including dates, were entered into the catechetical register, they partly, but not completely, duplicate the information in the usual demographic sources with births, deaths, marriages, and migrations. After completing a catechetical volume, the clergyman set up a new one into which he transferred current information from the old volume. Like in a census, families were kept together on the same page as long as they lived together, but if a person moved to another place in the parish he/she was transferred to the relevant page for the new residence.

Parish Registers: The POPUM Database

Currently, data from the registers of twenty-five parishes are available in the POPUM database. In the 1970s work started with seven separate parishes, from Nedertorneå in the north to Fleninge in the south. The experiences from these first parishes made DDB realise the problems of working with single parishes. Many of them were quite small, and migration made it difficult to follow people

¹ For more information and a picture of one of the catechetical registers, see: <http://www2.dbb.umu.se/DDB/eng/parishre/keyvar.htm>.

over time, which impaired one of the main advantages of the material. It was, therefore, decided to rather computerize contiguous parishes. The first complete region chosen was Sundsvall, which constituted an interesting area for studying Swedish industrialisation. In this previously agricultural district, the sawmill industry expanded rapidly from the middle of the nineteenth century. The region soon became Europe's leading sawmill industry region, for a time having the most rapidly increasing population in Sweden. The Sundsvall region also became a centre for the early Swedish labour movement, for example leading to the large sawmill workers' strike in 1879.

The Skellefteå region up north was originally registered for a large genetics project. This region covers the most extended time-period: from the early eighteenth century to 1900. Sources from eighteenth-century Skellefteå are, unfortunately, incomplete or are fuzzy with respect to some vital variables, making studies of that period problematic at least for some research topics. Careful controls of the quality of the material are necessary.

The region now being registered is Linköping in the county of Östergötland to the south-west of Stockholm. With 35 parishes and a time-period from approximately 1760 to 1900, this will ultimately become the largest DDB region. About 20 parishes are already digitalised, but record linkage (cf below) has to be done before they will be made available for research.

For the time being, the POPUM database contains about 365,000 individuals, corresponding to a total of 11 million entries. When the Linköping region is included, it is estimated that the sum of individuals will be about 700,000. This roughly corresponds to about 3–4 percent of all nineteenth-century Swedes. However, the population included in the regions cannot be considered as representative of Sweden in a statistical sense, since the selection of parishes did not take such criteria into account. Rather, priority was given to working with complete regions. In spite of this restriction, it can be established that the digitalised regions and parishes represent many different environments. Two larger (by Swedish standards) and two minor towns are included, while some parishes are dominated by agriculture and others by early industrial development. Furthermore, many of the agricultural parishes differed in their characteristics. Large estates dominated

some parishes, while others had more small- or medium-sized farms.

Digitalisation of Population Data

Much preparation is needed to make the population data available for research. It is crucial that the different steps in the process are well defined and the whole process regulated by general principles. The following are some of our basic principles:

1. Data input must be *controlled*. A stepwise control system was therefore developed. A data program checks if invalid information occurs. A certain percentage of the digitalised material is scrutinized and if mistakes are too many, the whole batch has to be redone. In this way, the quality of the data input can be guaranteed. It is however important to emphasize that this is only a control of the DDB input. The clergy could make mistakes when entering the data in the registers and mistakes could also occur at other points in the process.
2. The material and the whole process must be well *documented*. It is for example, very important that the researcher can access documentation of the original sources. What information is found there, what is the quality of the source, etc? Another important aspect of documentation is how the digitalisation has been performed—the decisions taken at each step. There is always information that does not fit in with the usual contents and therefore demands special treatment, and such decisions must be well documented. In that way the risk of misinterpreting the information can be minimized, and it is possible to analyse if the digitalisation procedures caused problems with certain variables.
3. DDB works with *complete* registration. Almost all information found in the relevant parish register is digitalised. This is based on the principle that it should be possible for a researcher to use all the potential information in the parish registers. It should be open for all sorts of questions.

4. The data registration must be *faithful* to the sources. The information should not be modified. Names, causes of death, and places of residence are registered exactly as they are spelled in the source. Birth dates, death dates, and so on, are accepted even if they seem to be inaccurate. Controls of their accuracy must be made at a later step, for example when generating information to construct variables or when a researcher is defining his or her files. Of course, the database cannot be an exact copy of the source, but the content of the information is not modified. The information is accepted with both its richness and its deficiencies. These last two points make the work at the DDB different from digitalisation for specific research projects. Those can be designed to fit their specific research purposes, but in that way the resulting files become difficult to use for other researchers. Information from the DDB needs to be open to an array of different research questions.

Database Structure

In order that data sets from DDB should be accessible to the whole research community, large resources have been devoted to database management and programming. The data registered by the DDB is stored in the relational database POPUM. The information in the different tables can be combined by using the unique identities of individuals as keys.

The following basic demands of the database have been defined:

- It should be possible to identify the pieces of information in the original sources, so that a researcher will be able to reconstruct them.
- It should be possible to find all basic information about a person over time.
- It should be possible to find all information valid for a person at a certain point in time. As a consequence, a census population can be simulated.
- The quality of the information should be limited to the quality of the original sources.

- It should be easy to extend the database structure.

The unique strength of our population database is that the records on individuals have been linked. This has already been done for the Skellefteå region and each parish in the Sundsvall region, while record linkage for Linköping is only starting. The linking process makes it possible to follow individuals through their lives, since all entries concerning the same individual are combined through a unique personal identity. Two types of linking are required, interior to and exterior to each parish. For the time being, the Sundsvall region is only linked within each parishes. Later the links will also be made between the parishes. When this has been done, it will be possible to follow individuals through their lives in the whole region. Previously, linking has been done manually, and the results have proven excellent. Manual linking is, however, very time consuming, so we can no longer afford the large amount of resources required.

In the future, we aim for an automatic linkage program, now being developed to be used, for instance regarding the Linköping region. Records related to the same individual will be brought together and the parents and other related persons identified. This creates specific problems. If too restrictive criteria are used, much of the information will never be linked. If the criteria are too permissive on the other hand, records related to different persons can be linked. Name variation often causes problems: surnames may be replaced; a person may have several names, which may not always be registered accurately; birth dates can vary. In fact every piece of information, perhaps with the exception of gender, might vary over time.

Coding of some of the other variables is needed to make their use in research easier. Occupations give one example, which will be discussed more in detail below. For the time being, classification of causes of death is left to each researcher. In the future, we plan to develop a classification system for this too.

The population database consists of three main parts - Kyrkobok (Church book), Befolkning (Population) and Service. "Kyrkobok" contains information connected to the source. From the tables in "Kyrkobok" it is possible to reconstruct the information from a specific record, since all entries have a unique number as identifier. The "Befolkning" part consists of generated information, that is information not connected to a specific entry

but that can be constructed by combining information from many entries. This could, for example, be time periods of residence in different places, occupations, migrations, related individuals, and marital status. It is usually a complicated task to generate these tables and much effort at the DDB is concentrated on defining rules for this.

POPUM – Some of the Variables

As the digitalisation covers almost all information found in parish registers, the database contains a wealth of information. Individuals can be described according to a vast number of variables, making feasible studies that are seldom realistic in historical research. For example, occupations can be used as an indicator of social status or class in the analysis, a variable that seldom is available in historical studies. Other variables are gender, age, marital status, illegitimacy, place of residence, family size, birth parish, etc. These are only a few examples. Many different types of variables could be constructed according to research interest. Some of the variables are not time-specific, as for instance gender, while others may change over time, typically occupation, residence, marital status. Brief comments on some the variables, especially viewed in relation to the information found in census material from other countries, follow. It should be noted that even if the catechetical registers are not real censuses, to a large extent, corresponding information can be constructed from them. In addition, the catechetical registers contain a great deal of other information, some of which will be referred to in the following part.

Household and Family

The catechetical registers are based on family and household information, which makes them a rich source for studies of connected topics. A page in the ledger was often reserved for a household. The name of the head of household is found on the first line, and names of his family on the following lines. Names of household servants were listed at the bottom of the page. When the population increased and there was much migration, this system became difficult to maintain. Where several families were listed on the same page, it became more difficult to identify who belonged to

each household. This was especially the case in the cities and for the servants.

Geographic Variables

Place of residence is usually stated in the catechetical register headings. In most cases, this is the village but it may be more specific, for example the name of a farm. Sometimes specific groups are listed together without any information on place of residence. This could be the case for the poor, sailors, or soldiers. All villages and places of residence in the registered parishes are coded by the DDB.

Sex, Age, Marital Status, and Relationship to Family Head

This is generally quite clear from the catechetical registers. Age, sex, and marital status are generally given. Explicit information on unmarried persons may be lacking, so to decide the marital status with complete certainty may be difficult. Information about the relationship to the head of the family does not usually create any problems.

Education

Information on formal education is not presented in the parish registers. Marks indicating literacy and comprehension are, however, given in the catechetical registers.

Economic Conditions/Occupations

There is no information on income or private means in the sources. There is good information on occupation but it, too, has limitations. It was not the clergymen's primary concern to provide twenty-first century researchers with information about occupational structure. Instead, they tried to characterise the family socially. As this suggests, married women were seldom given an occupational title, despite the fact that female labour was common. Since the family was characterised by the husband, only his title or occupation was of interest. Furthermore, the information is often of a mixed nature. Both formal titles—such as count or churchwarden, and occupations—such as farmer and smith, are included. For certain purposes these titles are difficult to place in a classification system and can be difficult to incorporate in the analysis. Another problem is that occupational changes are related to the keeping of books and not necessarily to an actual change of workplace.

Updating the source was usually connected with demographic events, so the exact time when changes occurred may not be known.

All occupations in the database have separate six-digit codes, making it possible to extract social status and occupational sector according to a classification scheme constructed by the DDB. Social status classification is to a large extent inspired by earlier research in Swedish social history. The DDB is also participating in an international collaboration, the Historical International Standard Coding of Occupations, called HISCO, which aims to develop an international classification system for historical occupations. This classification is based on ISCO-68 and the results of the project will be presented in 2000. Naturally, researchers can alternatively group occupations according to their own interests

Fertility

Families can be followed through their lives, and the total sum of children can thus be estimated, at least as long as the family stays in the same parish. All births occurring among the parishioners were recorded in the birth and baptism books. It is possible to calculate age-specific fertility rates.

Migration

The migration registers and the catechetical registers represent all the migration between parishes. Migration within parishes is also entered in the latter source. Information on parish of birth is common, at least during some periods. Therefore gross as well as net migration rates can be estimated.

Social (Race, Ethnicity, Nationality, Language, Religion)

It is sometimes possible to identify different religious or ethnic groups in the parish registers, even if that information was not kept consistently. Sometimes, persons belonging to a certain denomination could be listed separately, at other times this information was assigned to the head of the household, or as a remark in the birth and baptism book. The Sami population had registers of their own in some parishes, while in other cases they belonged to parishes with mixed populations. While it is usually possible to identify them in the eighteenth and early nineteenth centuries from the information in the sources, later on this becomes

more difficult. Ethnic information may thus not be entirely straightforward, and can be difficult to use.

POPUM in Research

One of the basic demands of the population database, POPUM listed above, is that it should be possible to find all information concerning a person at a certain point in time. In census material this is easier, since the population was already defined at a certain date. The catechetical records, which hold the parish population register, do not refer to points in time but instead to periods of time. The original source states both the date when the individual was registered on the observed page and why the information about him or her was entered. (This may for example be because of birth, transfer from an earlier volume, or migration.) Corresponding exit information exists for the end of the entry. Since all personal information in these registers is related to time periods, it is possible to construct a picture of the population at a certain time and to generate, for example, occupation or marital status valid for a person at that time. This makes it possible to define a population at risk and to describe the population in terms of the variables chosen.

Another basic demand is the need to follow individuals over time. Since records concerning the same person have been linked, it is possible to reconstruct a life span from the vital information, of course with the restriction that no information is available after a person leaves the registered parishes. Accordingly, the material is well suited for advanced event history analysis. In this way, family and demographic history can be studied with more precise techniques than usual, leading to improved possibilities for performing new types of studies.

The database structure permits us not only to follow individuals over time, but also to connect these individuals to their relatives and thereby reconstruct families and follow these over generations. Families can thus be used as the unit for studies.²

² The family entries for Sigrid Stina Nilsson and Lars Pehrson illustrate how this linking of demographic records can be used to tell a family's history. See <http://www2.ddb.umu.se/DDB/eng/parishre/keyvar.htm>.

The characteristics of the data discussed above make the DDB a valuable resource for many different types of studies. Experiences to date have shown the varied uses of the material. It has been used for studies in many different academic disciplines and there are still many potential research fields where it has been used only sporadically. Even in research fields where many studies have already been published, much remains to be done. Research often leads to new questions demanding new approaches, and these approaches have often been possible to implement with the material from the DDB. Some examples of possible research approaches follow:

Historical Studies on Mortality

A number of studies in this field have already been published. The material permits studies based on individual data for all age groups. Information on cause of death, although it was not always demanded and for a time was only mandatory in cities, is found in the death registers. It is possible to analyse mortality differentials in relation to, for example, social status, gender, and marital status. Mortality can be studied at a low residential level, so that different environments could be compared regarding, for example, access to sanitary facilities. One interesting possibility is to study long-term effects on mortality. Could conditions in childhood, for example, have an impact on survival later in life? Many studies in medicine indicate such a relationship. If similar results can be confirmed in historical studies, our understanding of the mortality decline may need modification.

Historical Studies on Social Mobility and Occupational Career

Information on occupation can be found in all types of parish registers. If a person lived all or most of his or her adult life in any of the parishes included, there are usually many records containing information on occupation, making it possible to follow an entire occupational career. There are good opportunities to use this material in social mobility studies, where different research strategies may be used. The researcher might choose the times in life for which individuals' occupations would be studied. Since it is possible to study several generations from the material, intergenerational mobility could be investigated.

Although there certainly are restrictions of the kind discussed above that are important to be acquainted with, the information on occupations is very rich. Occupations can be chosen in many different ways, and it is important to choose a definition best fit for the specific research purposes. It is for example, possible to select the first own or last occupation, best occupation, occupation at a certain age—for example 30 or 40—or in a time span around these ages, occupation at marriage, most common occupation, and so on.

Family and Household

Even if there are some problems, the material allows many possible research strategies in the field of family and household studies. The survival of children can be analysed from a family perspective. Family (and sometimes household) composition can be studied over time. Other interesting research topics include at what age children left the parental household, how property was transferred over generations, and living conditions among the elderly. Marriage patterns are another interesting research field where material from the DDB can be valuable, for example for studies of age at marriage, the proportion ever married, and choice of marriage partners. Until now, the research field of fertility studies has been largely untouched. The population database can supply researchers with material where fertility patterns can be studied over time, in different geographical areas, between social groups, etc.

Migration

Population movements are often difficult to study in historical populations. Migration registers are missing and continuous information on where people resided is generally lacking. The DDB is well suited to this topic, since migration was registered both in the catechetical registers and migration registers. The catechetical registers not only allow the study of migration between parishes but also of migration on a micro level between villages in the parishes or between households in the villages. We can, for example, analyse the local migrations of household servants. In this way it is possible to analyse the complete extent of migrations. Local movements can be followed by the way information in the catechetical registers was structured. The name

of the village where a family lived is found in the page heading. If a person changed households, his or her name was crossed off the old page and transferred to the page of the new household, where information on the name of the new village also appears.

These are some examples of research fields where the digitalised parish registers from the DDB could be of great value. There are of course many other potential fields where the digitalised parish registers from the DDB could be of great value, many of them less connected to demographic and family history studies, as for example, genetic studies where the tracing of families over generations is of great interest. Another potential for demographic and family history may be to develop new methods from this material.

The Database TABELLVERK

Our second large database, TABELLVERK, provides basic information on the population in each of the 2500 Swedish parishes between 1749 and 1859, totalling data from about 250,000 forms. Unfortunately, not all of the original forms were preserved, but the lacunae on the whole cause minor problems.

In the same way as with the parish registers, our initial ambition was to include all the information from the original forms. However, some economic information and some monthly data on the population have not been registered.

“Tabellverket” consists of two main types of forms: mortality tables and population tables. The mortality tables were to be sent in every year, containing information on births, deaths, and other vital events divided by sex, age groups, and so forth for all parishes. Causes of death were included from 1749–1830. At the beginning of a new year, the clergyman calculated last year’s figures from the parish registers (birth and baptism books, death and burial books, and marriage books). The population structure was reported every fifth year with data taken from the catechetical registers. These tables contain the population separated according to sex, age, marital status, occupations, and some other variables as for example households according to economic standing.

Even if the basic information in “Tabellverket” seems rather straightforward, there are problems even with this material. One is how to handle changes in the layout of the tables. Different forms

were used during the 110 years of "Tabellverket" causing comparability problems over time. Another problem arises from unclear identify of the geographic levels to which the information refers - sometimes a table refers to two rather than a single parish or occasionally a whole deanery.

The digitalisation of "Tabellverket" offers the research community a unique source for studies of many aspects of demographic history from a long-term perspective. The historical demography of Sweden can be analysed from the middle of the eighteenth century until 1859. Earlier, the population statistics was scrutinized at a national and sometimes provincial level. The digitalisation of "Tabellverket" at the parish level makes possible the study of development at more local levels than previously. Differential development of mortality, fertility, and nuptiality can be analysed on a regional and even parish level. Population structure regarding age, gender, marital status, or occupations can be analysed in different environments and comparisons can be made between different parts of Sweden.

Mortality can be more precisely analysed from the information in the forms. For a better understanding of the epidemiological transition, studies need to be made on specific causes of death. When "Tabellverket" was introduced in the middle of the eighteenth century, it became mandatory to report causes of death in the death and burial books. The diagnoses were made by clergymen - making the accuracy open to question. During the eighteenth and early nineteenth centuries, the diagnoses mainly described symptoms and are difficult to translate into modern medical language. They are, however, a valuable and interesting source of information and very useful for some diseases. Accidents, murders, and some epidemic diseases, such as smallpox, were probably well recorded.

From 1830 only a few causes of death were pre-printed in the forms for "Tabellverket," bringing to an end the recording of cause of death in most parishes, while some clergymen continued to register the causes fully. Accidents and some epidemic diseases were still to be recorded. The need to record cause of death was, however, soon recognized, but it was considered necessary to have professionals determine the information; only deaths certified by physicians were of interest. Since most physicians had their practice in towns, only those environments were included. From

1860, all deaths occurring in towns were to be certified by a physician, who was also expected to diagnose the cause of death. The death certificate was then sent to the clergyman who entered the information in the death and burial book. From that time, good information is available about causes of death in towns. The DDB currently works with a project for classifying historical causes of death.

A Short Note on Swedish Population Censuses

Sweden did not have the same sort of population censuses as in many other countries. Until 1855, population statistics were based on the parish statistics submitted to “Tabellverket.” From 1860 this changed. “Tabellverket” was reorganised as the Swedish Central Bureau of Statistics (SCB). Instead of relying on parish statistics, SCB demanded excerpts from the catechetical registers of all Swedish parishes every tenth year. During the late nineteenth century and early twentieth century, the variables included in the “folkräkningar” (population censuses) were name, occupation (generally of the head of household), age, sex, marital status, and place of residence. Some other variables could also occur, for example handicap, religious affiliation, and nationality. Family groups are easily identified. The census of 1890 is currently being digitalised, and for five provinces is available on the Internet at <http://www.foark.umu.se/folk/>. Future access to the census is now being discussed.

Getting Access to Material from DDB

The structure of the databases at DDB is intended to simplify the use of the information. However, the complexity of the databases usually requires expert help when files are prepared for research. This is especially the case for the POPUM database. There is generally no immediate, standard access to the data, since it must be arranged according to the specific research strategy before it is accessible for analysis. The researcher also needs to have good knowledge about the sources, both about their structure in general and about the specific sources. Is there any information missing during the studied period in that parish? Are there any peculiarities about how the books were kept that make it difficult to

extract information in the ordinary way? If there are specific problems that make it impossible to use a variable, can information be attained in any other way? If considerations of this sort are neglected, misinterpretations can easily result. Members of the DDB staff are willing to use their expertise and experience to assist in solving problems.

The usual procedure to gain access to material from the DDB is to contact one of the researchers or the Centre for Population Studies. During the initial discussions with one of them, the suitability of the data sets for the planned research project is assessed. It is important to emphasize that the material is generally not stored in formats that make it easy for researchers themselves to extract and analyse the data. Since special programming is generally needed, in most cases the knowledge and experience of the staff of the DDB proves essential.

After the initial contacts a draft file specification is discussed, defining the research questions, suggesting the population or material for the study, and specifying the variables. From this first specification, a cost estimate for the production of the file is computed. If the estimate is acceptable, a work order will be queued. Thereafter, a thorough file specification is discussed in collaboration between the ordering researcher, the DDB researcher, and a programmer from the system department in Umeå. Decisions are taken about the exact set of variables in the planned data file as well as how to structure the data. Often it is necessary to specify exactly how complicated variables should be defined. It is crucial that the ordering researcher must take the important decisions, of course with support from the DDB staff.

When the file specification is complete, the programmer will create the files. Some general controls are made by the DDB to check for major mistakes in the file. It is, however, important to emphasize that the main responsibility for the controls still rests in the hands of the researcher receiving the files. Since the material is very complicated in many ways, inconsistencies and data that do not meet the original intentions can easily appear. A response is therefore requested as a confirmation that the files have been controlled and accepted.

The database "TABELLVERK" is generally easier to handle, in spite of problems with the identification of geographical units. It is planned that a table with the basic demographic information,

maybe available on the Internet, will be created. To facilitate access, the DDB has also developed a web-based viewing system, allowing the user to study single forms.

The POPUM database can be accessed over the Internet with the search program "INDIKO." More information is available on the DDB's website at: <http://www2.ddb.umu.se/DDB/eng/index.html>. census data comparability in Europe and North America, was adopted as standards for harmonization and recoding.