

11 The Netherlands

Historical Sample of the Netherlands

Kees Mandemakers

Introduction

The Historical Sample of the Netherlands (HSN) strives to compile life history data as completely as possible for a representative portion of the nineteenth and twentieth century population. The sample for this purpose is based on the birth register from 1812-1922 (n=77,000). In this way, the HSN serves as a national database to be used as the core for other databases, assisting researchers who otherwise would have had to start from scratch. It functions as a framework for comparing data from specific research into special communities, social groups, regions, etc. In the process of the construction of the database, cooperation with other institutes and individual researchers was a high priority.

At present, the HSN database contains individual data from birth certificates, death certificates, personal cards, marriage certificates and population registers. By mid-1999, data entry of nearly all 77,000 birth certificates had been completed. Ultimately, the HSN-database will include information on

Kees Mandemakers is Senior Research Fellow at the International Institute of Social History (IISH) in Amsterdam where he heads the Historical Sample of the Netherlands. He also teaches Historical Informatics at the Free University of Amsterdam. His research interests include the social history of education; social mobility; and history and computing.

For more information on the Historical Sample of the Netherlands, see its website at: <http://www.iisg.nl/~hsn/english/welcome.html>.

individual or the research person (RP); the marriage and family of the following subjects: the occupational history of the sampled history of the RP, including age at marriage, religious affiliation and number of children; the parents and parents-in-law of the RP, including occupation, birth place and indication of illiteracy via absence of signatures; social network of the RP through the same information about the marriage witnesses; and the migration history of the RP. With these characteristics, the dataset will form a basic resource for research into historical questions concerning problems in the areas of demography, sociology, epidemiology, socio-economics and social geography.

The importance of the HSN for the researcher is fourfold: first, the HSN provides a representative dataset with which research can be done into social developments in the nineteenth and twentieth centuries. Secondly, the HSN provides a control group or groups which researchers can compare with their own research population. Thirdly, the HSN is developing the expertise which individual researcher can usually not acquire in the short time at his or her disposal. Finally, the HSN offers the possibility for researchers to use the existing HSN dataset as a base for their own research projects.

This last case involves collaborative projects. These are projects wholly or partially financed by third parties, which use HSN expertise and/or the HSN dataset as a starting point. In this way, the current dataset can be enriched with new data or enlarged with new research subjects. The extension of the number of subjects beyond the "HSN basic sample" can be done on the basis of the birth certificates (oversampling) as well as on the basis of other sources (thus far, the population register and marriage certificates).

In the following, the unique character of the HSN-database and the way in which the sample is composed will be described. The next paragraph will deal with the sources that are used, the variables which can be distilled from them, and the availability of datasets. Furthermore, attention is given to the possibilities to extend the database, and a few examples will be given of research projects which are done in cooperation with the HSN. Finally, the way in which the HSN is organized will be described as well as the conditions which are to be fulfilled before scientists can make use of the HSN-database.

Sampling Design and Size

Design

The HSN is based on a research concept in which the Netherlands is treated as one area and the civil registry offices, population registers, land registry archives, church rolls, tax records, etc. kept there are treated as one source, despite the geographical spread and occasional local differences in survival, quality and accessibility of this research material.

The sample of the HSN has been drawn from all persons born in the Netherlands between 1812 and 1922. In 1810, French civil law, which provided for the standardized recording of vital events, was introduced. At varying dates in the course of the year 1811, the Dutch municipalities introduced this form of civil registration. Almost every certificate of birth ever made in the Netherlands is still available. Each certificate was made in duplicate, one remained in the municipality and one was meant for the regional level of the province. Alternative sample designs like sampling from the certificates of marriage or death were considered, but rejected. The marriage certificates had the obvious disadvantage that only married persons would enter the sample and sampling death certificates would introduce very difficult if not impossible procedures for weighting the dataset to an inflow database. Sampling from the population registers could only have started from the period from 1850 onwards and would also create very difficult weighting procedures because of the differences in handling of these registers by the over 1300 municipalities existing in that period. There is also the problem that not all the registers have survived due to floods and the ravages of war and fire.

However, because data concerning date and place of birth are included in the marriage and death registers, sampling in these registers would at first glance have been more practical in the process of constructing life histories. The choice for a sample design from the certificates of birth was made easier by the rather unique situation of the Netherlands concerning demographic sources. There are only a few countries in the world which have a population register like the one in the Netherlands. This was even more true in the past. In principle, the introduction of the population register in 1850, with its

dynamic character makes it possible to follow everyone anywhere in the Netherlands from birth to death. Only in Belgium, small parts of Italy, Norway, Sweden and Japan did comparable systems exist. In other countries the best available data are original census schedules which give no continuous accounting but only a still picture every ten years. The Dutch population registers make it possible to follow persons from the cradle to the grave. This situation differs from other countries where all sources must be entered after which the data of individual persons can be collected by way of record linkage. Besides the Dutch population registers, there are also civil registry offices which record births, marriages and deaths. These are not unique to the Netherlands. What is remarkable is the great accessibility and outstanding possibilities for identification. This rather unique source situation results in countless possibilities for tracing research subjects who--despite the data on migration in the population registers--"get lost," for example, by means of the CBG Personal Card Archive (explained later) or the increasing availability of indexes to the civil registry offices. However, these systems make tracking down persons through the archives of population registers more efficient, since it is possible to track somebody's life history both forward and backward.

A choice for birth certificates excludes one category from the sample. These are immigrants who were not born in the Netherlands. For this group we intend to design an additional sample in the census and population registers. However, a person who came from abroad and stayed only for a relatively short time in the Netherlands would probably always be underrepresented in the survey and there would be only a small chance of entrance in the registers of marriage and death. Besides, a sample based on the certificates of birth has the advantage of including all persons who emigrated.

Stratification and Size

The actual sampling of the HSN is based on tables with the official number of births for each year and each municipality. As a rule the whole period under investigation (1811-1922) is stratified into cohorts of ten years (except the first cohort, which contains the eleven-year period 1811-1822). These periods correspond to the administrative order of the certificates

themselves. Secondly, the sample is stratified according to the regional levels of province and population density (countryside versus cities). In the case of the province of Utrecht, for example this means that a division was made between the two most important cities (Amersfoort and Utrecht) and the rest of the province (countryside). Another example: in the case of the province of South Holland, a fourfold division was made in a) the city of Rotterdam, b) the city of the Hague, c) the smaller cities Leiden, Dordrecht, Gouda, Delft and Schiedam and d) the rest of the province (small cities and countryside). All in all, the sample was stratified in 25 regional parts and eleven ten-yearly cohorts which made 275 strata all together. So, random sampling took place on much lower levels than the national one. This was not only practical—it divides the workflow into more handsome proportions—but is also a way to guarantee the representativeness of the sample on regional and temporal variables.

The tables with the number of births for each municipality were created from two sources. For the period 1811 to 1850, the numbers of births were provided by the so-called Hofstee database, which can be found at the Netherlands Interdisciplinary Demographic Institute (NIDI). For the period 1851 to 1922, data were used from the Historical Ecological Database (HED, Department of Social Geography, University of Amsterdam).

The sample of the HSN originally consisted of a half percent of the birth certificates. A sample of one of every two hundred births would have resulted in a database with a total of 70,000 main entries. Taking into account a loss of data of about 40 percent, mainly because of early deaths (about one-third) and migration, this is statistically sufficient to make significant statements on substrata which cover only 2% of the sample; in absolute numbers this is a substratum of about 800 people. In the pilot project, which was carried out for the province of Utrecht, it was found that this would lead to rather unequal numbers per ten-yearly cohort which survived the age of twenty. Confronted with this result, it was decided to shift to a more sophisticated design and to differentiate the sample ratios according to the period as follows:

1812-1872	0.75 percent
1873-1902	0.50 percent
1903-1922	0.25 percent

Table 11-1 gives the actual numbers of the sample. Instead of the division in 275 strata only the more important regional and temporal levels are shown.

In the context of collaborative projects with other researchers, the HSN database has furthermore been enhanced with about 10,000 certificates. These certificates are drawn from the provinces of Utrecht and Sealand (oversampling the period 1903-1922 to 0.5percent), the region of Twente and the cities of Rotterdam and Zaanstad. Another way of enlarging the number of research units is including parents, children and siblings. This has been done in several projects which have been worked out in collaboration with the HSN. For more information on these projects, see the section "Enlargement of the HSN database (collaborative projects)."

Sources, Data, Variables and Projects

The main sources for the HSN database are the birth certificates, death certificates, marriage certificates and the population registers, which aimed at a continuous registration of the composition of households and the whereabouts of each individual. On the basis of names and addresses it is possible to link other sources to the database, like tax registers, militia rolls, church records, notarial deeds and land registry archives. In the following, the possibilities of these sources will be presented. For a more systematic overview of all variables, see Appendix 11-1.

Almost every certificate of birth, marriage and death ever filed in the Netherlands is still available, as each was made in duplicate. The original remained in the municipality and the copy was sent to the registry of the county court after the closing of the register on December 31 of each year. The duplicates have been assembled in the provincial archives, at least for the nineteenth and early twentieth centuries. Of course, other sources are not as complete for the whole of the Netherlands. The population register of Arnhem, for example, was destroyed during the battle in and around this city in September 1944. Not

Table 11-1. Sample Size HSN-Database, by Region and Period, 1812-1922^a

TOTAL of the Netherlands	76,753
SUBTOTALS per region (provinces and big cities)	
Northern provinces: Groningen, Friesland, Drenthe	12,078
Eastern and Central provinces: Overijssel, Gelderland, Utrecht	17,455
Big cities: Amsterdam, The Hague and Rotterdam	11,443
Western provinces: Noord-Holland and South-Holland (without big cities)	18,278
Southern provinces (Zeeland, Noord-Brabant and Limburg)	17,499
SUBTOTALS per period	
1812-1872 (sampling fraction 0,75%)	45,769
1873-1902 (sampling fraction 0,50%)	22,524
1912-1922 (sampling fraction 0,25%)	8,460

^a As of January 1, 2000; about 200 certificates remain to be entered into the database.

every municipality has sufficient tax records, some people who migrated have disappeared from sight because of inadequate registration in some population registers, and so on. However, the sample is large enough and the research consistent enough to overcome a lot of these problems.

Certificates of Birth, Death and Marriages

As a consequence of the annexation of the Netherlands by the French Empire, French civil law was introduced in 1810. The Code Napoléon provided for the compulsory, standardized recording of vital events in certificates. These certificates served an essential legal function by recording the changes in individual status occurring through birth, marriage, divorce, and death. The certificates had to be drawn up in the municipality where the vital event occurred. The persons named in the certificates fulfilled separate and clearly described tasks. We can distinguish between the registrar, informants and witnesses. The registrar's function was to record in a prescribed way the content of a declaration. He could only refuse a declaration when he was completely sure of its falseness. Examples of informants are a father who notified the birth of his child, a bride and bridegroom who stated their intention to marry, or an undertaker declaring an event of death. Witnesses were chosen by the declarant or by the prospective spouses. The witnesses simply had to be present when the declaration was made in front of the registrar, and afterwards, when the certificate was read aloud. In a strict legal sense the witnesses did not bear testimony to the truth of the declaration itself. Obviously, in the case of a birth, they were not expected to have been present at the delivery. Yet, to a certain extent they were supposed to affirm the contents of a declaration. The registrars were therefore urged to take care that witnesses were friends or relatives of the informants. They could refuse untrustworthy persons. In practice however in the case of birth certificates, the witnesses were frequently either clerks or other municipal officials, or even persons who made a living of signing certificates for a small fee.

The registrars had to note the name, age, occupation and (actual) municipality of residence of the informants and witnesses. This information served the correct identification of these individuals. Certificates had to be signed by informants and witnesses. In case they were not able to sign, the reason was made explicit (usually because of illiteracy). The information concerning the other persons mentioned in the certificates was less detailed. For instance, it was not necessary for the registrar to state the age of the mother on a birth certificate. In Dutch birth certificates, besides data on the baby itself, we find the names,

address, ages and occupations of the parents. In the death certificates are found the last residence, age and final occupation of the deceased and data on the spouse(s). The marriage certificates give information on the occupations, illiteracy (absence of signature) and places of residence of the bride, the groom, their parents and the witnesses (usually four), often family or friends of the marrying couple. By means of these certificates, it is possible to research such topics as social and geographical mobility, marital mobility and illiteracy.

Population Registers

The Netherlands is one of the few countries in the world that has kept a continuous population register from as early as the mid-nineteenth century. Its functions were, among others, to serve as a basis for the franchise and to facilitate the systems of poor relief and conscription. In the early registers, each household was entered on a double page, with the head of the household first; he was followed by his wife where the head was a married male, children, other relatives, and other members of the household. Date and place of birth, relation to the head of the household, sex, marital status, occupation, and religion were recorded for each individual. All changes occurring in the household were recorded in the register. New household members arriving after the registration had started, were added to the list of individuals already recorded, and those moving out by death or migration were deleted with reference to place and date of migration or date of death. In fact, the population register combines census listings with civil registration in an already linked format for the entire population. Families and individuals could, in principle, thus be followed on a day-by-day basis for a long period. In most municipalities, registers cover a time span of ten years between the censuses. Population registers remained in use until 1910 or 1920, after which a new form of continuous registration was introduced, consisting of single sheets, so-called family cards. The registration unit was no longer the household but the family (for an introduction to the old Dutch population registers, see Gordon 1989).

In the late 1930s, the population register was replaced by the *personal card*; from that time on, the *individual person* became the registration unit in all municipalities. Since then the

population register in each municipality consists of a collection of personal cards, containing nearly the same information as the population register. All persons who were alive in 1939 or were born after that year received a personal card. In 1940 the system changed from family cards to individual cards and these cards are sent to the Dutch Bureau of Genealogy after one's death. More than five million cards, sometimes containing information dating from the beginning of the nineteenth century, are available for research at this bureau. Copies of the cards are made available to the public, although without the kind of information that might infringe on privacy (see Van den Brink 1966). In this way, many people who were lost in the system itself before the introduction of the individual record can be tracked down. In addition, many of the research subjects born after 1870 can be found in the personal cards archive of the Central Bureau for Genealogy, so that migration patterns can be followed for these subjects from the opposite direction and the chance of losing the research subject is minimized. Since October 1996, this system has been computerized on a central level. This means that sampled persons who died October 1, 1996 can be found in a computerized way.

Other data on the personal cards include: indication of occupation (from 1940 on), the entire migration path (all addresses), family composition and religious affiliation. Publicly released copies of personal cards have been partly used for the dataset (dates of death and marriage). However, they contain much more information: the name, municipality and date of birth of the person concerned, as well as those of his or her parents, marriage partner(s) and children. The addresses as well as the dates of moves (sometimes going back to the late nineteenth century) are given. Finally, the cards give information on the children of the person concerned, but only when he or she was head of the household. Obviously, the cards mention only those children who were alive in 1939/1940 when the cards were made. They stated whether a child is a son, daughter, stepson or stepdaughter. Place and date of birth are given, date of death, date of marriage and name of the partner, administrative removal by migration and place of destination. The information on the Personal Cards is so sufficient that it is not necessary to collect death certificates dating after 1940.

So, the outstanding feature of the population register is that it presents the RP in constantly changing stages in the life course. The following example gives a typical sequence:

1. as a son or daughter of the head of the household,
2. living independently or living with another household (for instance as a servant),
3. as a head of one's own household (or as a wife), or
4. living as an elderly father or mother with the household of a child.

The population register has a dynamic character. Frequent movers may have up to fifty entries in the registers during their life. The HSN distinguishes between four types of data collection:

- A. First entry in parental household
- B. Itinerary (survey of all entries without actual storage of the information)
- C. Complete storage of all entries in the population registers of own household of RP
- D. Complete storage of all entries of the parental household (from birth of the RP onwards)

In short, the research in the population registers provides us with data on migration, religion, occupation, moves and family structure for the complete life cycle. Other kinds of sources can be linked to the database as well; enumerating them here would lead us too far afield, but it is clear that e.g. tax records are easily accessible because of the availability of names and addresses.

Other Sources

In principle all possible sources in which individuals are named and have an address can be linked to the HSN. In the following, only those sources will be mentioned which have actually been linked to the databases or will be linked in the next two years. This linkage of other sources to the central database usually happens within the context of the so-called collaborative projects with other researchers (see further in the next section).

Census Registers 1811-1840.

In the period up to 1850, there was no population register. The census registers can be considered precursors. They are the result of the censuses that were held and have been preserved in most municipalities. These registers were not supposed to be updated after they were filled in. However, in some regions some dynamic registration was undertaken and noted in these registers.

Militia Rolls.

This source is widely distributed around the country and is rich in data concerning all men just before their twentieth year of life (height, health, occupation, place of residence, father's occupation, etc.).

Tax Records. Many tax records have been preserved, particularly for the period 1880-1920. The taxes were levied on at least half the heads of household. Use is being made of these sources in Hilde Bras's research and the TTA (Twente Textile Workers) project.

Land Registry Archives.

These archives indicate possession of real estate (land and houses) and their taxable appraised value. These data have been used for the TTA research and are also being collected by Hilde Bras.

Church Records.

One can ascertain from church records whether someone was a professing member or not. At the same time, some distinctions can be made within the group of Dutch Reformed, and migration patterns can also be refined through the attestation register. The HSN has not yet worked with church records.

School Records.

Especially for the higher forms of education, school registers with the names and the school results of the pupils have been preserved on a large scale.

Current Contents of the Database

Work on the database began in 1991 with a pilot project in the province of Utrecht, financed by the Ministry of Education and Science. This was completed in mid-1993. With continued

support from the Ministry, further work was done in the provinces of Sealand and South Holland during 1993-1995. Presently the database is being extended to the whole of the Netherlands with support from the Investment Fund of the National Scientific Organisation (NWO, period 1996-2001). At the moment all birth certificates needed for the HSN have been entered (n=77,000), as well as data from death certificates of those dying prematurely (n=20,000) and from personal cards and other death certificates (n=25,000). During the next two years, about half of the marriage certificates and the first entrances in the population registers will have been entered as well. Completion of these sources is scheduled for January 1, 2004, however, at the moment, funding is guaranteed only for the first two years.

As a result of the pilot projects, entry of the certificates and personal cards is almost finished in the provinces of Utrecht and Zeeland. For the province of Utrecht, a public database, *Utrecht release UT.03*, is already available and can be ordered at the HSN (see appendix B for an overview of HSN-releases). For this release the following data structure, which consisted of five normalised Dbase-4 files (see Figure 11-1), was used.

Figure 11-1. Structure of the dataset UT.03

```

6444444444444447
5 GEBOVL.dbf 5<)))))))))))))))))
5 birth/death 5<))))))))))))))))) * n:1 6444444444444447
94L44444444444448 /)Q*S))))))M HSNBERP.dbf 5
*1 * * 5 occupations 5
*: 6444444444444447 * * 94444444444444448
*n 5 HUW.dbf 5<))))))))))1 *
.)>M marriages 5<))))))))))))))1
94L44444444444448 * * 644444444444444447
* * * n:1 5 HSNPLTS.dbf 5
* 6444444444444447 * /))))))M locations 5
1:1 * 5 GTG.dbf 5<))- * 94444444444444448
.)>M witnesses of 5<)))))))-
5 marriage 5
94444444444444448
    
```

GEBOVL.dbf.

This table contains variables from the birth records of the sampled persons. Total number of cases: 3,669. This file also includes a number of variables from the death certificates (n=1,676). Most of the information concerns persons who died young in the community of birth. In addition to dates of death, dates of last observation have been gathered from personal cards, population registers and marriage certificates.

HUW.dbf contains variables on marriage couples from marriage records. Total number: 1,288. This figure includes second and third marriages.

GTG.dbf contains variables on witnesses of marriage records. Total number of couples: 1,288. These include second and third marriages. There are up to four witnesses per marriage record, totaling 4,031.

HSNBERP.dbf specifies all occupational titles and is enriched with classification codes.

HSNPLTS.DBF contains the specific names of municipalities and other types of locations with classification codes and geographical coordinates.¹

*Enlargement of the HSN Database:
Collaborative Projects*

In addition to functioning as an important source for research and as a source for control groups for interpretation of research into specific groups, the HSN database serves as the basis for collecting new data. In practice this is realized through: a) designing and maintaining a data structure for use by individual researchers; and b) taking the database as a starting point for further research, both through increasing the number of individuals included (oversampling) and by deepening by means of recording supplementary variables for a specific group of research subjects. Scholars thus kill two birds with one stone.

¹ For the publications based on this dataset, see Appendix 11–C (esp. Mandemakers and Boonstra 1995, Van Leeuwen and Maas 1997, Kok 1997, Van Poppel and Mandemakers 1997, Delger and Kok 1998).

Not only can they use the data already recorded, but the software and expertise developed by the HSN are also available. This expertise can also be considered an important byproduct of the data entry of the past ten years. For using its software and already recorded data, the HSN sets the precondition that new data must be added to the dataset, so that they will eventually become available to other researchers too.

Essentially, the basic set can be enlarged in two directions: a) enlargement with new research subjects, and b) enlargement with new data. In the foregoing, an overview of possible data has already been given. In what follows an impression is given of five directions in which the HSN basic set is being or may be enlarged with new research persons. Enlarging the database usually happens in cooperation with other research projects. Only a short description can be given of these projects.²

A. New Research Subjects: Enlarging the database with new research subjects can be done by an extension of the HSN sample period itself. The period can be extended backwards, including the period before 1812, by using the baptismal registers which are still kept on quite a large scale, at least for the eighteenth and second half of the seventeenth century. A forward link, including the period 1923-1940, is intended to realize a better link with modern sociological surveys.

B. Oversampling: There are certain research questions for which the HSN database contains too few sampled persons. For this reason the database is extended at several occasions *by oversampling birth certificates for certain periods and regions*. This happened through, among others, the RDN research (Janssens and Engelen) and the OVF control group (Smits and Zielhuis).

The RDN research is a joint project with the department of History of the University of Nijmegen (KUN). Its aim is to build a data file for use in two projects: Women, work and demographic behaviour in the Netherlands, 1900-1960 (Dr.

² For more information on these examples of collaborative projects, please see our website at <http://www.iisg.nl/~hsn/english/a/ac.html> and Appendix 11-B, which lists the most important publications to date regarding the HSN.

A.A.P.O. Janssens) and Changes in nubility and fertility in the Netherlands, 1900-1960 (Dr. Th.L.M. Engelen). The central question of Janssen's is: to what degree does the participation of women in the labour market have effects on marriage and reproduction in the lives of women? The goal is to show the historical development of labour participation by married and unmarried women for the cohorts 1881-1885 and 1911-1915 in four urban areas. The second project will try to understand the causes of demographic modernization of the Netherlands between 1900 and 1960. In comparison to other countries, modernization in the Netherlands started late and was relatively slow. The research focuses on two important discussion points in historiography: the role of cultural-mental variables and the reduction of the aggregation level to the level of individuals and couples. The data file RDN has been composed using a sample of two cohorts (1881-1885 and 1911-1915) from the birth registers in Rotterdam, Enschede, Tilburg and Zaanstad. The sample counts up to 400 persons of both sexes, except in Tilburg and Zaanstad where the research population will consist solely of women.

The project 'Reduced fecundity because of maternal high-risk conceptions' (OVF) is a joint project with the department of Medical Information Science at the University of Nijmegen (KUN). The research was testing indications that persons conceived in situations of hormonal imbalance will run a higher risk of pathology. For women this would result in reduced fecundity. A hormonal imbalance is presumed to occur in the first years after the menarche, in the final years before menopause, in the first period after a pregnancy and after a long period of infertility. It may also occur in certain seasons. The hypothesis that female fecundity decreases when a woman is conceived in a situation of hormonal imbalance can be epidemiologically approached using historical demographic data. This is where the HSN data files play an important role. A sample of the female population in Rotterdam born between 1873 and 1902 has been constructed. This sample consists of four control groups and four risk groups. By means of family reconstruction, it can be checked whether or not risk groups show reduced fecundity. Data used are numbers of still-born children and twins, sex ratio, birth intervals and seasonal

fluctuations in the numbers of births. The largest portion of the risk groups was constructed by means of a direct sample from family cards. The basic HSN datafile and an oversampling thereof for the period 1873 to 1902 (n total = 2100) supplied the control groups. In total this database contained 2591 sample persons with family reconstructions, with 1500 of them coming from Rotterdam and the rest spread all over the country. This research was finished in 1998 (Smits 1998).

C. Sample Based on Other Documents: Enlargement of the dataset with people selected on the basis of other criteria than a sampling of the birth certificates. Examples are: the already mentioned OVF research in which the risk groups were found in the population register, Lucassen's pioneer project (DVI), whereby the research subjects were taken from the population register; and van Poppel and Schellekens's research (RCM), which is being done on the basis of a sample of the marriage certificates.

The pioneer project 'Settlement determinants for immigrants and their descendants in the Netherlands', 1860-2000 (DVI, under the supervision of Dr. L.A.J.C. Lucassen, financed by NWO, the University of Amsterdam and the Centre for Migration History) consists of two partial surveys. The first survey will examine how three groups of immigrants integrate with their new environment. These immigrant groups are Germans, Italians and 'internal' immigrants from the provinces of North Brabant and Saeland. The other survey will take a look at how these immigrants were accepted by their new societies. The HSN plays a role in the first survey. Family reconstructions will be carried out for two cohorts (1870-1880 and 1920-1930) for three generations of immigrants. Also their marriage certificates will be entered into a database. The total number of families investigated will be about 1000.

Another example of 'oversampling' from other sources and integration with the HSN database and software is the research of Dr. F.W.A. van Poppel (NIDI, The Hague) and Dr. J. Schellekens (Hebrew University, Jerusalem), 'Religious differences in infant and childhood mortality, The Hague, 1860-1920". This project focuses primarily on the detection of the causes for the relatively low numbers of deaths for Jewish children in the nineteenth and early twentieth centuries and the

relatively high death numbers for Catholic children in that period. The research will answer the question whether socio-economic and demographic characteristics (age of mother at time of birth, rank number of birth, social class, income, etc.) can explain the differences in the numbers of deaths. A data file containing about 3600 families from The Hague will be created for the period 1860 to 1920. Marriage certificates, population registers and death certificates will be used as sources. A part of the HSN basic sample will be part of this database. All data will enter the database using HSN software.

D. Including Relatives: Enlargement of the basic set with people who are related to the original research subjects, whether horizontally (spouses, siblings) or vertically (parents, children). The current sample period is, in principle, long enough to portray between five and six generations. This type of enlargement occurs in Hilde Bras's research, in which sisters of research subjects are included; and in the already mentioned research of Leo Lucassen, which covers three generations of immigrants.

The MFZ-project, 'Geographic and Social Mobility of Female Domestic Servants in Zeeland, 1860-1920' is a joint project with the University of Utrecht and the Inter-University Center for Social Science Theory and Methodology (ICS). The research by drs. H. Bras is aimed at the reconstruction of the life careers of women born between 1845 and 1915 on some specific regions. The HSN data file contains 600 of them, 450 of whom passed the age of fifteen. To enlarge the sample size, sisters were added to the sample, with a maximum of one. This led to a total sample of approximately 700 women. The researcher will focus on the chances in life of domestic servants compared to other female and male occupations. The research will try to answer the question of how opportunities depended on parental environment, migration and the employer's status. Data are being gathered referring to migration, marriage, family situation at the time of birth and at entry into the labour force, income of the fathers, employers and spouses.

E. Including Non-relatives: Enlargement of the basic set with people who are related in other ways than by family ties, for example, neighbours or employers. In the thesis research of Hilde Bras, data were gathered for the employers of the

domestic servants and in the research of Lucassen data are collected about the keepers of the boarding houses (landlords).

Organizational Structure and Policy

Organization

The HSN started in 1987 with the formation of a work group whose members were from different social science disciplines like sociology, socio-economic history, social geography, economics and demography, all sharing a relatively strong historical interest. The members of this group came from several Dutch universities. In 1989 this group was formally structured as a foundation with independent legal status, the HSN foundation.

The foundation manages the HSN data with the goal of making these available to scientific researchers in the Netherlands and abroad. The HSN Board strives not only to make the data available, but also to distribute them as widely as possible. The only restrictions concern any overlapping of the research inquiries in question and the protection of data confidentiality. In order to guarantee continued existence and accessibility of the HSN database, the HSN Foundation has linked itself by contract to the International Institute of Social History (IISH) in Amsterdam, which forms part of the Royal Netherlands Academy of Sciences (KNAW). The IISH is an internationally renowned archive and research institute in the field of social history. It is devoted to the acquisition, management and accessibility of collections in that area.

The International Institute of Social History (IISH) provides housing for the HSN's activities and assumes the burden of the resulting costs. The IISH has guaranteed a permanent 0.5 FTE position for coordination tasks. The actual data entry is done on the basis of projects which are externally funded. The enlargement of the database to the whole of the Netherlands is granted by the NWO (Dutch Scientific Organization) Fund for Medium-Large Investments. Data entry is organised on the basis of unemployment schemes. This means that only fifteen percent of the wage costs of the data entry personnel has to be provided by the projects themselves. About

half of the personnel works on the spot, that is at the provincial archives, the other half works directly from the IISH.

Policy

The HSN sample is based on the birth certificates from the civil register. The number of variables that are added and can be added for the research subjects on the basis of other sources is quite large. The HSN will also be supplied with new research subjects by means of collaborative projects. This is being done both on the basis of birth certificates (oversampling) and on the basis of other sources: up to the present, population registers and marriage certificates.

The HSN has a three-part strategy. In the first place, it intends to promote the use of the existing dataset as much as possible. It aims to enlarge the existing dataset, both by launching its own projects and by establishing collaborative projects with others. The HSN intends to maintain and where necessary develop its existing expertise. This expertise is twofold: a) historical knowledge of the sources used, and b) knowledge of computerization techniques for scientific information. This expertise can also be described as the systematic and responsible entry, documentation, filing and making available of historical micro-data.

Secondly, the HSN sample can be used as a basis for more focused data collection, through which data from other historical sources, such as church registers, tax records and spatial context, can be added to the HSN data. This does have consequences for linkages with other collections; if these collections want to use the "birth bank" as a starting point for further (partial) research, then the researchers concerned will have to follow the design of the HSN database. This means that researchers must adopt the data structures employed by the database. Their data will also, in principle, have to meet the same quality standards as the data used by the Foundation for its own database. The doctoral thesis research by H. Bras on "Female Domestic Servants in Zeeland" is an example of such research. In this project, the life courses of women and their sisters (if any) are reconstructed. The project investigates the extent to which life chances depend on parental milieu, migration and employer status. The data collection concerns migration paths, marriage certificates (if any), the

parental family situation at the time of birth and at first employment, and the income situation of fathers, employers and spouses.

Thirdly, the HSN not only provides data but also software to enter and process the data. Use of this software not only leads to a standardization of research data entry, it also means enormous savings in expenses and time on the part of potential researchers, who are no longer compelled to design and install the necessary programs themselves.

Public availability of the data is regulated by a set of rules established for this purpose (Dutch Registration authority personal records #0-0030426), see Appendix D. These privacy rules are based on the principle that in deciding which data will be made public for purposes of scientific research, the same rules will be followed as pertain in the archive from which the specific data were collected. This may mean that part of the database can only be made available in anonymous form. For researchers who need access to data arranged by name, a certain number of additional confidentiality guarantees are built in.

Cited References

- Brink, T. van den, 'The Netherlands populations registers', *Sociologia Neerlandica* 3 (1966), 52-53.
- Gordon, C., "The *Bevolkingsregisters* and their use in analyzing co-residential behaviour of the elderly" ('s-Gravenhage: NIDI, 1989).

Appendix 11–1

Overview of HSN Database Variables

The following lists the most important variables used in reconstructing life histories, by the source in which they are found. These variables do not pertain solely to the Research Person (RP), but also to persons associated to the RP. *In this survey, the variables which have already been found in preceding sources and which do not change (such as the name) are not repeated, nor are those which can be predicted (such as age).* Variables which stand for locations, occupational titles, surnames and first names are collected into reference tables which are used for controlling the newly entered data. The table with occupations is in this way linked to diverse coding systems for occupational titles. In the same way the locations are systematized and coded into standard codes for locations and enriched with X- and Y-coordinates for calculating distances etc. This system is under construction, but releases have been realized for the data of the provinces of Utrecht and Zeeland.

RP=Research Person

Birth certificate

Father RP	Name, age, occupation, address, signature
Mother RP	Name, age, occupation, address, civil status
Witnesses (two)	Name, age, occupation, address, signature
RP	Name, date and municipality of birth, sex, address

Death certificate

Informants (two) ³	Name, age, occupation, address or municipality, signature
RP	Occupation, civil status, address of decease, municipality of residence
Spouses of RP	Name, age, occupation, civil status, address or municipality of residence

Marriage certificate

RP	Date and municipality of marriage, municipality of residence and address, occupation at marriage, names of previous spouses, signature
Illegitimate child RP	Names, date and municipality of birth
Spouse RP	Name, address, age, occupation, signature
Parents RP	Occupation, municipality of residence or date and municipality of death, signature

³ Often related to the RP.

Parents in law of RP	Name, age, occupation, municipality of residence or date and municipality of death, signature
Witnesses (four) ⁴	Name, age, occupation, municipality of residence, signature

Personal card (only for RP's still alive in 1940)

RP	Occupations, civil status (with dated changes), religion, addresses and municipalities of residence (dated), date and municipality of death
Spouses RP ⁵	Name, date and municipality of birth and death
Parents RP	Municipality and date of birth and decease
Children RP	Name, date and municipality of birth, date and municipality of administrative removal due to death, marriage or departure

Population register (1850-1940)

The population register always contains the following variables for all persons included in the family: parents, siblings, children, grandchildren and (after 1900 less systematically) other relatives, boarders and servants

- date of entry
- full name
- date and municipality of birth
- sex
- relation to the head of the household
- civil status (with possible changes including date and municipality, on which and where these changes have taken place)
- religion (with possible changes)
- occupation (the same)
- date and municipality of provenance
- date and municipality of destination
- (possible) municipality of legal residence and further remarks

The following sources are not actually recorded yet, except for collaborative projects.

Census registers (period 1810-1840)

The census registers are quite comparable to the population registers except for the lack of dynamics in the data. Generally they exist only for the years 1810, 1825 and 1840.

⁴ Usually related to the RP.

⁵ When RP is still alive.

Registers of conscripts

RP	Occupation at the time of examination, address, description, examination report (e.g. height), date of conscription
Father RP	Occupation, address

Capitations

RP	Assessment and estimated income (also important for evaluating occupational titles)
Father RP	The same

Land register

RP	Possession of real estate, assessment land tax
Father RP	The same

Church records

RP	Migrations (attestations), type of membership (only baptism or also confession)
----	---------------------------------------------------------------------------------

Appendix 11–2. HSN Data-Releases

The HSN distinguishes two forms of data-releases: public and not public. Public releases are available to everyone who fulfills the conditions mentioned in the licence agreement.

Additional data-releases are made available to certain researchers on the basis of a cooperation contract in which terms of use are established for a specific lecture or congress. These contract-specific data releases usually contain rough, minimally checked test versions of future public releases. An embargo period usually precedes the release of public data so that the researchers can finish their investigation.

The following list contains only contains the most recent version of the data-release concerned. For a description of the research, please see the section on “Enlargement of the HSN Database: Collaborative Projects” presented earlier in this article.⁶

Public Use

1. *Dataset Utrecht UT.03*, data basisset HSN province Utrecht (n=3669, period of birth: 1812-1912).

This dataset can be ordered from the HSN after paying handling costs and filling out the licence agreement.

Not- or Limited-Public Use

1. *OVF-release nov. 1996*, dataset OVF-research
2. *RDN-release 97.1*, dataset RDN-research
3. *HSN-release Zeeland 98.2*, data basisset HSN province Zeeland (n=4443, period of birth: 1812-1912).
4. *DUM-release 98.1*, dataset DUM-research
5. *Release MFZ.02*, dataset MFZ-research

⁶ See also: <http://www.iisg.nl/~hsn/english/a/ac.html>.

Appendix 11–3.

Selected List of Publications Based on the HSN

- Bras, H. (1998). 'Domestic service, migration and the social status of woman at marriage. The case of a Dutch sea province, Zeeland 1820-1935,' *Historical Social Research* 23, nr. 85, 3-19.
- Delger, H. and J. Kok. (1998). 'Bridegrooms and biases: a critical look at the study of intergenerational mobility on the basis of marriage certificates,' *Historical Methods*, 31, no. 3, 113-121.
- Janssens, A. (1998). 'Class, work and religion in the female life course - the case of a Dutch textile town: Enschede, 1880-1940', *Historical Social Research* 23, nr. 1/2, 254-274.
- _____. (1997). 'Mujeres y trabajo en la economía doméstica de los Países Bajos, 1880-1960: El caso de Enschede', *Boletín de la Asociación de Demografía Histórica*, XV, no. 2, 55-78.
- Kok, J. (1999). 'Migratie als gezinsstrategie, Midden-Nederland 1850-1940', in: J. Kok, A. Knotter, R. Paping & E. Vanhaute, *Levensloop en levenslot. Arbeidsstrategieën van gezinnen in de negentiende en twintigste eeuw* (Groningen/Wageningen, Nederlands Agronomisch Instituut, 1999), 89-156.
- _____. (1997). 'Youth labour migration and its family setting, the Netherlands 1850-1930,' *The History of the Family. An International Quarterly*, 2:4, 507-526.
- _____. (1997). 'Estrategias familiares en el abandono del hogar y la migración juvenil. Los Países Bajos 1850-1940', *Boletín de la Asociación de Demografía Histórica*, XV, no. 2, 79-111.
- Kok, J. and H. Delger. (1998). 'Success or selection? The effect of migration on occupational mobility in a Dutch province, 1840-1950', *Histoire & Mesure* 13, nr. 3/4, 289-322.
- Maas, I. (1998). 'Coding Dutch occupations into ISCO68. The most frequent male occupations in the HSN-Utrecht dataset', in: M. H. D. van Leeuwen, I. Maas & A. Miles, *Historical international standard coding of occupations: status quo after coding 500 frequent male occupations* (Berlin: Max Planck Institute), 95-117.
- Mandemakers, K. (1993). *Eindverslag HSN-proefproject provincie Utrecht* (Amsterdam: IISG), 154 p.
- _____. (1994). 'Historical sample of the population of the Netherlands (HSN). Backgrounds, objectives and international context', in: H. J. Marker en K. Pagh (ed.), *Yesterday. Proceedings from the 6th international conference Association of History and Computing Odense 1991* (Odense: Odense University Press), 174-181.
- _____. (1995). 'New Approach to the Study of Migration in the Netherlands during the 19th and 20th Century. First Results of

- the Historical Sample of the Netherlands', in: G. Jaritz, I. H. Kropac and P. Teibenbacher (eds), *The Art of Communication. Proceedings of the eighth International Conference of the Association of History and Computing*, Graz, Austria, August 24 - 27. (Graz: Akademische Druck- u. Verlagsanstalt), 64-73.
- Mandemakers, K. & O. Boonstra, ed. (1995). *De levensloop van de Utrechtse bevolking in de 19e eeuw* (Assen: Van Gorcum), 186 p.
- Mandemakers, K., O. Hoogerhuis and A. de Klerk, eds. (1998). Over Zeeuwse mensen. Demografische en sociale ontwikkelingen in Zeeland in de negentiende en begin twintigste eeuw. Themanummer Zeeland 7, nr. 3. Inleidingen HSN-congres Middelburg, 20 maart, 64 p.
- Mandemakers, K. and J. Kok. (1998). *HSN. Historical Sample of the Netherlands. Dataset Release Utrecht UT.03* (Amsterdam IISG Publishers), 54 p.
- Smits, L. J.M. (1998). Preconceptional determinants of the fecundity of female offspring (thesis KU Nijmegen), 205 p.
- van Leeuwen M. H. D. & I. Maas. (1997). 'Social mobility in a Dutch province, Utrecht 1850-1940,' *Journal of Social History* 30, 619-644.
- van Poppel, F., J. Kok & E. Kruse. (1995). 'La cause principale des différences c'est le crime': Essai d'explication de la surmortalité des enfants illégitimes à la Haye au milieu du xix^e siècle', *Annales de Démographie historique*, 241-275.
- _____. (1997). 'Mortality among illegitimate children in Mid-Nineteenth-Century The Hague', in: C.A. Corsini and P.P. Viazzo (eds), *The decline of infant and child mortality. The European Experience: 1750-1990* (The Hague: Martinus Nijhoff Publishers), 193-211.
- van Poppel, F. & K. Mandemakers. (1997). 'Differential infant and child mortality in the Netherlands 1812-1912: First results of the historical sample of the population of the Netherlands', in: A. Bideau, B. Desjardins & H. Pérez-Brignoli, *Infant and child mortality in the past* (Oxford: Clarendon Press), 276-300.

Appendix 11–4. Privacy Regulations HSN

Regulations governing access to data owned by the Historical Sample of the Population of the Netherlands (Stichting Historische Steekproef Nederlandse Bevolking), as adopted at its Board meeting of January 21, 1993, and as deposited with the Personal Data Protection Board (Registratiekamer persoonsregistraties), and registered there with no. 0-0030426.

1. The Historical Sample of the Population of the Netherlands (Stichting Historische Steekproef Nederlandse Bevolking, hereinafter called HSN) has its registered office in Utrecht (Register of Foundations Chamber of Commerce [Stichtingenregister Kamer van Koophandel], file no. S 183871).
2. These regulations are obligatory under article 5 of HSN's Articles and the Personal Data Protection Act (Wet persoonsregistraties); (Act of December 28, 1988, Book of Statutes [Staatsblad] 665, 1988).
3. The term 'data of the HSN foundation' refers to all data collected within the bounds of the foundation's objectives. They are 'data on a sample of the Dutch population, as they may be found in certificates of births, deaths and marriages, records of the population register and other historical source material,' HSN's Articles, article 2, paragraph a.
4. A distinction is made between data available to the public without restriction (hereinafter called 'public data') and data which may be made public subject to restrictions (hereinafter called 'scientific data'). In a register, a record will be kept of a) which data are contained in the archive and b) which status these data enjoy (public or scientific).
5. Public data are all data which may be consulted in public archives (as referred to in the Archiefwet [Archives Act] 1962, Book of Statutes 313, article 1, paragraph c) without restrictions of any kind.

6. Scientific data are all data not covered by article 5 of these regulations, and collected within the bounds of the foundation's objectives (see article 3). These data are managed in accordance with article 17 of the Decree of January 2, 1990, designating those registers of persons to which the articles 19, 24 and 25 of the Personal Data Protection Act do not apply.
7. With the passage of time, the data covered by article 6 may be released in strict accordance with the rules of the archives from which these data have been drawn (see article 5). A record of this will be kept in the register mentioned in article 4.
8. The data referred to in article 6 are accessible only for the purpose of scientific research, as envisaged in article 2, paragraph b of HSN's Articles. Researchers using these data will have to sign a mandatory statement to this effect. Only after this statement is signed will the Board be authorized to release the data concerned.
9. The provisions of article 8 imply, among other things, that use of 'scientific' data is allowed exclusively on condition that the anonymity of data is guaranteed, this within the meaning of article 33 of the Personal Data Protection Act.
10. The provision of article 9 grants an exemption from meeting the requirements of the Personal Data Protection Act as regards the provisions contained in articles 28 up to and including 32, which concern the rights of interested parties to inspect and amend personal data.
11. The Board may authorize in writing another legal entity to manage access to the data, in accordance with the above articles.

