Dutch Lives. The Historical Sample of the Netherlands (1987–): Development and Research

By Kees Mandemakers and Jan Kok

Historical Life Course Studies is the electronic journal of the European Historical Population Samples Network (EHPS-Net). The journal is the primary publishing outlet for research involved in the conversion of existing European and non-European large historical demographic databases into a common format, the Intermediate Data Structure, and for studies based on these databases. The journal publishes both methodological and substantive research articles.

**Methodological Articles**

This section includes methodological articles that describe all forms of data handling involving large historical databases, including extensive descriptions of new or existing databases, syntax, algorithms and extraction programs. Authors are encouraged to share their syntaxes, applications and other forms of software presented in their article, if pertinent, on the EHPS-Net website.

**Research articles**

This section includes substantive articles reporting the results of comparative longitudinal studies that are demographic and historical in nature, and that are based on micro-data from large historical databases.

Historical Life Course Studies is a no-fee double-blind, peer-reviewed open-access journal supported by the European Science Foundation (ESF, [http://www.esf.org](http://www.esf.org)), the Scientific Research Network of Historical Demography (FWO Flanders, [http://www.historicaldemography.be](http://www.historicaldemography.be)) and the International Institute of Social History Amsterdam (IISH, [http://socialhistory.org/](http://socialhistory.org/)). Manuscripts are reviewed by the editors, members of the editorial and scientific boards, and by external reviewers. All journal content is freely available on the internet at [http://www.ehps-net.eu/journal](http://www.ehps-net.eu/journal).

**Co-Editors-In-Chief:**
Paul Puschmann (Radboud University) & Luciana Quaranta (Lund University)
hislives@kuleuven.be

**The European Science Foundation** (ESF) provides a platform for its Member Organisations to advance science and explore new directions for research at the European level. Established in 1974 as an independent non-governmental organisation, the ESF currently serves 78 Member Organisations across 30 countries. EHPS-Net is an ESF Research Networking Programme.

**The European Historical Population Samples Network** (EHPS-net) brings together scholars to create a common format for databases containing non-aggregated information on persons, families and households. The aim is to form an integrated and joint interface between many European and non-European databases to stimulate comparative research on the micro-level. Visit: [http://www.ehps-net.eu](http://www.ehps-net.eu).
HISTORICAL LIFE COURSE STUDIES  
ONLINE FIRST (2020), published 15-06-2020

Dutch Lives  
The Historical Sample of the Netherlands (1987–): Development and Research

Kees Mandemakers  
International Institute of Social History, Amsterdam & Erasmus University Rotterdam

Jan Kok  
Radboud University Nijmegen

ABSTRACT

The HSN was initiated during the period 1987–1989 when an interdisciplinary and interuniversity group of Dutch scholars started discussing the foundation of one large database with data on individuals. Building one general prospective database with multiple research possibilities was considered as the only way to realize a cost-effective and properly documented tool for historical research from economic, social, demographic, epidemiological and geographic perspective. The birth registration was considered the most adequate sample framework. The new database should be ‘open’ in the sense that extension should be possible in all kinds of ways: more sources or variables, more persons and larger time periods. The HSN was deliberately created as a nationwide sample covering the whole 19th and 20th century. Since 1991 about 12 million Euro has been invested in the database and related projects. Besides the basic sample about 25 additional projects have been realized that created all kind of extensions to the database. A special project is LINKS by which the indices of names from the Dutch civil registration are used to reconstruct pedigrees (for the period 1780–1940) and complete families (1811–1900) for the whole of the Netherlands or parts of it. In this article we will present an overview of the research that was done with the original themes and the new fields that were introduced over the years. We will also go into methodological issues that were picked up by the ‘HSN community’ and we will point out the present and future challenges for the HSN.

Keywords: Historical databases, Life courses, Demography, Sociology, Epidemiology, History, Economy

e-ISSN: 2352-6343
PID article: http://hdl.handle.net/10622/23526343-2020-0001?locatt=view:master
The article can be downloaded from here.

© 2020, Kees Mandemakers, Jan Kok
This open-access work is licensed under a Creative Commons Attribution 4.0 International License, which permits use, reproduction & distribution in any medium for non-commercial purposes, provided the original author(s) and source are given credit. See http://creativecommons.org/licenses/.
INTRODUCTION

In their recent overview of the history of historical demography in the Netherlands, Theo Engelen and Ad van der Woude (2016) describe the years around 1990 as a ‘turning point’, a period in which Dutch historical demographers began to move away from aggregated data based on censuses to individual-level data. More than class, group, local community or region, the individual was taken as the center of research. In this development, the Historical Sample of the Netherlands (HSN) is seen as ‘the most prominent example’.

The HSN was initiated during the period 1987–1989 when an interdisciplinary and interuniversity group of Dutch scholars started discussing the foundation of one large database with data on individuals to be used in existing research and for exploring new research themes. Starting a database with data on the appropriate level of research was indeed the main motivation behind the formation of this group of scholars. Out of necessity much research was using data on the aggregative level of provinces or municipalities while decisions regarding life course issues were taken on an individual level. Building one general database with multiple research possibilities was considered as the only way to realize a cost-effective and properly documented tool for historical research from economic, social, demographic, epidemiological and geographic perspectives. During this initial period some important decisions were taken. The sample should be taken from an individual perspective nationwide and be limited to the 19th and the early 20th centuries. This creates standardized biographies for the whole 19th and 20th century, even until the present day. This choice for a prospective database also implied that birth certificates were the best documents to start with and to form the sample framework. The new database should be ‘open’ in the sense that extension should be possible in all kinds of ways: more sources or variables, more persons and larger time periods. Besides the original research persons, it should also be open to include additional research persons (such as siblings or a second and third generation) and to make oversampling from birth certificates for some specific region or time period. Other advantages of the database would be the systematization in data gathering and documentation, the reuse and multiple use of existing data, economies of scale by working on a national level and the possibilities to contextualize more in-depth regional research in a national context (Mandemakers, 1989, p. 96). The HSN was not the first database in this field; other databases already existed, in particular in Sweden and Québec (Hall, McCaa, & Thorvaldsen, 2000). The Dutch example distinguished itself by being a sample (compared with Sweden) working on a national instead of regional scale (see also the HSN website: https://iisg.amsterdam/en/hsn).

In 1991, the HSN found its home at the International Institute of Social History (IISH) in Amsterdam and a pilot project started with a sample from birth certificates in the province of Utrecht for the birth period 1812–1922. In 1993, the HSN started collecting data from population registers, as well. In the beginning this was done only for additional projects and on a limited scale. Since 1991 about 12 million Euro has been invested in the database and related projects: 3 million by the IISH and 9 million by way of external funding, especially the Dutch Research Council (NWO). Besides the basic sample, about 25 additional projects have been realized. This has created all kind of extensions to the database. A special project, LINKS, uses the indices of names from the Dutch civil registration to reconstruct pedigrees (for the period 1780–1940) and complete families (1811–1900) for the whole of the Netherlands or parts of it. The HSN and the additional projects have been widely used in hundreds of presentations and scientific publications. By March 2019 there has been about 400 scientific publications and 15 dissertations completed wholly or partly based on HSN or LINKS data. The first results of these studies were communicated in over 700 worldwide presentations. The large output can be linked to the fact that the data has always been freely available. In a quantitative sense, HSN can certainly be considered to have been a game-changer in the fields of Dutch historical science, especially historical demography and historical sociology. But what has been its impact in a qualitative sense? With HSN,

1 The first overview was given by Hall, McCaa, and Thorvaldsen (2000). For the most comprehensive overview see the website of the European Historical Population Samples network: https://ehps-net.eu/databases.


3 The data of both HSN and LINKS are freely available for scientific research after signing a license agreement. License forms can be requested by writing an email to hsn@iisg.nl. By signing a license the researcher guarantees the private character of the data, to use the data only for science in a non-commercial way and to share the results of his/her research with the HSN, for more information, see https://iisg.amsterdam/en/hsn/privacy-statement; https://iisg.amsterdam/en/hsn/products.
Dutch social science historians have added a new perspective to Dutch history. But they also analyzed the Dutch experience in the context of more international debates on ‘universal’ behavioral patterns. Regularly, this contribution took the form of comparative projects.

We especially focus on debates: to what (old) questions has HSN provided the answers? Have definite outcomes been reached or is research still going on? In this article we will present an overview of the research that was done on the original themes and the new fields that were introduced over the years. We will also go into methodological issues that were picked up by the ‘HSN community’ and which helped the field forwards. In the last section we will point out the present and future challenges for the HSN, both in terms of methods and data.

2 STRATEGY AND DATABASE DEVELOPMENT OF THE HSN

Before answering the main question, we will present a more detailed overview of the development of the HSN database. The incremental growth of the database implies that earlier studies were based on different releases (in terms of period, region and sources) than later ones. Thus, the development of HSN needs to be explained to understand shifts in research questions. What did the initiators actually have in mind when starting the HSN and were their ambitions realized? The HSN operates in a dynamic setting. Firstly, the scientific fields in which the HSN operates are constantly changing in focus and methods. Over and over again new questions have to be answered, leading to the need for extensions of the HSN dataset, especially by way of additional projects, resulting in extra samples and/or including data from new sources. Secondly, the HSN itself plays an important role in agenda setting and the dynamics of research, especially in the Netherlands. An example is the introduction of the integrating concept of the ‘life course’ in Dutch historical research for which the data collected by the HSN were extremely useful (Kok, 2000). Thirdly, the possibilities for research were also determined by the logic in the construction of the database itself in coherence with the way investment grants were allocated. Fourthly, the HSN became an international player in the development and documentation of best practices in the field of historical life course databases.

The database activities of the HSN can be distinguished in three categories: a) projects that are focused on the construction of the central database which is the HSN basic sample with life courses, b) projects that are initiated by researchers or institutions adding new research persons to the database (for example siblings or children of existing research persons) and/or adding data that are not considered to be part of the central database such as cadastral or tax data and c) projects to develop software for data entry and data integration and to document the expertise by which all projects are accomplished, archived and released for research (Mandemakers, 2001a, 2006a). The focus of the central database is the systematic collection of data from civil certificates (birth, marriage and death) and the population register. The research persons (RPs) of the HSN are selected by way of a stratified random sample from the birth certificates of the period 1812–1922. The strata are defined by ten yearly time periods and by region. The regional stratification was based on the administrative division in provinces and the provinces were further stratified in urban and countryside areas. The province of North-Holland for example was stratified in four parts: Amsterdam, large cities (Haarlem, Alkmaar, Hoorn and Zaandam), small cities and the countryside. Eleven time periods and 25 regional areas made 275 strata all together, making sure that the sample of the HSN would be representative in time and in geographical perspective. This representativeness was further guaranteed by varying the sample fraction from 0.75% for the period 1812–1872 to 0.5% for the period 1873–1922 (Mandemakers, 2000, pp. 151–155). The difference between these two levels was motivated by the wish to get a more or less equal absolute number of survivors at the age of 16 for each ten yearly cohort, depending on the number of births and the infant and child mortality which varied between 15 and 40% for the period till 1870 (Engelen, 2009; van Poppel & Mandemakers, 2002). The number of sampled persons for the central database amounts to about 85,500 persons. This is about 0.6% of the 14.5 million persons that were born in this period in the Netherlands. The goal of the HSN is the reconstruction of the life courses of all these sampled persons.

For a description of the variables that can be collected from these sources, see the website of the HSN (https://iisg.amsterdam/en/hsn/data/sources) or Mandemakers (2006b), Vulsma (2002).
Figure 1 shows the development of the construction of the HSN central database. In 1991 the HSN started with a pilot project in the province of Utrecht that did not only focus on entering birth certificates but was also directed at the data entry of death certificates of persons who died before the age of ten, of personal cards administered for all persons living in the Netherlands from the 1st of January 1940 onwards and of marriage certificates. In 1993 the pilot project ended successfully (Mandemakers & Boonstra, 1995) and new funding was received to continue in the same way with the provinces of Zeeland and Zuid-Holland. By way of grants from investment funds of the Dutch Research Council (NWO) the HSN was scaled up to the nationwide level between 1996 and 2002. During this period most of the sampled persons were entered into the database. Only the period 1903–1922 lagged behind, since initially for this period a sample of 0.25% was constructed. Getting a sample with a comparable number of surviving 16-year-olds for each time period, on average 0.35% to 0.4% would be the norm for this period. For reasons of comparability and because three provinces had already been sampled at the 0.5% level, it was decided to bring the whole period to the same level of 0.5%. Data entry of death and marriage certificates was limited to these certificates that were easy to find, searching only in the municipality of birth and neighboring municipalities and in case of the death certificates only for those who died at a young age. From 2010 onwards it became much easier to find death and marriage certificates because of the digital indices that were created by the Dutch regional archives mobilizing thousands of volunteers. And nowadays more and more persons can be traced in the scans of all civil certificates made available by the municipal and regional archives (see https://www.wiewaswie.nl/en). This means that deaths outside the municipality of birth are getting more and more complete, especially for the births after 1862 when all life courses have been followed by way of population registers.

Soon, researchers showed up expressing the wish to collect data from population registers. In a limited way this was started up in 1993 in the form of three additional projects: Migration in the province of Utrecht (Jan Kok), the epidemiological project Reduced fecundity because of maternal high-risk conceptions (Luc Smits, Gerard Zielhuis and Piet Jongbloet) and Regional differences in demographic behaviour, the Netherlands, 1900–1960 (Angélique Janssens). The system of population registers was introduced in the Netherlands in 1850, following the Belgian lead. Basically, the registers updated census information on a day-to-day basis: this means that every change in household composition (through birth, death, marriage, arrival, or departure) was to be recorded with a proper date. The registers also provide information such as place and date of birth, religion, occupation and (family) relation to the head of the household. As people can be traced in all their households across the country, this source allows for a complete and very detailed life course reconstruction of RPs born after 1850 (Mandemakers, 2006b).
With a great investment impulse from NWO, for the project *Life Courses in Context* (2003–2008), it became possible to collect all population registers for the then sampled persons from the birth period 1863–1922 in a systematic way (Mandemakers, 2004). The outcomes of this project also provided sufficient information to collect death and marriage certificates for these persons. Priority was given to the sampled persons from the provinces Utrecht, Friesland, Zeeland and Rotterdam. This priority was meant to quickly offer researchers workable datasets, since for these regions a lot of data from the registers had already been collected within the framework of the project *Early-life conditions, Social mobility and Longevity* (George Alter and Frans van Poppel). This also implied that for this group the period of birth was extended to 1850–1862. Besides the RP itself, other relatives are included in the database as long as they lived together in a family or household context with the RP. For an average RP who reached the age of 60 years we count about 26 other persons (2 parents, 2 parents-in-law, 1 spouse, 8 siblings, 6 children, 6 witnesses from the certificates and 2 other kinds of kinship).

Since the early 1990s, more than a thousand volunteers have digitized all names, ages and partly also occupational titles and birth places from birth, marriage and death certificates as far as they are open to the public. Initially the database was known as GENLIAS, now it is called WieWasWie (WhoWasWho’, abbreviated as WWW) and presently it includes over 200 million persons’ appearances. All digitized data are publicly accessible through the website of WieWasWie (‘WhoWasWho’, https://www.wiewaswie.nl/en), based at the Dutch Center for Family History (CBG).

The main part of WWW consists of indices from civil certificates. For the whole of the Netherlands the civil registration system started in 1811 and the registers are public with a delay of 100, 75 or 50 years depending on the type of certificate. At present, the registers of birth until 1919, the marriage registers until 1944 and the registers of death until 1969 are public. In the middle of 2018, about 26 million civil certificates have been digitized, containing information on about 110 million appearances of (not unique) persons. As figure 2 shows the marriage registers have been indexed quite completely, lagging behind are the birth and to a lesser extent the death certificates, especially those of Amsterdam. Basically, the system includes the full names of the main actors in these certificates: the birth, the

---

5 For more information about all additional projects, see [https://iisg.amsterdam/en/hsn/projects](https://iisg.amsterdam/en/hsn/projects) and the annual reports of the HSN ([https://iisg.amsterdam/en/hsn/annual-reports](https://iisg.amsterdam/en/hsn/annual-reports)).

6 CBG is the national center for knowledge, documentation and publicity for genealogy and heraldry. It is a non-commercial organization, supported by about 40,000 contributors. The work of the volunteers is managed by more than twenty cooperating regional archives.
death, the bride and the groom and their parents. Since 2005 the HSN has distributed data from this system. With the LINKS program a more systematic approach was started in 2010.\(^7\) LINKS stands for \textit{LINKing System for historical family reconstruction}. By way of record linkage software certificates are being linked to create pedigrees and families. This matching is relatively easy, because the parents are always included in the data from the WhoWasWho system and because the original names of females are always kept in Dutch certificates. Other included variables are the age of the death and of the bride and groom and for some provinces the occupational titles of all actors (about 60\% part of the data from the marriage records).\(^8\) After all certificates have been entered, it will encompass the reconstruction of all 19th and early 20th century families in the Netherlands from 1840 onwards.

The products of the LINKS software system are delivered in specific releases, but the dataset as a whole is known as the LINKS database. And although these data must be seen as a separate database, it is very much connected to the HSN, in the sense of creation, content and dissemination, so we will integrate the results from LINKS with the results from the HSN in the following chapters. Based on birth certificates, all HSN Research Persons and its family members are included in this kinship system as well. This opens up enormous possibilities to enrich the existing life histories with kinship in the second, third and fourth degree. The first release with an integration of the HSN Life Courses and the LINKS database is expected for 2020. A systematic comparison of both HSN and LINKS has been made by van den Berg, van Dijk, Mourits, Slagboom, Janssens, and Mandemakers (2020). On the basis of a sample of an integrated dataset from the province of Zeeland they compared the average number of children within families, the average age at death for RPs and their kin and other demographic indicators. They concluded that RPs own families were reconstructed very thoroughly in HSN, which is not biased by out-migration as is the case with LINKS. However, information on parental families is more complete in LINKS.

This overview is limited to the basic concepts of both HSN and LINKS. Two separate articles are prepared to be published in this journal. In these articles both databases will be described from a more technical point of view. For the HSN this will include an overview of all projects that have been undertaken since the beginning including a description of all datasets that have been created as oversamplings from the main one (Mandemakers, forthcoming). For LINKS we will describe in more detail the way the software system works and the amount of data that are available (Mandemakers, Bloothoof, & Laan, forthcoming). Till now the HSN released about 70 data releases, often in relationship with a specific project. The first release of LINKS data appeared in 2005, since then over 30 data releases have been made available.\(^9\)

### 3 NEW ANSWERS TO CLASSIC RESEARCH QUESTIONS

#### 3.1 THE AMBITIONS OF THE HSN

In 1989, when the plans for the HSN were publicly announced, five research fields were explicitly mentioned that would profit on what was also called the 'birth-bank' (Mandemakers, 1989). The following examples were given:

1. Social stratification
2. Composition of the workforce
3. Social mobility
4. Migration
5. Demographic themes (mortality, marriage frequency, marriage age, marital fertility and bridal pregnancies)

\(^7\) LINKS was granted through the NWO CATCH program, see https://www.nwo.nl/en/research-and-results/programmes/Continuous+Access+To+Cultural+Heritage+(CATCH) and https://iisg.amsterdam/en/hsn/projects/links

\(^8\) For the death and especially the birth certificates the percentage of included occupational titles is much lower.

Although we know that these five themes were not meant to be exhaustive but were only presented as examples of possible research fields, we are still surprised by the original modest expectations compared to the underlying vision of the plans. The first two themes are quite related and we will band them together under the umbrella of social stratification. The HSN proved fruitful in getting a grip on the old theme of social mobility as it enabled Dutch researchers to build on the paradigm change by Blau and Duncan (1967) redirecting the focus from social mobility tables to an approach with status achievement models based on micro data (Heath, 1981). Classic debates in historical demography concern social inequality in relation with death, household composition, marriage (and its ‘Malthusian’ connection to survival and economic independence), and, above all, the demographic transition (Coale & Watkins, 1986). From the onset, HSN researchers used the new data to contribute to these debates. Together with social mobility they form the nucleus of this section.

The migration theme we consider as quite new since in the beginning this was only envisaged in a rather limited way by combining civil certificates (birth, marriage and death). This supplied a rough indication of migration histories. During the development of the database the use of the population register for migration studies became a very innovative aspect of the HSN. Since section 4 focuses on new types of research, we will handle migration there. But, of course the difference between new and old research themes is somewhat arbitrary.

3.2 LIFE CHANCES AND SOCIAL BACKGROUND

We will first briefly discuss the theme of social stratification. This topic was quickly superseded by research interest in the process of social mobility. The changing society as such did not become central in research rather it was chances of an individual to proceed into another social layer than the one he or she was born in. More and more social mobility was considered a process of status attainment. Research into processes of social background and mobility is interwoven with all kinds of other topics such as migration, nuptiality and mortality. So, also within the HSN community, social mobility is studied from all point of views. For the sake of argument, we have made a distinction into different sections: marriage mobility, status attainment, intragenerational or career mobility, educational mobility and female-specific aspects of mobility. However, since part of the mobility studies are overlapping with studies in the field of demography, some of the studies will not be discussed in this section but in the one on demography.

3.2.1 SOCIAL STRATIFICATION

The first theme focused on the solution of what became known as the ‘great question of the social stratification’ of the Netherlands between 1850 and 1940. The discussion was started by Giele and van Oenen (1974) and circled around two questions: should a social stratification be constructed from the perspective of social class or of social status, and subsequently how did this social structure develop (Boonstra & Mandemakers, 1995)? For all research persons born from 1812 until 1872 it would be possible to reconstruct a complete census for each moment in time until the age of 59 from 1872 onwards, for those born from 1812 until 1882 until the age of 69 from 1873 onwards, etc. However, until now the HSN-database has been developed mostly for persons born after 1863, so it is not surprising that the HSN has not been able to revitalize this debate. Anyway, not much new work has been published either. Exceptions are articles by Mandemakers (2001b) covering the period 1850–1990 and van Leeuwen and Maas (2007a) on economic specialization in which only the last authors made use of the HSN-database. Besides, the introduction of micro-data made the focus shift from changing class structures to individual chances of upward or downward social mobility.

Related to the theme of social stratification is the composition of the workforce. A problem here was the lack of the occupational censuses of 1869 and 1879. The 1869 census did not take place at all and the results of 1879 were only used as a mechanism to check the locations of residence (van Maarseveen, 2002). Again, as in the case of social stratification, the HSN was expected to be able to make a simulation of the census, and for the same reason (no life courses from the birth period before 1863) it did not work out. However, the HSN dataset was used to gain more knowledge on all kinds of aspects of the labor market, such as the position of migrants (Delger, 2003) or the occupations of women. The occupational titles on marriage certificates of the HSN were especially suitable to trace changes in registered female labor (van Poppel, van Dalen, & Walhout, 2009); combined with certificates of the LiNKs database (Maas & van Leeuwen, 2006) and for a sample of twelve municipalities (Walhout & van Poppel, 2003).
3.2.2 INTERGENERATIONAL AND MARRIAGE MOBILITY

Intergenerational mobility has mainly been studied by using data from the marriage certificates which include the occupational title of both bride and groom and all four parents. Marriage mobility is a good indicator for the ‘openness’ of societies, and researchers expect that status and class differences in partner choice have diminished after 1850. This supposedly resulted from the increasing importance of ‘romantic love’, and from growing possibilities to reach a better social position than their parents, thanks to education and their own achievements (Kok & van Leeuwen, 2005).

Van Leeuwen and Maas (1995, 1997) started their research on the openness of society with the data of the Utrecht pilot project. They concluded that although there had been an increase in absolute social mobility this was a consequence of the growing opportunities in society while the relative mobility (mobility controlling for changes in the occupational structure of society) remained the same till 1940, at least in the province of Utrecht. Since then, several studies have been published about the development of marriage mobility in the Netherlands, also in comparison with other countries. More in general, it was shown that both for the country as a whole, as for urban and countryside areas, the association between social mobility and economic modernization could be researched and differentiated. Van Leeuwen, Maas, and Mandemakers (2005) showed that endogamy decreased in the second half of the 19th century, because children from the diminishing group of farmers had to move elsewhere to find marriage partners.

Bras and Kok (2005) researched marriage mobility to measure which marriages were homogeneous or heterogeneous in terms of social class background by comparing the social background of fathers and fathers-in-law. They used one of the first versions of the GENLIA ZEELAND dataset with marriages from 1796 till 1922. They did not find much influence of parental characteristics on having a heterogeneous marriage or not, except for the farmers who showed a strong inclination to marry in their own group (about 60 to 70%).

Kok and Mandemakers (2008) concentrated on the marriage market as a mechanism of partner selection and concluded that personal circumstances such as poverty, illiteracy and low social status decreased the chance of finding partners outside the local community. This effect was stronger in geographically more isolated communities. Concentrating on farmers, they found that it were wealthy farmers who were most able to place their children on a farm of their own, not through the choice of a successor to the existing farm, but by negotiating and bringing together the assets for a new farm household. This transmission required that both the father of the groom and bride were still alive (Kok, Mandemakers, & Damsma, 2010).

On the basis of a LINKS dataset of six provinces, Maas and van Leeuwen (2019), tested the modernization hypothesis again and confirmed that the occupational status of the father had become less important for partner selection in the second half of the 19th century. In cities, they found evidence that achieved characteristics became less important. This is a partial confirmation of the ‘romantic love’ hypothesis stating that both achievement and ascription became less important in partner selection.

3.2.3 STATUS ATTAINMENT

One of the first LINKS datasets consisted of the linked marriage certificates of five provinces of the Netherlands including about 0.4 million certificates. This dataset was widely used by sociologists to test the modernization theory which states that modernizing and industrializing societies become more open in terms of social mobility. Knigge (2015) used several status attainment models not only looking at the direct relationship between fathers and sons but including the association between brothers in his models as well. Covering the period 1827 to 1897 he found that the status correlation between father and son was relatively high (0.57) compared to 20th century societies. Remarkable was the high sibling correlation of 0.53 of which only 0.32 could be explained by the effects of the father alone. This implies strong family effects not directly passed on by the father. Parts of the family effect shown by siblings could be explained by direct effects of grandfathers and even great-grandfathers and these were relatively stable effects which endured over time and were independent of the chance of having direct or only indirect contact with the siblings (Knigge, 2016). The average effects decreased gradually after 1850 under the influence of some but not all modernization processes that changed Dutch society; after 1850 only the effects of urbanization and mass transportation showed significant results. Sibling cohesion proved to be weaker when age differences were relatively high, the status of the father shows more fluctuations and migratory moves interrupted the effect of birth ranking (Knigge, Maas, & van Leeuwen, 2014; Knigge, Maas, van Leeuwen, & Mandemakers, 2014). Sprok (2013)
found a negative effect of the number of siblings on one’s social status, especially brothers. Bras, Kok, and Mandemakers (2010) also studied siblings’ effects on status attainment through the perspective of resource dilution. They showed that the effects clearly depended on the social and temporal setting.

Zijdeman (2010) researched status attainment in several ways, using amongst others the LINKS dataset of the province of Zeeland covering the period 1811–1915. He not only included micro-data but also a large number of contextual data to control for changes in environment on the status attainment process. For the father-son relationship he found a more or less linearly decreasing association in social status; for the father-in-law relationship this was the opposite: as soon as the context became more industrialized and mass transportation developed, the association with the status of the father-in-law started to increase. However, as soon as the educational expansion became serious, the association with the parent-in-law began to decrease again (Zijdeman, 2009; Zijdeman & Maas, 2010).

To study the influence of religious background on life courses a special sample with an oversampling of Jews was made for The Hague. It was concluded that the 19th century social differences between Protestants, Jews and Roman Catholics were a consequence of a slowly changing class-based society and not of discrimination, at least the chances to change the social position proved to be the same for all religious groups (van Poppel, Liefbroer, & Schellekens, 2003). Within the context of a study into the occupational structure of Jews in Amsterdam between 1850 and 1940, Tammes (2012b) investigated the effect of different decisions regarding Jewishness on social status. He found that marrying a gentile or converting to Christianity had no effect on social status, however giving up one’s religion was strongly associated with upward social mobility.

3.2.4 SOCIAL AND ECONOMIC CAREERS

Career mobility, the possibilities for persons to change their occupational position during their lives, could also be studied by using the HSN database. Maas and van Leeuwen (2009) researched to what extent occupational careers between 1865 and 1940 were to be considered as stable, upwardly or downwardly mobile both for males and females. They concluded that over the whole period mobility increased, both upward and downward. Where the average level of the occupational structure became higher as well, the male’s total upward mobility equalled downward mobility. For females they found net upward mobility, although this could probably be explained by selection issues whereas females with lower estimated occupations disappeared from the registrations when they grew older. So, the influence of the social background did not change where it was expected that this influence would become less important over the years.

In her doctoral thesis, Schulz (2013) investigated the development of individual careers during the process of modernization and industrialization in the Netherlands. She tested if the role of the parents in the status attainment process diminished in favor of the role of education. She used the HSN dataset by employing multilevel growth models and found that career mobility decreased at older ages and that a father with a higher social status provided a much better start than other fathers, although this effect diminished over time. However, for females this was less clear and — unexpectedly — it turned out that being married had a positive effect on occupational status (Schulz & Maas, 2010). When looking at the development over time by comparing three cohorts active between 1865 and 1920, she found that intragenerational effects became increasingly more important for successful careers whereas the influence of social background diminished. Most mobility occurred at the start of the career, although for females this was less clear than for males maybe as a consequence of underreporting, especially by lower status groups (Schulz & Maas, 2012). Modernization indicated through, for instance, the level of educational opportunities and the number of steam engines proved to have no positive influence on career mobility, and the proximity of train stations even had a negative effect (Schulz, Maas, & van Leeuwen, 2015). More in general, females having a career proved to be more successful than males (Schulz, 2015).

The HSN was also used to zoom in on specific occupational groups. Boonstra (2011) used the HSN to study the intergenerational mobility of teachers, searching for real ‘educational’ families. He found indeed some ‘hereditary transmission’, but in general he concluded that it was a quite open profession during the 19th and early 20th century especially attracting newcomers from lower social backgrounds. Heerma van Voss and Vermeulen (2000) compared the life courses of active union members in the textile industry from the Twente region with a control group and concluded that the union members showed more upward social mobility. This could be explained by the large group of workers with a farming background in the control group. For them going back to a career in agriculture
was an opportunity that did not disappear when they started working in the textile factories. Finally, Bras (1998) studied females working in domestic services and concluded that this typical job, while it resulted in more geographical mobility, did not result in a higher social status except when associated with long-distance migration to the cities.

3.2.5 EDUCATION AND SOCIAL BACKGROUND

It was already known that literacy levels were relatively high for the Netherlands, compared with other European countries. In general, Protestant countries showed high levels and Roman Catholic countries rather low levels. In the Netherlands, the Roman Catholics who lived in Protestant-dominated areas caught up with the Protestants. Only the two southern provinces with a population of about 90% Roman Catholics, Noord-Brabant and Limburg, showed relatively low levels of literacy around 1800 (van der Woude, 1980).

Boonstra (1995) used the signatures in the first entered marriage certificates to make an overview of the development of literacy in the province of Utrecht. He concluded that at the start of the 19th century illiteracy had already receded to the lower social economic classes and to some remote areas in the province. After the middle of the century, illiteracy was not a factor of importance anymore. In a second study, Boonstra (2009a) could use the complete set of HSN birth certificates covering the whole of the Netherlands. Based on whether or not the father was able to sign, he studied the development of illiteracy from the birth year 1775 onwards. Illiteracy dropped from an average 30% to almost 0% for persons born in 1880. In this development the males took the lead. As said, in the early 19th century illiteracy levels were already very low in the north of the country except for the villages and towns that specialized in sea fishing. Boonstra also showed that the peat areas and the very conservative orthodox Protestant areas had rather low literacy levels as well.

Boonstra (2009b) extended his study of illiteracy and social and regional background into social mobility by combining the data on signatures and occupations from the certificates. He found that chances of both downward and upward intragenerational mobility were not influenced by being literate or not. However, on the intergenerational level it was important: being a son of a literate father gave a significantly higher chance of upward mobility and having an illiterate father gave a higher chance of downward mobility. Vandezande, Matthijs, and Kok (2011) tested the so-called resource dilution hypothesis on education by combining data on the number of children in a household and the capability to sign a marriage certificate. Having more brothers lowered the chance of being literate. The number of sisters had a different effect in the sense that having older sisters improved the chance of being literate for the younger sisters, but not for the younger brothers.

Zijdeman and Mandemakers (2008) researched the relation between social mobility and education. They compared a dataset with marriage certificates from pupils with a higher general secondary education background (HBS and Gymnasium) with a control group from the HSN database. They concluded that the direct importance of the parental background for the later social position of sons became less with a constant indirect influence by way of the educational achievement between 1880 and 1920. The influence of the educational achievement as such on the social position at marriage remained constant. The same dataset was used in an investigation of the impact of the HBS system on the labor market for white collar and professional occupations (Schalk, 2015).

3.2.6 GENDER, WOMEN AND WORK

During the 19th century the ideal of the ‘male breadwinner’ and the ‘housewife’ spread through the Netherlands. Van Poppel, van Dalen, and Walhout (2009) researched the way this diffusion of a cultural ideal took place on the basis of a GENLIAS dataset of marriage certificates. They found that the ideal diffused from the upper to the lower classes starting around 1830 in the urban areas and around 1850 in the rural areas of the Netherlands. Around 1830 on average about 30% of the marriage certificates from the different layers in the lower classes showed brides without an occupation at the moment of marriage, around 1910 about 60%, and within the group of farmers, 80%. For the upper classes this had always been the norm with a percentage that moved from 90% in the beginning of the 19th century to 97% around 1880 to start declining back to 90% around 1910.

Schulz, Maas, and van Leeuwen (2014) continued research on this question with the same dataset, concluding that not only the inclusion of females in the labor market dropped until about 1885 but
also the social prestige of the work they practiced. This trend was reversed by growing educational participation and dissemination of more egalitarian gender values after 1900, females changing from occupations such as housemaid to primary teacher, telephone operators, clerks, etc. On the basis of an oversampling of the HSN for the textile city of Enschede, Janssens (1998) researched the relationship between social background, occupation and religion in the female life course. Her main conclusion was that religion was the main factor in decisions regarding marriage and fertility and not the type of work in which women were engaged.

Van Leeuwen and Zijdeman (2010) used the Zeeland marriage certificates in combination with all kinds of data on the level of the municipalities. Their main question was how society became more open in terms of social mobility during the modernization of society. They showed that ascribed characteristics became less important while the effects of achieved characteristics were stable in the process of marrying a bride with a relatively high social status. Although the ‘logic of industrialization’ still proved the moving factor in this process, it turned out that on the municipal level characteristics of industrialization were less systematically related with changes in ascription and achievement than one might expect.

The LINKS marriage certificates dataset was also used to research why there was a decreasing female participation labor rate in the Netherlands in the course of the 19th century (Boter, 2017; van Poppel, van Dalen, & Walhout, 2009). As both brides and grooms were asked to state their profession at the moment of marriage, it was possible to research this phenomenon on the micro-level. Cultural reasons like the advance of the male breadwinners ideal and the increased income for large parts of society made it for women less necessary to work. Where van Poppel, van Dalen, and Walhout used only the provinces and period as independent variables, Boter researched the developments at the municipal level. She showed that also the changing economic structure was causing diminishing female labor participation, whereas in some areas the reverse trend was visible (e.g. textile industry).

For her dissertation, Bras (2002, 2004) concentrated on the career mobility of maids from Zeeland and found that for half of them the career did not last longer than five years and that there were no serious mobility routes in terms of the social background of the employers. The social background of the maid was also positively related with the social background of the family in which she got a position. In general maids from middle class background migrated more to cities which gave them higher chances of upward social mobility and gaining better living conditions.

### 3.3 DEMOGRAPHIC TOPICS

#### 3.3.1 MORTALITY ACROSS THE LIFE COURSE

Since the middle of the 19th century scholars have debated the effect of industrialization on living standards, for which (infant) mortality quickly came to be seen as a good indicator. ‘Pessimists’ predicted that increasing inequality during industrialization — combined with the receding effects of epidemics — would lead to increased social class differentials in mortality. ‘Optimists’ pointed at higher real wages, which would soon be translated into better housing and food, and in a lower social gap in mortality rates. Yet others predicted that social class differences would remain constant over time, as the rich would always be able to translate their resources into better health (Bengtsson & van Poppel, 2011). Already from the start, HSN was used to add to this debate. The first papers dealt with infant and child mortality (e.g. van Poppel, 1995; van Poppel & Mandemakers, 1997; van Poppel & Mandemakers, 2002; Woods, Løkke, & van Poppel, 2006). The fact that the data covered large regions, with distinctly different health environments and traditions, made it possible to study class differentials in infant and child mortality across regions. Van Poppel, Jonker, and Mandemakers (2005) discovered clear social gradients in both infant and child mortality. However, the region in which a child was born was more important than the class position of its father. Contrary to what was expected, they found that in a ‘healthy region’ (higher quality of surface water), in combination with breastfeeding practices, social class differences were limited compared to more unhealthy regions. A recent study using the LINKS data confirmed that limited breastfeeding in Zeeland, associated with female field work, led to an outspoken pattern of high infant mortality in the first summer after birth (van Poppel, Ekamper, & Mandemakers, 2018). High temperatures would increase mortality risks even more (Ekamper, van Poppel, van Duin, & Garssen, 2009; Ekamper, van Poppel, van Duin, & Mandemakers, 2010).
Infant mortality is also the subject of a long-term debate on the role of the Roman Catholic clergy, who, supposedly, in the late 19th century began to discourage mothers to breastfeed their babies thus causing higher infant mortality in Catholic regions (on religious differentials, see also van Poppel, Schellekens, and Liebrot 2002). Supposedly, this prudishness was part of their ‘moral campaign’: in their fierce competition with Protestants, Roman Catholics aimed to demonstrate they had the highest moral standards. Although this particular role of the clergy has recently been put into question, it is still possible that Catholic mothers were less inclined to breastfeed their infants. Inspecting birth intervals by religion, Janssens and Pelzer (2014) concluded that in several places (in the period 1880–1920) Catholic mothers were less likely to breastfeed than mothers from other denominations. However, this did not translate into Catholic excess mortality. In their view, region was the most important factor in explaining infant mortality. But in her recent doctoral dissertation, Walhout (2019) concluded that religion, after all, superseded region in explaining the higher infant mortality risks of Catholic babies.

Illegitimacy has always been related to high risks of dying in infancy. The question remained whether these high risks could be fully explained by the social conditions of the single mothers — poverty, exclusion from poor relief, social isolation and inability to combine work with breastfeeding — or whether infanticide was also involved. By comparing with poor mothers, van Poppel, Kok, and Kruse (1997) concluded that infanticide, or at least, willful neglect played an important role in high mortality of illegitimates.

Social class also appears to be the most important factor in explaining excess mortality of teenage girls, a phenomenon that occurred in most European countries before 1940. In the Netherlands, the problem seems to have been limited to daughters of unskilled workers. Van Poppel, Schellekens, and Walhout (2009) explain the relatively favorable situation for Dutch girls from "the dominant position of small family farms, the preponderance of mixed or dairy farming, the well-integrated position of women in market production and, more generally, the higher degree of equality between men and women" (p. 37). In contrast, the wives of farmers and skilled workers ran higher risks of maternal mortality than women from other social groups, probably because women in these groups tended to perform heavy labor in advanced stages of their pregnancy (Ory & van Poppel, 2013).

As to mortality at older ages, social class differences turned out to be even less conspicuous. Although somewhat paradoxical, van Poppel, Jennissen, and Mandemakers (2009) found relatively strong differences for the provinces of Zeeland and Limburg. Van Poppel and van Gaalen (2008a) concluded that there was no sign of living standards of workers deteriorating during industrialization and urbanization which they interpret, rather reluctantly, in line of the ‘optimist school’: the lack of class differences relate to relatively high real wages and a quite generous poor relief in The Netherlands. Their results were corroborated by Schenk and van Poppel (2011), who state that Dutch economic growth benefited the lower classes, who “took advantage of the new possibilities created by increased medical knowledge, improved sanitary standards, which in their turn were not independent of the increased economic growth” (p. 415).

### 3.3.2 HOUSEHOLD COMPOSITION AND ITS DEMOGRAPHIC IMPLICATIONS

One of the oldest topics in historical demography is household composition, for which the Cambridge Group for the History of Population and Social Structure created elaborate schemes to make distinctions between types of household extensions (Hammel & Laslett, 1974). One of the most important findings in this line of research was that the nuclear family was not the outcome of processes of industrialization and urbanization, but that it had antedated them by at least several centuries. The Netherlands clearly fits into the English pattern of nuclear families, but the HSN made it possible to study regional differences and the role of agriculture in household extension in greater detail. HSN data allowed to map two regions of extended families. One along the eastern border with Germany where a stem family system persisted, even in the face of legally prescribed equal division of inheritance. And another one in the south, which was related, ironically, to strict partibility coupled with struggling small farms. In this region, adult (unmarried) children tended to stick together in order to hold on to their share (Kok & Mandemakers, 2009). The HSN also makes it possible to look at extended households from the perspective of persons potentially in need of care, such as orphans, unmarried persons, widows, widowers and the elderly. Research showed that many of those found recourse, at least temporarily, in the homes of family. Especially in the northern and western regions, more commercialized and urban areas of the Netherlands, household extension was often related to a crisis situation. E.g. a couple took
in an ageing parent who could no longer live independently. This research put the notion of ‘nuclear hardship’ in perspective (Kok & Mandemakers, 2010, 2012).

Another approach to households in the HSN is to take a long-term perspective and to compare the (prospective) data from the HSN with (retrospective) panel surveys. In this way, van Poppel, Schenk, and van Gaalen (2013) created a continuous birth cohort from 1850 until 1985. They studied the living arrangements of children, and were able to dispel the notion that nowadays more children are living in broken or reconstituted families than ever. They discovered that, due to the decline in adult mortality, the decrease in out-of-wedlock fertility, and rather low divorce frequencies, the cohorts born between 1900 and 1965 were spending more time with near kin than “at any time in history” (p. 255). Also, the traditional, complete nuclear family in which children live with both their biological parents is nowadays still stronger than it was in the middle of the 19th century (see also van Poppel, Tammes, and Schenk, 2010; on urban-rural differences in the living arrangements of children, see also van Gaalen, 2007; van Gaalen and van Poppel, 2009). Lin (2011) studied female headship using the HSN data for Rotterdam, and contrasted it to patterns in Taipei, Taiwan. Incidence, timing and duration of female headship differed strongly, which Lin sees as a (new) indicator of differences between European and East-Asian family formation systems.

In his recent dissertation, Rosenbaum-Feldbrügge (2019) used the HSN to study the effects of parental loss on a number of outcomes: the remarriage or migration of the widow of widower (also Rosenbaum-Feldbrügge, 2018), and the impact on the timing and incidence of leaving home and marriage of the children. Furthermore, he looked at the long-term effects of bereavement on the occupational status of (half-)orphans in early adulthood. The findings support to some extent the ‘niche’ hypothesis that the marriage of sons of farmers and skilled workers was advanced by parental death (Rosenbaum-Feldbrügge & Debiasi, 2019). The study shows that the death of the mother had more devastating and lasting effects than the death of the father, e.g. in terms of status attainment. The results point at the more pivotal role mothers played in care and upbringing — combined with the generous Dutch poor relief which may have made widows less vulnerable than might be expected.

### 3.3.3 Nuptiality

The study of marriage frequency and timing has always been core business of historical demography. How was marriage related to the requirement of an independent household? Why did property-less people seem to abandon traditional restraint in the second half of the 19th century and began to marry more often and younger? To be sure, the data collection of marriage certificates in the HSN was soon superseded by the massive collection of marriage data available through the LINKS project which allowed not only the study of marriage timing, but also of intergenerational transmission of the age at marriage (van Poppel, Monden, & Mandemakers, 2008), the impact of birth order (Suanet & Bras, 2014), the development of the marriage market in a geographic sense (Ekamper, van Poppel, & Mandemakers, 2011) and the seasonality of marriage as a marker of secularization (Engelen, 2017).

One of the first possibilities with the LINKS dataset of marriage certificates, was the study of kin marriages (Bras, van Poppel, & Mandemakers, 2009). Kin marriages were rather common among farmers and the higher social strata, clearly as a way to consolidate and merge properties. The increase of cousin marriage among orthodox Protestants in the Bible Belt area of the Netherlands was interpreted as a form of identity preservation. Overall, kin marriages took place more often in the relatively isolated, inland provinces of the Netherlands. Finally, sibling set exchange marriages were a consequence of the enlarged supply of same-generation kin as a result of the demographic transition.

What could the more limited collection of marriage certificates in the HSN add to the impressive LINKS findings? Of course, the HSN datasets offers much more variables than available in the LINKS database. To begin with, Kalmijn (1995) used the first HSN release to link the decreasing marriage age with social class, urban areas and migration. Van der Velden (2012) used HSN data to compare the age of marriage of seafarers with the inland workforce and concluded that marriage ages did not differ widely while he had expected that seafarers would marry later than the inland males. Other groups of whom marriage behavior was studied include farmers (van Poppel, Ekamper, & van Solinge, 2007) and workers in Rotterdam (Kok, 2006a). The HSN was also employed to relate timing and incidence of marriage to the RPs family situation both early in the life course (Engelen & Kok, 2003) and at age 18 (Kok, 2014b). Thus, these researchers assessed the influence of religion, living in a city, region, social economic status, and the family composition on the likelihood to marry early, at a normal age, late
or not at all. This procedure also allowed to answer the question how late marriage and permanent celibacy were related; did the unmarried persons have the same 'profile' as those who married late? It was found, that indeed several groups of singles (e.g. youngest sisters) simply had waited too long in seeking a marriage partner. But others had a different profile, especially women from the upper classes, who could not combine education and fitting work with marriage. Finally, in the eastern part of the country late marriage was normal, but celibacy was not. Marriage was sure to come, but it depended on the timing of success (Engelen & Kok, 2003, p. 27; on the life courses of urban singles, see Kok & Mandemakers, 2016). A more sophisticated way of studying life course 'pathways' into family formation, and the shift in these pathways of time was performed by Bras, Liefbroer, and Elzinga (2010). They used sequence analysis to uncover life course trajectories into adulthood, and could demonstrate the gradual emergence of the typical mid-20th century 'standard' life course of early family formation. Forerunners in this process appear to have been "laboring class youths, farmers' daughters, youngsters of mixed religious background, and the urban-born" (p. 1030).

As mentioned earlier, HSN procedures have been applied in specific projects. An example is the reconstruction of two complete generations (including migrants) of the people who married in the North-Holland rural municipality of Akersloot (first generation), which was selected primarily because it had already started a population register in 1830. The sample has been used to study, among others, the effects of the differences in sibling sets on marriage chances. Bras and Kok (2016) found, among others, that resource dilution in this community was gender-specific: "the more sisters in the household, the more the marriage of boys and girls is postponed or forfeited. Clearly, girls meant a drain on the resources at home" (p. 202). Another research using this dataset studied the intergenerational transmission of the age at marriage in relation to social control. The hypothesis is that such transmission effects are strong in groups with weak social control mechanism. This hypothesis was supported by the finding that mother to daughter transmission was weaker among farmers and among Roman Catholics than among working and middle classes and Liberal Protestants (Kok & Van Bavel, 2009).

The Akersloot database has also been used to study partner choice and social reproduction among farmers (Kok, Mandemakers, & Damsma, 2010). Partner choice was also the subject of research based on an oversample of Amsterdam Jews (religious intermarriage, Tammes, 2010) and a sample of Germans living in Utrecht (intermarriage with other ethnicities, Schrover, 2001, 2003, 2004, 2005). A large special project entailed the reconstruction of the life courses of Germans, Italians and persons from the Dutch provinces Zeeland and Noord-Brabant migrating to Rotterdam. Two migration cohorts were created which were followed up to the third generation. This research into assimilation processes obviously also dealt with the likelihood of marrying natives, people from the home region, or other migrants (Lucassen, 2003, 2005; Chotkowski, 2006).

Partner choice studies using the HSN have dealt with the various preferences for endogamy (by locality, religion, and class) and how they interact (Kok & Mandemakers, 2008; using GenLias, Ekamper, van Poppel, van Duin, & Mandemakers, 2010; Maas & Zijdeman, 2010; Zijdeman & Maas, 2010). The question whether family pressure leads to more social endogamy was studied by van Leeuwen, Maas, and Mandemakers (2005) who looked at the presence of parents and other family as witnesses at the wedding. They conclude that in general, farmers excepted, there was little family pressure leading to endogamy. More recent studies take up the theme of migrant integration through marriage again. The city of Rotterdam is compared to Stockholm and Antwerp to enable a more detailed study of how characteristics of cities enable migrant integration (Puschmann, Van den Driessche, Grönberg, Van de Putte, & Matthijs, 2015). The authors conclude that migrants faced many difficulties and social exclusion, resulting in low chances of marrying natives. The situation in Rotterdam and Antwerp appears to have been better than in more industrialized Stockholm, which suggests that port cities do offer some routes to escape exclusion.

Marriage can of course be triggered by a pregnancy (see Kok, 2011 on forced marriages in HSN). Bridal pregnancies in relation to regional courtship norms are the subject of a recent paper by Kok, Bras, and Rotering (2016). Local customs, however, were not strongly related to levels of bridal pregnancy. The phenomenon was strongly concentrated in proletarian as well as in Protestant groups. There is evidence of parental tolerance for the sexual urges of (endogamous) young couples who posed no threat to the property transmission. But there are also indications that youths deliberately advanced a marriage (using a pregnancy as leverage) to gain independence.
Recently, Jennings and Gray (2017) related marriage data from the HSN to climate fluctuations culled from the archives of the Dutch meteorological institute. Contrary to Malthusian expectations, people did not marry because they felt secure of their prospects. On the contrary: "Adverse environmental and food security conditions consistently increase marriage in this historical context" (p. 257). The authors explain this outcome from intra household competition over scarce resources, a situation in which children might see leaving home and marriage as a way to improve their situation. Possibly rural children migrated to cities where it was easier to find a partner and marry.

Finally, HSN has been used to study remarriage (van Leeuwen & Maas, 2007b) and divorce. Kalmijn (2008) noted a clear social gradient in the likelihood of a marriage ending in divorce. Forerunners of the recent divorce explosion appear to have been the 'cultural elite'. Higher professionals, but not managers, were more likely to divorce than others. He notes that the lower classes might have chosen other options, such as simply leaving, to end an unhappy marriage.

### 3.3.4 FERTILITY

As elsewhere, entire generations of Dutch historical demographers have been puzzled by the fertility decline, and we can safely say that HSN has enlivened but not concluded the debate. Dutch demographers faced the challenge of having to explain two conspicuous features of demographic change. First, the onset of the transition had a marked regional pattern, with the north-western parts of the country being the first to show change and the south-eastern parts the last. Second, within Western Europe, the Netherlands as a whole was a latecomer, characterized by relatively high fertility levels until the 1960s. Initially, the explanation followed a classic modernization discourse, in which fertility was held in check by late marriage, as household formation was contingent on inheriting a farm or workshop. The spread of wage labor supposedly led to an 'intermediate' period of earlier marriages and high fertility, until people adopted modern birth control techniques to limit family size (for an overview of the debate, see Engelen (2009)). The willingness to do so presupposed an attitude of self-determination, which was part of modern culture, which supposedly first emerged in the North and West, and gradually spread across the country. Critics of this view claimed that the regional pattern in Dutch fertility decline could be explained more convincingly by the predominance of Roman Catholics in the South, in combination with the Catholic opposition to birth control. Throughout the 1960s to 1980s, scholars have tried to solve the debate by weighing the relative contributions to fertility decline of social and economic 'modernization' on the one hand and religion on the other.

Internationally, the debate has been between those who see fertility decline as an 'innovation' of new ideas and techniques, which means research had to focus on the venues and mechanisms of diffusion, such as social learning. On the other side, scholars still see merits in the older conception of fertility decline as a response to decline in mortality.

Ethnographers have for a long time documented the contraceptive methods employed in non-Western societies. Recently, historical demographers have demonstrated that conscious birth spacing, for instance due to expected increases in food prices, was practiced by pre-transition European populations (e.g. Van Bavel, 2001; Dribe & Scalone, 2010). Birth spacing intensified during the first stage of the fertility decline, for example through protracted breastfeeding. Even in the 20th century, traditional contraceptive methods such as abstention and withdrawal remained the most important techniques for married couples. The HSN has contributed to this field by providing family reconstructions which were not, as usual, limited to specific regions or social groups. Moreover, the database makes it possible to weigh the relative impact of socio-economic and cultural factors (especially religion) at the level of individual couples. And not only parity-specific stopping behavior was studied, but also spacing and childlessness. Probably the most important contribution of register data to the study of fertility is the possibility to study the conditions triggering agency or the willingness and ability to limit one’s fertility.

Spacing studies using the HSN have revealed that (working class) couples burdened with many young children delayed the arrival of a next child (Van Bavel & Kok, 2004, 2010c). However, a deliberate response to rising food prices was not found (Van Bavel & Kok, 2005b). Time and again, religion proved a prime factor in determining whether people exercised agency to control births or not (Schellekens & van Poppel, 2006). Especially Roman Catholics stood out with short birth intervals, which may be related to a lower incidence of breastfeeding, the 'conjugal duty' (spouses were not supposed to refuse to have sex) or active pronatalism (Van Bavel & Kok, 2005b). Recently, van Poppel, Reher, Sanz-Gimeno, Sanchez Dominguez, and Beekink (2012) demonstrated that couples actively responded to
deaths of their children, in other words spacing and stopping was contingent on childhood survival (also Reher, Sandström, Sanz-Gimeno, & van Poppel, 2017; Shepherd, Kok, & Hsieh, 2006). Thus, they step in on the 'adaptation' side of the fertility debate. They observed that "childhood survival had clear effects on reproduction, the chances of having another child, and the length of the intervals between births, which indicates that this variable was crucial for fertility decisions" (p. 300). The pattern was especially strong after 1900. This form of adaptive agency was not found in all groups. Skilled workers and Liberal Protestants show the reaction much stronger than farmers, Roman Catholics and Orthodox Protestants. A similar conclusion was reached by Schellekens and van Poppel (2012) who attribute the decline in fertility in the Netherlands before 1940 to the decline in child mortality, rising real wages and unemployment during the Depression. These see no room for innovation/diffusion models: "ideational change that is independent of social and economic change was not a major determinant of the decline" (p. 982).

The Interwar period has also been subject of research into the remarkable rise in the number of couples who had no, or only one, child. With hindsight, low fertility in this period has been interpreted as the effects of unemployment and looming war, but contemporary commentators were blaming the remarkable increase of childless or one-child families on changing life styles. Supposedly, many couples were hesitant to give up newly-won luxuries and holidays, and supposedly decided to have no or few children. To check whether childlessness was a reaction to adverse circumstances or a reflection of a 'modern lifestyle', Van Bavel and Kok (2010a) combined the reproductive histories of almost 3,000 HSN couples with information on their municipalities (e.g. level of unemployment, number of stores). Although, indeed, married men without occupation often remained childless, they also found indications that more or less deliberate childlessness was involved in order to defend a career or a modern lifestyle. Controlling for age at marriage, white-collar workers had relatively high hazards of childlessness, although the difference with other groups of workers gets small and statistically insignificant after controlling for regional and municipal characteristics. As to the municipal effects, people in localities with many stores tended to have no children, suggestive of a 'luxury effect'. The same applied to people in urban environments. Furthermore, couples who had a mixed marriage or who were both without a religion had relatively high odds of childlessness (also Van Bavel & Kok, 2010b and Van Bavel, Kok, & Engelen, 2008). The 20th century also witnessed the emergence of the ideal of a mixed offspring set, consisting of a boy and a girl. Kok (2018) showed that this emerged first among Liberal Protestants.

The HSN has also been used to demonstrate that there is still room for a 'blended' approach to fertility decline in which diffusion of new ideas and attitudes plays a role alongside structural factors. Bras (2014a) studies social networks of HSN RPs by looking at the witnesses at the marriage ceremony. Age, occupation and relationship to bride or groom of these (four) witnesses were recorded. Bras discovered that when witnesses consisted of siblings, age-peers and/or women, the couples in question were more likely to practice birth control. She sees these reference groups as "salient role models for social learning" (p. 178). These ‘controlling’ couples stand in contrast to couples who had more ‘vertical’ ties, that is witnesses from their parent’s generation — in particular unskilled workers and farm workers. Using an oversample of four industrial cities, Janssens (2009; 2014) studied the impact of female working experience before marriage on their reproduction. She assumed that women coming into contact during their work with upper classes (such as domestic servants), or women working alongside other women (factory girls) would have had more opportunity for social learning than, for instance, women working at home. She showed that (former) servants were actually late in adopting birth control. Interestingly, a high education before marriage did not have the expected effect of lower fertility (through an increased bargaining position with the husband) either.

Most evidence from the HSN points into the direction of fertility decline being an adaptation to changing circumstances, in particular declining child mortality. But there are also indications that some ideational change was going on, diffused through networks of peers. However, the old Dutch debate on the striking regional variation in fertility decline in the Netherlands is still far from solved.

Apart from explaining socio-cultural variation in fertility and its decline, innovative research using longitudinal micro-data has come from molecular epidemiologists and researchers testing hypotheses from evolutionary biology. A large oversample of Rotterdam women, who were traced from birth to their own families has been employed in epidemiological studies to discover negative effects from the season of birth (Smits, Jongbloet, & Zielhuis, 2001; Smits, Zielhuis, Jongbloet, & Straatman, 1998); the age of RPs mother (Smits, Zielhuis, Jongbloet, & van Poppel, 2002) and the interval after which
the RP was born (Smits, Jongbloet, & Zielhuis, 2000) on her own fecundity (also Smits, 1998). With respect to the preceding birth interval, it was discovered that RPs born after very short birth intervals (less than 1 year) showed higher likelihood of childlessness and stillbirth in their offspring, compared to those born after intermediate intervals (21–32 months). The age of the mother was important as well: RPs whose mother was over forty had higher risks of stillbirths, childlessness and multiple births than RPs born from a mother of intermediate age (24–30). The explanation lies in ovarian maldevelopment, which already occurs in fetal stage. Recently, the LINKS data were used together with other databases on ‘natural fertility’ populations to chart the distribution of the final age at childbirth (Eijkemans, van Poppel, Habbema, Smith, Leridon, & te Velde, 2014).

In recent studies, the relevance of evolved human traits for understanding (historical) demographic behavior has been tested. One example is the notion that the human race stands out compared with other animals by the role kin groups play in rearing children. It has been suggested that through evolution, women have developed a relatively early menopause in order to help their daughters with raising children. The notion of co-operative breeding or the ‘grandmother hypothesis’ has been tested using data from several villages, studies in developing countries and historical family reconstitutions (Sear & Coall, 2011). Overall, maternal grandmothers appear the most reliable ‘helpers in the nest’. This hypothesis was tested using household information contained in the HSN by Roterink and Bras (2015). They studied the effect of living-in kin on birth intervals, and expected that ‘helpers in the nest’ would lead to shorter intervals. Some kin effects were found, but only in the low parities. Also, a ‘grandmother effect’ was not found. Living-in widowed grandfathers even had the opposite effect of delaying the arrival of a next child, probably because he was using up too much of the couple’s resources. As to siblings it was not, as could be expected, sisters who would stimulate fertility, but brothers — apparently because they brought in additional income.

4 NEW RESEARCH FIELDS

In the wake of poststructuralism and responding to renewed interest in individual agency, the early 1990s saw the emergence of new topics, such as individual life plans, individual and family strategies of betterment, etc. Researchers began to use the HSN to study how people responded (e.g. by the timing and reason for leaving home, and the timing of marriage) to parental needs in relation to their own preferences. Apart from parental social class, family size and sibling order were seen as key indicators to study to what extent demographic choices were determined by parental resources and family constraints. HSN researchers could also join emerging new fields of research, such as the role of siblings and the wider kin network in influences on life courses, the topic of ‘death clustering’ at the level of families, and the study on how early life conditions affected health in later life.

The HSN’s rather unique research design — following individuals through the entire country, and even beyond — also made it the perfect database for in-depth studies of migration. Until the early 1990s, research into migration was always hampered by skewed comparisons (leavers contrasted to stayers, in-migrants contrasted to natives), and the fact that generally only one move per individual was observed, as research was always limited to one locality. Not surprisingly, the HSN stimulated diverse studies of migratory behavior.

4.1 MIGRATION

Being a servant as part of the life cycle, has always been an important element of the North-Western European family system, and as such, a key factor in historical demographic interpretations of late marriage (Hajnal, 1982). However, the ‘life cycle servants’ themselves have hardly been subjects of research: from what families did they come, at what age and why did they leave home, how long did they stay away, did they remit their earnings to their family, did their experiences away from home — e.g. living in urban, upper-class households — affect their choices later in life? Using the data on the first completed province (Utrecht), Kok (1997) charted the family background of leaving home, and traced the domestic servants and farm workers in their subsequent jobs (also Bras & Kok, 2003). Recently, the various destinations (ranging from an interlocal move to emigration) of rural youths leaving home were studied using a competing risk event history analysis (Mönkediek, Kok, & Mandemakers, 2015; Mandemakers, Mönkedie, & Kok, 2016).
Bras focused on girls from the province of Zeeland, and included sisters in her research project. Her aim was to find out whether and how life chances of women were affected by (possible) service, the location and employers, and the migration trajectory, controlling for parental environment. One of the hypotheses was that domestic service was a ‘bridging occupation’, allowing girls from rural lower-class families to achieve upward mobility because of their increased social and cultural capital gained in cities. However, service mobility in Zeeland remain predominantly local, with only a minority of the girls actually working in cities. Those who did were indeed able to marry a higher status husband, but local city girls fared even better (Bras, 1998; also Bras, 2002). Girls moving to a city were already a selected group, being able to draw on parental resources. The girls came from families of teachers, supervisors, lower civil servants, and skilled laborers. These families could afford losing direct income by having a family member working in the urban sector and could possibly spread their risks in times of economic hardship. They may also have seen a position in urban domestic service as an opportunity for education and upward mobility for their daughters. Siblings already living outside Zeeland enabled a move out of the province (Bras, 2003).

Sibling effects on the likelihood and direction of migration was also the topic of a comparative study contrasting Zeeland to the Pays d’Herve (Eastern Belgium). Bras and Neven (2007) demonstrated that the type of labor market, which differed strongly between these regions, affected whether women would be stimulated by their brothers or sisters to migrate. In Zeeland, girls were affected primarily by their sisters, because of the gender division in the labor market: girls were sent into domestic service, boys to work in agriculture. The marriage of a sister implied that another girl was sent into service, to make up for the loss of income. In Eastern Belgium, service was not common. Here, migration of females was influenced primarily by the size of the sibship, by the presence of younger siblings at age 12, and by the births and deaths of new siblings, in other words the parents aimed to regulate the size of the household. The effects of siblings were less sex-specific in the Belgian area. Since the women migrated into urban employment opportunities in Walloon industry, the contacts and resources of both their sisters and their brothers were important (Bras & Neven, 2007). Taking another approach to siblings’ migrations, Kok and Bras (2008) used the Akersloot oversample (see above) to study how families dispersed. By looking at the residences of parents and their adult children in different stages they charted dynamic ‘family territories’. Marriage was the prime factor in explaining dispersal of siblings.

The study of internal migration also included family migration, and the HSN register data allowed scholars to add previous migration experiences and the (dynamic) composition of the household to their models (Kok, 2004; Kok, Mandemakers, & Mönkediek, 2014). Detailed migration information spanning the life course is hard to come by, and the findings from HSN were also used to ‘calibrate’ migration studies based on genealogical information. As genealogies may be biased by the social class of the descendants reconstructing their family trees, and by the fact that they depend on vital events (marriage and child births), a check with register data can be useful (Kok, Adams, Ericsson, & Moch, 2002; Kok, Lucassen, Kasakoff, & Schwartz, 2002). In particular, a comparison was made with American genealogies, by simulating the Dutch data in such a way that it became clear how many moves were missed by reliance on vital events, and in what stage of the life course (Adams, Kasakoff, & Kok, 2002). Recently, Jennings and Gray (2014) related migration to weather conditions (time series of precipitation and temperature), thus relating to current concerns of the impact of climate change on migration. They showed that only internal mobility of specific social groups could be labelled ‘distress migration’, whereas emigration was actually limited by adverse conditions. Over time, people were less ‘trapped’ by bad weather conditions and moved more freely (Jennings & Gray, 2014). Finally, HSN also allows the study of residential mobility, a topic much neglected in historical demography. Kok, Mandemakers, and Wals (2005) analyzed the address changes of Amsterdam dockworkers, people on poor relief, and the HSN sample of Amsterdam. They interpreted residential mobility as a ‘coping strategy’, when cheap dwellings were available, poor people moved frequently to save on the rent, which was only an option when the move was over a very short distance and did not entail a break with their social network on which the poor relied for survival. Van der Harst (2006) traced the moves and residential clustering of Zeeland and Brabant migrants in the city of Rotterdam.

By definition, foreigners are absent in the HSN sample, since it is based on the birth records. Life courses, however, are ideally suited to study integration and assimilation through, for example, residential clustering, chances of upward mobility and intermarriage. Following HSN procedures, cohorts of migrants have been followed through the registers in Rotterdam and Utrecht. The focus in these projects was on more permanent migration, whereas many immigrants soon moved on or returned,
probably because the city offered them no prospects. Perhaps the migration literature has paid too much attention to this ‘floating proletariat’ of migrants who left quickly (Thernstrom, 1973). Germans who stayed in Rotterdam actually fared very well. They were positively selected to begin with, but they managed to reach higher positions than the native born (Lucassen, 2004; also Delger, 2003; Schrover, 2002; van de Laar, Lucassen, & Mandemakers, 2006). Many Germans came to the Netherlands to occupy specific occupational niches (e.g. file-makers or stucco-workers), and their migrations are clear examples of chain migration, linking a few places in the German hinterland to Dutch cities. The study of German communities offered interesting insights in how these niches developed and how they remained dependent on incoming migrants from the home community, preferably both men and women (Schrover, 2000a, 2000b, 2001, 2003). However, the study of German immigrants has also revealed that many of them were not connected to other Germans, and apparently managed perfectly well to find housing and work on their own. Lesger, Lucassen, and Schrover (2002) conclude that ethnic ‘chain migration’ has received too much attention in the migration literature — in which American scholars predominated — probably because ‘community studies’ in the US were more attractive than studies of loosely attached migrants.

The study of (r)emigration on the basis of HSN is still limited. Following migrants into Germany is hardly possible, because of the lack of register data. Migrants crossing the border into Belgium have been followed, but they have not been studied separately. An important exception is a study into the careers of Dutch (HSN) men who joined the colonial army in the Dutch East Indies. The commonly held opinion is that such colonial armies consisted of the dregs of the society. It turned out that many of these migrants had an urban background, in contrast to the migration of the Dutch to the United States, which had a strong rural basis. However, the milieu of the unskilled urban proletarian was under-represented among the migrants to the Dutch East Indies. Apparently, engaging for colonial military service, even as an ordinary soldier, was considered to be a serious option for sons from the lower middle classes and artisans of the Dutch cities. The higher classes were actually over-represented among those who left for the Dutch East Indies (Bosma & Mandemakers, 2008; Bosma, 2009, 2010). Recently, an attempt has been made to trace HSN emigrants in the US censuses (Paiva & Anguita, 2017).

### 4.2 DEATH CLUSTERING

Increasingly, scholars have become aware that early-life mortality was often clustered in families. In other words, a relatively large proportion of all deaths occurred in a relatively small number of families. To properly understand the trends and local variation in infant mortality we need to know more about this phenomenon. Factors that may account for the clustering include the mother (her age, fecundity, inherited genetic disorders, disease profile, level of education, training in hygiene and breastfeeding habits), the household (income and composition) and the community (ecology and disease environment) (Edvinsson & Janssens, 2012). For the Netherlands, a study based on the LINKS dataset has shown that infant death clustering was unevenly spread, being very strong in the province of Zeeland. The study revealed that part of the explanation lies in the competition between young children for care, attention and resources within households. The mother’s health was also important, which was indicated by a strong association between stillbirths and infant mortality; if a mother has poor health, she has more births with low gestational ages and low birth weights (van Poppel, Bijwaard, Ekamper, & Mandemakers, 2012). Using registers data, Janssens and Pelzer (2011, 2012) were able to attribute infant deaths in Enschede and Tilburg completely to the ‘usual suspects’ such as twin births, the preceding birth interval and the age of the mother. Therefore, there appears to be limited explanatory space for bad childrearing practices, at least in these two cities.

The recent dissertations of van Dijk (2019), Mourits (2019) and van den Berg (2020) all dealt with various aspects of death clustering. Van Dijk used extended family reconstitutions of Zeeland (based on LINKS) to trace the importance as well as the intergenerational transmission of familial death clustering (also van Dijk, Janssens, & Smith, 2018; van Dijk & Mandemakers, 2018). She found that children growing up in high-mortality families suffered lasting consequences in terms of lower survival chances of themselves and of their children. The mother’s experience of sibling mortality turned out to be more important than the father’s. Mourits (2019) focused on mortality after age 50 and individual chances to belong to the top 10% survivors of their birth cohort. He demonstrated that in longevity as well there was a clear element of familial clustering. This implies that ‘average’ life tables can be quite meaningless, as the distribution of survival chances are strongly determined by family effects.
In understanding longevity clustering, Mourits extensively explored social (resource transmission) and geographical components (see also Mourits, 2017). Finally, van den Berg (2019) used LINKS and HSN to develop new concepts and methods to study the familial component in longevity (also van den Berg et al., 2019a, 2019b). For his research, a case and control group design based on HSN was followed which implied tracing all descendants of exceptionally old RPs as well as a control group from late 19th century to the present day. This unique dataset contains five generations connecting living persons to their deceased ancestors. Van den Berg unequivocally showed that longevity is only transmitted as a quantitative genetic trait if persons belong to the oldest 10% survivors of their birth cohort when at least 30% of their ancestors also belonged to the oldest 10% survivors (van den Berg et al., 2019a, 2019b). A comparative study, covering five databases, including the LINKS database for the province of Zeeland, showed a strong relationship between levels of infant mortality of grandmothers and their daughters (Quaranta & Sommerseth, 2018).

4.3 EARLY AND MID-LIFE EFFECTS ON HEALTH AND MORTALITY

Since the early 1990s, the notion that later life diseases already have their origins in the fetal stage (e.g. Barker, 1992) has stimulated a lot of research. A number of important contributions to this field were based on HSN register data. Early life conditions were studied for birth cohorts, which were exposed to economic crises or epidemic diseases. For the survivors, having experienced these adverse conditions could still imply a deterioration of health in later life. Van den Berg, Lindeboom, and Lopez (2006) looked at the effects of the business cycle in the years HSN RPs were born. They discovered that an individual born during a recession lived a few years less than an individual born in a boom. The outcomes suggest that the very first year of life is already crucial. As for the causal pathway, they suggest that the cyclical turbulence creates stress among parents, which in turn creates high mortality later in life among their infants. In a follow-up study, Yeung, van den Berg, Lindeboom, and Portrait (2014) merged the HSN life course data (cohorts 1880–1918) with cause-of-death information held at the Statistics Netherlands (CBS). This showed, surprisingly, that adverse condition in early life led to higher mortality due to cancer, as well to chronic respiratory diseases (females). Overall, the effects of economic crises in childhood were stronger for women than for men. The effect of nutritional defects on fetal growth — leading to impaired health in later life — was tested using an HSN cohort that had experienced the potato blight of the 1840s. Van den Berg, Lindeboom, and Portrait (2013) discovered that men exposed to severe famine during pregnancy (at least four months) and directly after birth have a significant lower residual life expectancy at age 50 than others, but not at earlier ages. The effect was not found for women, or for men experiencing the crisis when they were slightly older (also Lindeboom, Portrait, & van den Berg, 2010; Popławska, 2015). Schellekens and van Poppel (2016) use aggregate mortality rates and height as indicators of early life conditions and estimate that improvements have contributed more than five years, or about a third, to the rise in life expectancy of females at age 30 between cohorts born in the periods 1812–29 and 1910–21. And such improvements contributed almost three years, or more than a quarter, to the rise in life expectancy of males at age thirty.

Research has also focused on ‘intermediate’ events in the life course which could exacerbate or weaken effects from early life (on the role of marriage, see van den Berg and Gupta (2015); also Gupta (2010)). Effects from mid-life experiences on health in later life were also studied. Kaptijn et al. (2015) test the hypothesis of a trade-off between women’s fertility and their longevity. Fertile women may have shorter lives, e.g. through maternal depletion. Such a trade-off was found, but only for the oldest cohorts. Women with few, or with many children, lived shorter. Possibly, improved wealth and health conditions explain the disappearance of the trade-off in the younger cohorts. Stressful periods can also lead to impaired health. Alter, Dribe, and van Poppel (2007) showed that widows, especially those with large families, lived shorter than other women. Migration might also be such a stressful experience. However, Puschmann (2015), Puschmann, Donrovich, Grönbek, Dekeyser, and Matthijs (2016) and Puschmann, Donrovich, and Matthijs (2017), using among others HSN data on Rotterdam, demonstrated the existence of the ‘healthy migrant’ effect; migrants were already positively selected and actually had lower mortality risks than the natives. This proved especially true for Zeeland and Brabant migrants to Rotterdam. However, they also detected specific vulnerable groups with higher mortality risks, such as Italian migrants.
4.4 FAMILY AND SURVIVAL

In line with the already discussed interest in ‘cooperative breeding’ and related hypotheses stemming from evolutionary biology, researchers have looked for the effects of living-in kin on child survival. Clearly the presence of parents, especially the mother is of paramount importance (van Poppel & van Gaalen, 2008b). Kok, Vandezande, and Mandemakers (2011) showed that living-in kin could have very beneficial effects. The positive effects of the presence of grandparents could compensate for the loss of the father or mother. Uncles and aunts as well played a positive, altruistic role. Apparently, a family setting which was optimal for infants’ and children’s well-being included parents, grandparents, (unmarried) uncles and aunts, possibly servants, boarders and lodgers. However, the family should not have too many young children, as the infant would have to compete with them (also Vandezande, Kok, & Mandemakers, 2011).

Exploring this ‘competition’ further, Riswick (2018) used the HSN to study how infant and child mortality was affected by the interaction between resource dilution and the specific family type prevailing in different regions in the Netherlands. He shows, firstly, that the number and gender of siblings played an important role in determining child mortality, but were less significant in determining infant mortality. After the age of one, boys experienced more competition from their brothers. Riswick surmises that this effect is related to their assigned roles within the household and on the family farm. A similar effect was found for girls in the nuclear family region in the northwest. Here, fewer children were needed to work inside or outside the household, and girls and boys may even have been interchangeable.

4.5 RELIGION AND THE LIFE COURSE

The Dutch population registers are among the few that record religious denomination. And they did so quite meticulously: more than a hundred different denominations can be found in the Dutch registers. This offers the rare opportunity to study how religion can affect demographic behavior and life course transitions. In principle, it is possible that such effects were caused by specific religious norms and prescriptions, coupled with effective sanction mechanisms. However, McQuillan (2004) has argued that another condition is necessary, which is the identification of people with their clergy. The latter could differ from country to country, and was often due to political reasons. But apart from religious norms, sanctions and identification, belonging to a church community may have an effect in itself, e.g. through the support derived from the community. Observed differentials by religion may also stem from differences in social characteristics of the members. And finally, the observed connections may not derive from individual beliefs and associated behavior, but from an underlying effect of the presence of specific churches on the environment (Kok, 2017). Kok compared the outcomes on the religion variables in regressions models on twenty different life course transitions, ranging from leaving home to old age mortality. Differences in moral prescriptions and sanctions between the Churches, but, perhaps even more important, in ‘mentality’ (e.g. fatalism or openness to the outside world) led to marked variation in behavior, after controlling as much as possible for social characteristics.

One of the HSN subprojects has been an oversample of Amsterdam Jews. This material has been used to study how (during 1880–1940) assimilation into Dutch society at large changed their household structure (Tammes & Van Poppel, 2012). Furthermore, their life courses were studied to understand how and why Jews decided to leave their faith, and to become either unaffiliated or to convert to Christianity. Individual backgrounds were combined to period effects (changes in the level of assimilation of Jews through intermarriage, the rise in atheism, and anti-Semitism), and the outcomes suggest different pathways to either conversion to Christianity or to simply abandon the faith (also Tammes, 2012b; Tammes & Scholten, 2017). Tammes (2011) also used the data to study residential segregation of Jews before the War. The rise of secularization among the HSN population was the subject of a study by Knippenberg and de Vos (2008).

Jews and Roman Catholics had been subject to official discrimination before 1800. Van Poppel, Liefbroer, and Schellekens (2003) showed, on the basis of a specifically constructed HSN subset, that this indeed affected their (inherited) class positions (in the city of The Hague), but also that intergenerational social mobility patterns in the 19th century did not differ across the denominations. In other words, social disparities between religions were ‘path dependent’ and not the result of ‘new’ obstacles.
4.6 SOCIAL NETWORKS

Register data as such give very little clues about the social networks of the RPs. At best, we know who, for a longer or shorter time, lived in one household. With respect to kin networks, this lacuna can be dealt with by combining HSN to the extended kin traced through the LINKS matching system, see also below. But friends, colleagues and acquaintances remain out of view. However, the civil records, which are also part of the HSN database contain witnesses at birth, marriage and death. Especially those present at the marriage are of interest, as the relation to the bride and groom are specified. Although witnesses can be simply municipal clerks, or even witnesses from a just completed wedding asked to stay on for the next (Kok, Adams, Ericsson, & Moch, 2002), generally the data give a good insight in who were considered close and important. Van Poppel and Schoonheim (2005) used the witnesses to study the networks of Jews, in contrast to those of other religions. They observed a much stronger familial involvement with the marriage ceremony among Jews, a phenomenon that was observed in all social classes. Also, the network of Jews included people coming from a much wider area than was the case among Dutch Reformed and Catholics. Again, these differences remained when social class of the groom was controlled (also Tammes, 2012a).

Bras (2011) uses the marriage witnesses in HSN to study a transition in the meaning of kinship. During the 19th and 20th centuries, an intensification of family relations took place (see also Bras, 2014a). During the period 1830–1950, lateral kin (siblings, siblings-in-law and cousins) were increasingly selected as marriage witnesses, at the expense of professional witnesses and patronage relations. The process started in urban higher and middle classes and among farmers, but cultural diffusion took place whereby the choice for brothers(-in-law) and cousins increasingly spread to other social groups. Also, familial age peers were particularly selected when the mother (of the bride) was still alive and when the bride was literate, pointing to the important role of women in cultivating family relations (see also Zijdeman & Corten, 2012).

4.7 BIOLOGICAL STANDARD OF LIVING

Economic history, as well as other disciplines, shows an increasing interest in new indicators to measure well-being, in terms of determinants, long-terms developments, societal variation and inequality. Prominent among these new indicators are human heights, as they reflect (net) nutrition, health and workload (Floud, Fogel, Harris, & Hong, 2011). HSN and the development of the LINKS dataset, in combination with an integrated index to the municipal militia records (thanks to a crowdsourcing project) enabled a unique research design: the linking of complete life courses to early adult heights of RPs as well as of their (male) kin. The project Giants of the Modern World aims to understand what changes in Dutch society (e.g. in family size, access to high quality food, economic growth and integration) were responsible for the unique position of the Dutch as the tallest in the world. The first results confirm that in the second half of the 19th century the coastal provinces and large cities recovered from the economic stagnation that had plagued them since the late 18th century, but national convergence was still not visible by 1940 (Tassenaar, 2019).

The HSN makes it possible to study the effect of (changing) early life conditions on stature, because all changes in household composition are known from birth of the RP onwards. Quanjer and Kok (2019) experimented with a consumer/producer ratio and showed that it captures resource dilution effects in households more effective than e.g. sibship size. They also demonstrated that children suffered more from the death of the mother than from the death of the father (see also 3.3.2.) and that boys had an advantage over girls when it came to food distribution in households (at least they grew taller when they had female siblings compared to male siblings). Finally, they conclude that the early introduction in the Netherlands of the homemaker-breadwinner model might be an important clue to understanding Dutch gigantism.

Stature not only reflects adverse or beneficial early life conditions, it can also have an independent effect on one’s life chances, e.g. by its relation to earning capacity. To unravel direct and indirect effects on later life outcomes, one needs ‘mediation’ models. Thompson, Lindeboom, and Portrait (2019) used such models to relate economic crises in childhood (the Potato Famine of 1817) to socioeconomic status in early adulthood, mediated by height. So far, no mediation effect was found.
5 META RESEARCH ON DATABASES, DATA STRUCTURES AND DATA

5.1 DATA STRUCTURE

From the start the HSN worked with custom made data entry programs. When the database became larger, we gained a better understanding of how to organize the data. Although numerous studies existed detailing the process of historical data entry and the difficulties of data interpretation, how the database itself should be organized remained quite implicit. In May 2001, the HSN organized a workshop in which all major database administrators were invited to comment on the sketches and ideas that existed within the HSN. This workshop was very well attended and resulted in a comprehensive list of best practices for the creation of large databases on historical populations (Mandemakers & Dillon, 2004). Most important in these guidelines was the systematic distinction between the data entry part of the database, the database itself with unstandardized, standardized, enriched and linked data into life courses and the part with all the different releases that have been realized during the history of the database.

Building life courses from dynamic data such as data from population registers is not an easy and straightforward job. The way the HSN constructed the life courses out of the raw data was first systematically described by Mandemakers (2006b). During this process several decisions about the structure of the dataset had to be taken that were not always in the interest of all researchers. It turned out that this was not a typical HSN problem but a more general one for all historical databases. And there was also the question of how to enlarge the potential use of all these databases. Several workshops were organized to tackle this problem, a first one in Montréal (2003), a second one in Amsterdam (2006) and a third one in Ann Arbor (2008). Although all participants in the meetings recognized the complex nature of their longitudinal databases, it was the general feeling that comparative research was possible if certain conditions were set. It was agreed, firstly, that standardization would be the basis for comparative research. And secondly, that a standard structure was necessary to create possibilities of data conversion from different datasets into common datasets for analysis. This resulted in a model for data sharing, which became known as the Intermediate Data Structure (IDS; Alter & Mandemakers, 2014; Alter, Mandemakers, & Guttmann, 2009). Several databases have been converted into the IDS-structure. Major developments are the integration of four Swedish databases into one IDS data structure (Swedpop project, https://www.vr.se/english/mandates/research-infrastructure/find-research-infrastructure/list/2018-10-18-swedpop---swedish-population-databases-for-research.html). The HSN is working to integrate the HSN register data with the civil certificates and the LINKS dataset. LINKS will provide more data on already included relatives and will also extend the number of relatives as far as third and fourth degrees. So-called extraction software to convert IDS-data into datasets for analysis have been developed by Luciana Quaranta (2015, 2016, 2018).

Within the context of the HSN the existing knowledge was also used to develop data structures of data entry or integration for data from former colonies of the Netherlands. Ideas to develop HSN like databases for Suriname and Dutch East Indies, especially Java, were developed and tested (Bosschaart & Lelieveld, 2016; van Bruggen, Bulten, & Maas, 2017). Also, a data structure was developed for the so-called thombos, a quite complicated population and land register of Dutch colonial Ceylon from the second part of the 18th century (van den Belt, Kok, & Mandemakers, 2011). The inaugural lecture of Mandemakers (2009) stressed the importance of standard formats and the cooperation of specific scientists from different academic fields, statisticians and methodologists to do cutting edge research.

The first initiative to develop the Intermediate Data Structure was supported by the Dutch Research Council (NWO; the Humanities program for internationalization). This initiative was strengthened by the so-called European Historical Population Samples Network funded by the European Science Foundation (ESF). This program ran from July 2011 to December 2016 and brought together in one program databases from about twenty European countries, whereas databases from outside the EU participated as much as possible. The program was chaired by the HSN and concentrated on several tasks: documenting of existing databases, promoting of the IDS and developing software that ran on the basis of the IDS structure and organizing summer schools. All documentation was concentrated on the website, including the launch of a journal to publish methodological and substantive research articles, especially if the used datasets were based on the IDS, see further https://ehps-net.eu/. A follow up of the program was Methodologies and Data mining techniques for the analysis of Big Data based on Longitudinal
Population and Epidemiological Registers (LONGPOP). It is a project within the framework of the Marie Skłodowska-Curie Innovative Training Network within the Horizon 2020 Programme of the European Commission and is led by Diego Ramiro (http://longpop-itn.eu/).

A new development in data formats is the introduction of so-called linked open data structures (LOD). By way of the semantic web it becomes possible to connect data from databases all over the world and to create ‘new’ datasets (Meroño-Peñuela et al., 2015). Initiated by DANS, the HSN and Free University Amsterdam cooperated in the CEDAR project to develop a LOD structure for the aggregate data of the Dutch censuses from 1830 till 1948. This resulted in a methodological description of workflows to convert source data into the new data structures in a systematic and repeatable way (Ashkpour, 2019; Ashkpour, Mandemakers, & Boonstra, 2016; Ashkpour, Meroño-Peñuela, & Mandemakers, 2015).

5.2 STANDARDIZING AND CODING DATA

HSN occupational data as such were intensively used during the construction of the Historical International Classification of Occupations, the so-called HISCO code (van Leeuwen, Maas, & Miles, 2002, 2004). This was not that remarkable since three researchers from the HSN steering committee, Marco van Leeuwen, Ineke Maas and Onno Boonstra, were strongly involved in this enterprise (Maas, 1998). The introduction of the HISCO code, by Theo Engelen and Hans Hillebrand, lamented the situation that in the Netherlands ‘each researcher constructed his/her own classification scheme’ (Engelen & Hillebrand, 1985, p. 255).

The HISCO code itself is not suitable for doing analytic research, but it forms the basis for stratification schemes according to social class or social status. Using characteristics of the occupational titles according to distinctions like manual or non-manual character, higher, medium, lower or no schooling, supervising or not and being part of the primary sector, Maas and van Leeuwen constructed a decision scheme based on HISCO resulting in twelve social classes (HISCLASS). Van de Putte and Miles (2005) constructed SOCPO which is a scheme of five status groups based on the principal of social power originating from both economic and cultural resources. Economic power is derived from factors such as self-employment, skill, and authority (command). Sources of cultural power are non-manual versus manual occupations and nobility and prestige titles. An interval scale called HISCAM created a scale from 0 to 100, ranking the HISCO codes on the basis of the occupations of the participants in marriage certificates (Lambert, Zijdeman, van Leeuwen, Maas, & Prandy, 2013; Zijdeman & Lambert, 2010).

During the last twenty years HISCO has become the basis for the classification and ranking of occupational titles for almost all research with HSN data. The HSN has also taken the lead in standardizing and classifying occupational titles into HISCO, HISCLASS, HISCAM and other systems. The latest release (Mandemakers et al., 2018) includes 134,964 different occupational titles found in the sources used by the HSN and LINKS datasets.

Besides the coding of occupational titles, standardization schemes have been developed for all major data types that are included in the HSN and LINKS datasets such as locations, religion, age, civil status, types of certificates, sex, first names and last names. Most important is the location file with standardized and spatio-temporal coded values of 7,925 location names (hamlets, villages, towns, cities and municipalities). This dataset was released by Huijsmans (2013) including the municipality to which they belong, X and Y coordinates as well as longitude/latitude coordinates within the framework of the LINKS project. And of course, the release also includes the AMCO code developed by van der Meer and Boonstra (2006) which is the way to link with all other datasets in the field of locations.

5.3 RECORD LINKAGE

In essence the LINKS project is no more than software for record linkage and standardization in which the indices of names from the Dutch civil registration are used to reconstruct pedigrees and complete families for the Netherlands in the 19th and first half of the 20th century. No wonder that a lot of effort went into improving record linkage procedures. The first record linkage on marriage certificates was done by Oosten (2008) who developed artificial intelligence to match equal but apparently different names such as Guillaume and William. The basic idea of his algorithm was that two of the same names will never appear in first names that are a combination of different first names (like ‘Cornelis Albertus’). The idea was that if a positive match was only dependent on two completely different first names...
and these two names did not appear together in combination of first names than the names probably would be synonyms of each other (if not they should appear together in combined first names).

In his thesis research Schraagen (2014) tested old and developed new techniques on record linkage with the context of the LINKS project. Together with Bloothooft he further developed the basic ideas of Oosten by learning name variants from matches that on the one hand have a high level of precision, but on the other hand lacked one element necessary to approve for a match (Bloothooft & Schraagen, 2015). Schraagen experimented with different and innovative record linkage techniques, such as using domain based information on the structure of families resulting in more valid evidence for proposed links (Schraagen & Kosters, 2014), using graph theory to predict if a link is missing because of failed record linkage or because of missing records (Schraagen & Hoogeboom, 2011) and developing methods to compare datasets which are structured differently for events (from certificates) or for individuals (from genealogies; Schraagen & Huijsmans, 2013). Within the LINKS project the emphasis gradually changed from record linkage as a technique to the reconstructions of populations as a result. In 2014 LINKS organized a conference with the theme ‘Population Reconstruction’. Several topics were presented and discussed on themes as standardization of names and locations, record linkage and population reconstructions (Bloothooft, Christen, Mandemakers, & Schraagen, 2015).

5.4 STUDIES ABOUT DATA AND DATA QUALITY

Throughout the years, researchers within the HSN domain learned to be critical of the historical data they were using. Using marriage certificates for social mobility, for example, will be associated with various problems. Firstly, the occupational titles of the bride and especially the mothers are underreported, secondly no occupational titles are included when a parent is not alive anymore and thirdly the occupational titles of father and son are only associable at different age levels. Delger and Kok (1998) researched this source of bias by comparing mobility tables for the same HSN RPs based on marriage certificates and based on data from the population registers. The last ones include more fathers with occupations and are also better comparable because the occupations can be compared at the same age levels. The comparison made clear that mobility is underestimated when only the marriage certificates are used as data source. The sons whose fathers had already deceased when they married were characterized by relatively more downward mobility. Walhout and van Poppel (2003) systematically researched when the brides presented an occupational title in the marriage certificate or not. And the sampled character of the birth certificates of the HSN made it possible to do research on the nature of the witnesses in the certificates. Were they family, or so-called professional witnesses hanging around municipality halls having not much to do (Mandemakers, 2017)?

From the onomastic perspective names were not only used as a way to link certificates, but as such they were studied as well. Bloothooft, Mandemakers, Brouwer, and Brouwer (2012) overviewed the way family names and first names are available for research from 1811 till the present. Doreen Gerritzen used the first names of the HSN basic sample to make popularity polls of first names during the 19th and 20th centuries for several provinces (Gerritzen, 1998, 2001). For the introduction, development and socio-cultural imbedding of naming by way of combined first names was researched from 1760 onwards till the present (Bloothooft & Onland, 2016).

Register data, although superior to many other historical sources on populations, obviously come with biases and lacunae of their own (Kok, 2006b). Biases may relate to the observations themselves, to the used administrative categories, to the information provided, and to the fixation on addresses of the population registers. Given the richness of Dutch administrative sources some of the mentioned problems could be solved.

As to observations: the thorough system of civil registration, introduced in the wake of the French occupation, does not leave much room for misidentification or ‘disappearance’ of RPs. A small number of RPs could not be traced to the municipalities where they lived, because their mother was either single (which frequently implied she was sent to an institution or hospital to give birth), a vagrant, or a skipper’s wife. In those cases, the child was not born in the place where the mother officially resided. More serious are (temporary) gaps in the observations across the life course, when persons failed to officially declare their move to another locality, or simply because registers did not survive the Second World War or flooding. The system of Personal Cards, which started in 1940, allowed HSN staff to pick up those ‘disappeared’ persons in 1940, and trace their moves backward in time. A systematic study of these so-called cold cases showed that about a third was traceable, but the rest was not mostly because of the destruction of population registers by war actions or flooding (Haarman, 2018). However, in
many life courses of RPs, shorter or longer periods are missing. Also, as has been said before, the sample only covers Dutch natives, not immigrants. Finally, RPs in the youngest cohort may still be alive today. This implied that no Personal Cards were available, and the end of observation, at best, had to be put at 1940. In particular early studies on mortality based on the first HSN releases were beset with this problem. Van Poppel and van Gaalen (2008a) surmise that mortality in the youngest cohorts was overestimated because of this problem.

Interpreting life courses and demographic behavior with HSN implies having to rely to a large extent on the administrative categories devised in the 19th century. Although most of them are rather straightforward, we do not always know how far they reflect historical reality. An administrative head of the household may not have been a person of real standing at the time. Moreover, as changes in headship are not always recorded, HSN assigns headship status according to an algorithm which may not always be accurate (Mandemakers, 2006b). Sometimes, administrative procedures were ideologically driven — as in the case of Rotterdam after about 1910. Here, the population registers were restructured to fit the ideal of the nuclear family. Not only were living-in kin assigned to cards of their own, single mothers and their children were also separated administratively. Thus, we need to look through ‘administrative’ lenses to critically assess the value of all information from the registers. For instance, official moves to another address were recorded, but lengthy stays from home (sailors, seasonal workers) remain invisible. Finally, and obviously, the official religious denomination is not a proper indicator for one’s religiosity.

The quality of information in the population registers may vary. Firstly, in contrast to civil records, the information is given by informants without the official approval of witnesses. Thus, especially in larger places where they were unknown to the officials, persons may have given false information. Examples are cohabiting couples who reported the female partner as ‘housekeeper’, or cases of single motherhood where the child is ascribed to the grandmother. Secondly, individual changes in occupation and religious affiliation were generally only recorded at the time of the decennial update, or when people migrated. Thus, to some extent the socio-economic profile of an RP depends on geographic mobility. Information in the registers on heads of household may be of better quality than of his or her dependents as occupations of women and children were often not recorded. Therefore, the official occupation of the head may give insufficient information on income and status of the entire household. Thirdly, the SES information contained in the registers is limited to occupations, which, although providing important information, sometimes is too broad (‘merchant’, ‘farmer’, ‘worker’) to indicate the income level and status of the family. Kok, Mandemakers, and Damsma (2010) showed that adding income level for the farmers led to a much more nuanced picture: their income level determined strongly the timing and incidence of the marriage of their children. Also, persons without occupations may be unemployed, retired, or actually living handsomely on non-labor income without us knowing. Finally, missing information may actually lead to biased conclusions. For instance, the strong association Van Bavel and Kok (2010a) found between religiously mixed marriages and childlessness may have resulted from an (unobserved) effect of education, if mixed couples had received relatively high education.

As to the fixation on addresses: in the research design of HSN only the Research Persons are traced in the population registers. This implies that other individuals and kin are out of sight when they don’t live at the same address as the RP anymore (Mandemakers, 2006b). For instance, very close kin may actually live next door (or even in the same house when the address was split), without the researcher being aware. The administration also ‘fixes’ persons to a specific household. The current situation of children from broken families spending time in different households may have occurred — in other forms — in the past as well (van Poppel, Schenk, & van Gaalen, 2013).

6 PROSPECTS

Recently, many historical demographers have joined forces and put forward their ideas for the future of the field (Matthijs, Hin, Kok, & Matsuo, 2016). How does the HSN fit in with these visions? Several scholars have argued for an extension backwards in time and in place, covering non-European areas. In principle, the datasets of the HSN and LINKS could go further backwards into the 18th and 17th century by sampling and linking parish records, but this is not anticipated in the near future. As to
space, a pilot study is undertaken to create a Historical Sample of the Dutch East Indies (van Bruggen, Bulten, & Maas, 2017). Also, the HSN is involved in the construction of a Historical Database of Suriname (van Galen & Hassankhan, 2018).

Some researchers (e.g. Kurosu, 2016) plead for more comparative research, following the lead of the Eurasia Project in which high-quality datasets on rural populations from several countries were compared. Indeed, this is an exciting prospect, as more life course datasets, genealogies, etc. are becoming available. HSN has already been involved in several comparative projects. Among the most detailed population registers in the world are the Taiwanese household records created by the Japanese colonial government. Since about 2000, HSN has been used to contrast Dutch family formation processes and mortality with the Taiwanese experience (Kok, Yang, & Hsieh, 2006; Lin, 2011; Riswick, 2020; Shephard, Kok, & Hsieh, 2006; Shephard, Pan, Kok, Engel, Engelen, & Brown, 2006; Shephard, Schoonheim, Chang, & Kok, 2011). Other comparative projects revolve around fertility reactions to child mortality (Reher, Sandström, Sanz-Gimeno, & van Poppel, 2017) or the intergenerational transmission of infant mortality (Quaranta & Sommerseth, 2018).

Others have advocated a more qualitative approach to demographic behavior, and have pleaded for more attention to culture and to the historical context in which, e.g. occupations and causes of death were recorded (Brandström, 2016). Indeed, a continuous effort is needed to a) understand the data we have collected in terms of their origin, specificity and local variation, b) to improve documentation and ensure awareness among users of the contextual nature of the data and c) to continue adding variables. For instance, variables on income, schooling and health can help to avoid an overreliance on occupations as the major indicator of SES. As for culture, new research into religiosity (e.g. by going into church archives or studying naming practices) can complement the already extensive findings on the effects of denomination on life course transitions. Another direction is adding data on local popular beliefs and relating them to demographic behavior. An example is research by Hilde Bras who linked a survey of the Meertens Institute (dedicated to the study of dialect and folklore) on beliefs regarding isolation of women in childbed and enchantment of children to local variations in fertility (Bras, 2014b).

Smith, Hanson, and Mineau (2016) argued that historical demography should link up with the trend towards Big Data, which would increase its relevance for demography at large. They point at the rapidly increasing possibilities to use massive numbers of digitized family trees. As soon as the HSN is linked to LINKS, the problem of ‘atomization’ by focusing on individuals devoid of their (kin) networks will be solved. It will become possible to trace ancestors, extended kin and descendants of HSN RPs and to study how the careers, migrations, family formation and health was affected by intergenerational transmission (see Bras, Van Bavel, & Mandemakers, 2013; Murphy, 2013) and by kin proximity, support and social status. Tracing descendants of HSN RPs will also stimulate collaboration between historical demographers and genetics (see also Hobcraft, 2006; Larmuseau, Van Gestelen, van Oven, & Decorte, 2013). In fact, descendants of HSN RPs are already being traced in order to link their DNA samples to the health status of their ancestors (project Genes, Germs and Resources, see https://iisg.amsterdam/en/hsn/projects/ggr).

By using remote access linkage to archives containing individual causes of death (after 1936), the HSN will also be useful for health-related research in the future. Another important step is adding the height of HSN men at age 19, determined at the physical examination for military service (research project Giants of the Modern World). HSN can also expand its mission by offering more information on the environment in which people lived: soil and water condition of the area, precipitation and temperature, population density, transport and communication infrastructure, presence of industry, and local causes of death (Eckamper, 2019). This would create the setting for a real multilevel analysis of demographic behavior. Another envisaged development is the connecting of the younger cohorts of the HSN with the Social Statistical Datasets (SSD), curated by Statistics Netherlands. The SSD integrates register and survey data on all inhabitants of the Netherlands since the 1990s, containing important socio-economic, demographic, health, crime, income and wealth variables of the complete population of the Netherlands. It includes day-by-day information on co-residence, marriage and parenthood and detailed information on parents, children, siblings, partners, colleagues, companies, communities and neighbors (Bakker, van Rooijen, & van Toor, 2014). The HSN will be connected by way of the second generations of the HSN research persons, born from 1900 onwards.

It can be foreseen that future generations of HSN users will work in even more interdisciplinary settings, e.g. by working with evolutionary biologists. But probably the largest challenge will be to translate
the findings of HSN studies to the larger historical processes. What does all the variation they find at the individual or household level imply for the trends in society observed at the aggregate level? The methodological bridge between the micro and macro levels can possibly be found in agent-based modelling. This technique goes beyond classic micro simulation (creating synthetic populations from demographic characteristics and rates culled from e.g. registers) by letting the simulated persons learn from their past, react to others, influence their surrounding and be influenced by their environment. Much more social variation can be built in such models than was possible in the past. In this way, different pathways of, for example fertility decline through diffusion of innovative ideas, can be tested (Billari & Prskawetz, 2003; Courgeau, Bijak, Franck, & Silverman, 2016; for examples of modelling the fertility decline, see Klüsener, Scalone, & Dribe, 2016; Nomes, Grow, & Van Bavel, 2019).

7 EVALUATION

We began this essay with the question on the ‘qualitative impact’ of the HSN. Has HSN changed our knowledge and perception of the history of Dutch society, and in what ways? What is HSN’s contribution to social science history in general and to the specialist field of (historical) database construction? A synthesizing social and demographic history of the Netherlands based on HSN still has to be written. As the foregoing has shown, many new insights are scattered across numerous articles. HSN-based research has shown the ‘micro-effects’, that is the implications for individuals and families, of ‘macro’ characteristics of Dutch society in 19th and early 20th centuries. Such characteristics are: a relatively wealthy country with a generous poor relief system; a late-comer in industrialization; an early adopter of the breadwinner-homemaker family model; and a country ‘pillarized’ along religious fault lines.

Although particular western areas of the Netherlands suffered economic hardship in the early half of the 19th century, we do not see the same strong effects of food price fluctuations as in other European countries. Apart from market integration, it seems that the poor relief system which was built up in more prosperous eras managed to mitigate effects of bad times and the restructuring of the economy. For instance, social class differentials in adult mortality were limited and widows seem to have been shielded against the loss of the breadwinner. Effects of industrialization (through the changing occupational structure, the valuation of achievement and female factory employment) have been studied on issues such as migration, social mobility and fertility decline. Clear effects were not found, probably because industrialization came rather late and coincided with increasing real wages and expansion of education. As for migration, the loss of opportunities in the countryside led to migration flows into the service sectors of the large cities, to emigration, and to a limited extent to the new industrial areas.

For many centuries, the Netherlands have been a typical ‘nuclear family’ country, apart from the border area with Germany. Its relative wealth made it also possible to implement the ideal of the homemaker model, with women withdrawing from the labor market and concentrating on the household. HSN had allowed to trace this development and to study the effects, e.g. on the decline of infant mortality and increased stature (read: health) of children. Finally, the Netherlands were characterized by increasing competition among religious denominations, in which (sexual) morality played an important role. In particular Roman Catholic and neo-Calvinist foremen endeavored to discipline their flocks and to moralize the country at large. HSN research shows how this played out in diverging sexual and marital behavior among different religious groups.

A typical feature of HSN research is its international character. Thus, many studies were comparative and/or oriented at an international audience. For example, several researches on fertility behavior (e.g. response to declining infant mortality, spacing of births) were done in comparison with other countries. Complete residential histories covering an entire country are extremely rare, and as such HSN has been used as a ‘benchmark’ to see where gaps and biases could occur in other data collections. Examples are the comparison of migration trajectories reconstructed from population registers (HSN) and those from genealogies (Adams, Kasakoff, & Kok, 2002) and using HSN to study the attrition process and selection mechanisms before boys could be measured at conscription age (Quanjer & Kok, 2020). Also, the Netherlands were presented as a ‘laboratory’ to study specific behavior that could not (or not yet) be studied elsewhere. A case in point is the relative impact of socio-economic status, household...
composition and religion on a host of life course transition and life events. HSN researchers eagerly joined international debates and innovations in methods and research topics. Examples are studies on (household strategies behind) leaving home (e.g. Bras, 2002), on ‘helpers at the nest’ and their effect on reproduction (e.g. Rotering & Bras, 2015) on extended kin and social networks (through witnesses) and their role on social mobility and diffusion of birth control (see Bras, 2014a), on life course approaches to anthropometric history (e.g. Quanjer & Kok, 2019) and so on.

A major spin-off of HSN is LINKS, which has quickly become an important player in the field of studies on intergenerational effects in social mobility and health. Examples are the recent studies on familial clustering of infant death and longevity. Social and life science struggle with disentangling environmental and heritable components of behavior and health, and LINKS promises to become a crucial tool, especially when the entire country is integrated in LINKS and when it can be enriched with more variables.

HSN and LINKS are much more than just datasets. The HSN is a center of expertise in collecting, harmonizing and documenting data from a variety of historical sources, in record linkage, and in creating standards for occupations and locations. Moreover, the HSN network of researchers have set an example for Dutch scholarly work (at least, in the humanities) by collaborating in interdisciplinary teams and by sharing data and tools.

8 EPILOGUE

When thirty-odd years ago the HSN was conceived, we anticipated that it was going to be a long-lasting undertaking. The data were available in Open Access from the start which ensured interest from many researchers both from the Netherlands and abroad. The database — or more precisely data warehouse — was designed to expand. By enlarging the geographical scope and the period covered and by constantly adding variables allowing for detailed life courses, the HSN ensured a growing (and even insatiable) interest from — by now — several generations of researchers. Furthermore, the many additional projects designed for specific researchers (e.g. an oversampling of city, the addition of siblings of RPs, an entire second and third generation of selected RPs) were added to the warehouse after an initial embargo period in order to be reused in further research.

What we could not anticipate around 1990 were the technological advances realized in the ICT sector in the ensuing years. At the time, we were happy with the 3.5 inch diskettes on which our employees in the archives could store the data. Now, we can effortlessly (in term of memory storage and processor speed) link HSN to other databases (e.g. using the common IDS format) or to enrich the data in the Cloud by record linkage in Open Linked Data. Obviously, the speed and scale of actual research has gone beyond our wildest dreams.

Another matter we could not foresee at the start was the sheer expansion of research questions, which of course stems from the dynamism of science itself. For instance, great strides in the sociology of the life course were made from the 1990s onwards. Also, emerging intersections between evolutionary biology, biodemography and historical demography lead to new questions being researched with HSN and similar databases. Another example is the strong rise of genealogical demography and health studies related to the revolution in genetics.

What we can foresee is that HSN and LINKS will remain at the core of an expanding ecology of national and international historical micro-level databases, provided researchers and funding agencies go on investing time, money and energy in keeping the datasets up to standards and continue catering to the ever changing needs of the research community.

ACKNOWLEDGEMENTS

We are grateful to Frouke Hansum who corrected the English text and Marja Koster who systematized our provisional reference list into the required format for this journal.
The article is written within the context of the project ‘Methodologies and Data mining techniques for the analysis of Big Data based on Longitudinal Population and Epidemiological Registers’ (LONGPOP). The project is funded by the EU Horizon 2020 research and innovation programme (under the Marie Skłodowska-Curie Grant Agreement 676060). DISCLAIMER: the results published in this website reflect only the author(s)’s view and the Research Executive Agency is not responsible for any use that may be made of the information they contain.

REFERENCES


