

## **APPENDIX D**

### **RESEARCH PROGRAMME**

#### **Life courses in context.**

A collaboratory based on Dutch population registers and censuses (19<sup>th</sup> and 20<sup>th</sup> century).

**8<sup>th</sup> August 2001**

## Introduction

The data made available in the proposed data sets will serve numerous research projects in the social sciences and history, especially social and economic history. This can be substantiated from present practice, through an international comparison and on the basis of ongoing and proposed research projects.

Present practice applies in the first place to the census data, most of which are already available in printed form. These are used intensively in historical and social research. This holds also true for the Population Registers on which the HSN dataset is based. Both will be used to an even greater length when they are readily available in digital form.

Most nations publish some form of census and have some form of population registration. The international community of scholars in the social sciences and history uses these data to test its hypotheses and compare research results internationally. This is done mostly on the basis of published censuses. In some cases the original returns to these censuses are still available, and individuals can be located in these. More often, researchers have to reconstruct data on the level of the individual from population registers or marriage acts. It is only in very few cases that Population Registers allow us to follow the individual life courses through every move of residence, as is possible with the Dutch Population Registers (Hall, McCaa and Thorvaldsen 2000). The combination of the two proposed data sets therefore makes it possible

- to compare Dutch data with all levels of data available elsewhere (and thus to execute all kinds of comparative research, or to replicate research performed elsewhere);
- to assess the error margin of foreign data through simulation, using the richer Dutch data as bench mark (Adams, Kasakoff and Kok 2001)
- to zoom in from the aggregate data on the national level to comparable data on the micro level, or to zoom out from individual data on the micro level to gauge their regional or national representativeness.

To sum up: a wealth of social and historical research executed all over the globe shows the usefulness of the kind of data to be included in the proposed data sets. Moreover, the specific Dutch situation, where very good data are available both through national censuses and through a unique Population Register, will make the combined dataset better than any available elsewhere.

The third way to substantiate the usefulness of the data is by presenting ongoing and proposed research projects. It is impossible - and unnecessary - to do this for the censuses, which are used in practically every scholarly study of the social and economic developments in the nineteenth and twentieth centuries. We therefore focus below especially on those projects, which can only be executed on the basis of the superior Dutch population registers and other sources, made available through the HSN.

Adams, John, Alice Kasakoff & Jan Kok. 2001. Family migration during the child-bearing period in 19th century Holland and the American North: A comparative analysis based on genealogies and population registers. *Annales de demographie historique* (forthcoming).

Hall, P. Kelly, R. McCaa & G. Thorvaldsen (eds.). 2000. *Handbook of International Historical Microdata for Population Research* Minneapolis: Minnesota Population Center: 149-177.

### A. Social inequality in health and mortality

Social and economic inequality is a persistent subject of social and political pressure, a continuous topic of academic interest, and a key issue of attention of policy makers. Usually, economic variables such as income and property are studied. Sen, and others with him, have argued that in judging a person's advantage or deprivation, one has to shift the attention from an exclusive concentration on income and commodities, to things that people value intrinsically, such as not dying prematurely (Sen 1998). In arguing for a wider use of mortality statistics in the analysis of inequality, Sen has also stressed that mortality rates can be measured over a longer period of time and much more unequivocal than other measures of inequality.

Long-term developments of inequality in mortality are not only part of core historical issues, such as the societal effects of urbanization, industrialization, and medicalization (Preston et al. 1981), trends in inequality over time are also an important issue in contemporary discussions on socioeconomic inequalities in health. Possible persistence of mortality differences indicates that efforts that have been made to improve the health of the lowest socioeconomic groups have not been fully effective (Antonovsky and Bernstein 1977; Smith 1991; Van de Mheen, Reijneveld and Mackenbach 1996; Woods and Williams 1995).

Studies on the long-term development of socioeconomic inequality in health and mortality have many shortcomings, among which are the usage of different measures of social class (income, occupation, or level of education), the use of rather crude mortality measures and the focus on limited age-groups. Furthermore, there is little awareness of the fact that social class and environment confound each other in influencing long-term changes. Studies have shown that the socio-historical context in which groups of individuals are embedded – socioeconomic status of the area, level of segregation of socioeconomic groups – has a strong association with their mortality level, net of individual characteristics (Hummer, Rogers and Eberstein 1998, Wilkinson 1996). A recent Dutch study, for example, showed that living in areas characterized by poor socioeconomic conditions had negative effects on health for both people with a high and those with a low socioeconomic status (Bosma et al. 2001). Present-day studies have shown that people in lower socioeconomic positions are generally worse off with respect to their health than people in higher positions (Kunst et al. 1994a, 1994b). The question which processes underlie these inequalities is still largely unanswered. Several authors have argued that socioeconomic health differences in adult life are partly explained by processes much earlier in life, in particular by childhood socioeconomic conditions (Davey Smith et al. 1994, Lundberg 1991, Lundberg 1993, Wadsworth 1986). In recent years, attention in the Netherlands, elsewhere in Europe, and in the US has in particular been directed to the role that early-life conditions play for socioeconomic health differences later in life (Kuh and Ben-Sholomo 1997). Mortality in later adulthood and old age appeared to be dependent on the opportunities for and constraints upon life styles during the first part of the life span.

To study long-term effects of early-life conditions, a prospective birth cohort study is the most appropriate research design. However, the results of these studies become available only after a long time. In addition, they suffer from high non-response and loss to follow-up which become more important as the study period is longer. That is precisely when the effect of childhood conditions on later ages might come to light. Retrospective data have a high risk of biased information, in particular as far as childhood conditions are concerned, and only describe the surviving part of the cohort.

Data from the HSN can be used to study the issue of socioeconomic inequality in a much more sophisticated way. The data set will consist of data for several provinces over the whole period 1863-2000. This was a time period during which an important epidemiological transition took place. The data make it possible to study time trends in completely differing epidemiological environments; they allow the use of identical socioeconomic status variables over regions and time; they make it possible to examine jointly the impact of area and of socioeconomic characteristics of individuals, and, finally, to study mortality differences over the whole age range. By focussing on specific dangerous occupations the effect of the working environment can be studied separately.

Further, data from the HSN will be used to examine the relationships between conditions in childhood and health in later life with mortality later in life as outcome variable. Three childhood conditions will in particular be studied in their later-life effect: the economic position of the child's family between birth and adolescence (Van de Mheen et al. 1997), the short-term and long-term effects of migration on survival (by following migrants and by comparing their mortality with that of persons from their region of origin, and of their region of destination; Coggon et al. 1997) and long-term effects of family circumstances of children during the first fifteen years of their life (Hansagi et al. 2000, Sweeting and West 1995). The length of time during which children were in advantaged or disadvantaged positions will explicitly be taken into account.

Although mortality is not registered in the Dutch censuses as such, detailed mortality tables have been published for several census years. For example, for 1899 extensive mortality tables were published in the introductory volume. They have been partially analysed by Van Poppel

(2001). These population data will especially be fruitful as a general background for the HSN data. For given points in time, it will be possible to check the longitudinal sample data of the HSN with the population data of the census.

A wide time perspective will be applied to study the effect of changes over time in the role played by various early life-conditions. Several Dutch provinces, each with its particular social and economic structure, will be studied during a time period in which the regions underwent radical changes in their economic, social and family structure. The period studied here covers the mortality transition in the Netherlands from the middle of the nineteenth century to the late twentieth century. The comparison of the mortality pattern of individuals remaining in their region of origin with that of migrants, and with the population residing in the region of destination offers the opportunity to study the effects on mortality of a variety of environmental conditions during different stages of life.

- Antonovsky, A. & J. Bernstein. 1977. Social class and infant mortality. *Social Science and Medicine* 11: 453-470.
- Bosma, H., H.D. Van de Mheen, G. Borsboom & J.P. Mackenbach. 2001. Neighborhood socioeconomic status and all-cause mortality. *American Journal of Epidemiology*.
- Coggon D, C. Martyn & C. Osmond. 1997. Mortality from aortic aneurysm in migrants between counties of England and Wales: evidence for causes acting early in life. *QJM* 90 (2):133-7.
- Davey Smith, G., D. Blane & M. Bartley. 1994. Explanations for socio-economic differentials in mortality. *European Journal of Public Health* 4: 131-144.
- Hansagi, H., L. Brandt & S. Andréasson. 2000. Parental divorce: psychological well-being, mental health and mortality during youth and young adulthood. *European Journal of Public Health* 10: 86-92.
- Hummer, R.A., R.G. Rogers & I.W. Eberstein. 1998. Sociodemographic differentials in adult mortality: A review of analytic approaches. *Population and Development Review* 24: 553-578.
- Kuh, D. & Y. Ben-Shlomo (eds.). 1997. *A life course approach to chronic disease epidemiology*. Oxford: Oxford University Press.
- Kunst, A.E. & J.P. Mackenbach. 1994a. International variation in the size of mortality differences associated with occupational status. *International Journal of Epidemiology* 23: 742-750.
- Kunst, A.E. & J.P. Mackenbach. 1994b. The size of mortality differences associated with educational level in nine industrialized countries. *American Journal of Public Health* 84: 932-937.
- Lundberg, Olle. 1991. Childhood living conditions, health status and social mobility: A contribution to the health selection debate. *European Sociological Review* 7: 149-62.
- Lundberg, O. 1993. The impact of childhood living conditions on illness and mortality in adulthood. *Social Science and Medicine* 36: 1047-52.
- Poppel, F. van. 2001. "Van vijf naar 16 miljoen. De groei van de Nederlandse bevolking in de afgelopen eeuw", in: S.M.R.N. van Maarseveen & P.K. Doorn, *Nederland een eeuw geleden geteld*. Amsterdam: IISG.
- Preston, S.H., M.R. Haines & E. Pamuk. 1981. Effects of industrialization and urbanization on mortality in developed countries. pp. 233-254 in *International Population Conference, Manila 1981, Volume II*. Liège.
- Sen, A. 1998. Mortality as an indicator of economic success and failure. *The Economic Journal* 108: 1-25.
- Smith, D.S. 1991. Mortality differentials before the health transition. Forum: Fatal years. *Health Transition Review* 1: 235-237.
- Sweeting H & P. West. 1995. Family life and health in adolescence: a role for culture in the health inequalities debate? *Social Science and Medicine* 40:163-75.
- Van de Mheen, H., S. Reijneveld & J. Mackenbach. 1996. Socioeconomic inequalities in perinatal and infant mortality from 1854 to 1990 in Amsterdam, the Netherlands. *European Journal of Public Health* 6: 166-174
- Van de Mheen, H., K. Stronks, J. Bos & J.P. Mackenbach. 1997. The contribution of childhood environment to the explanation of socio-economic inequalities in health in adult life: a retrospective study. *Social Science and Medicine* 44: 13-24.
- Wadsworth, M.E.J. 1986. Serious illness in childhood and its association with later-life achievement. Pp. 1-20 in R.G. Wilkinson (ed.), *Class and health: research and longitudinal data*. London: Tavistock.
- Wilkinson, R.G. 1996. *Unhealthy societies: the afflictions of inequality*. London: Routledge.

Woods, R. & N. Williams. 1995. Must the gap widen before it can be narrowed? Long-term trends in social class mortality differentials. *Continuity and Change*, 10: 105-137.

## B. Social movements and change

One of the key issues in the study of social movements is the question how mobilisation into a social movement takes place. Which are the characteristics of those members of society who - either under exceptional or under ordinary circumstances - organise themselves into movements aiming to change society? For several reasons, research into this question has recently turned to the individual level.

One, rather elementary, reason for this development is that recent research has suggested that the motivation for adhering to movements may - perhaps especially in the early phase of a movement - often be found at the level of the primary group. The main recruiting ground for members are the families and close friends of those who already adhere (Altena 1989, Van der Laarse 1989, Leenders 1991, Bos 2001). So far, this hypothesis has been buttressed mainly by noting family links, but it can be checked with the kind of data HSN collects.

However, a more general cause of this shift was the unsatisfactory character of earlier monocausal explanations of social movements, especially naive forms of class analysis. This was certainly the case in labour history, where the assumption of self-evident working-class solidarity was increasingly put in question from the 1980s on. As a response, scholars pointed to other major identities that competed with class identities in concrete situations, especially gender, ethnicity and religion (Geary 1999, Van der Linden and Heerma van Voss, forthcoming). Another response was a turn to micro-history. This took the form of analysing workers' ego documents, which are however very scarce, and of oral history, although this method could only be applied to recent periods. A third, more structural method consists of reconstructing workers' life courses with the help of standard sources of the type collected by HSN (some recent examples are: Kok 1997, Baines and Johnson 1999, Brooking et al. 1999, Olssen et al. 1999a and 1999b, Horrel and Oxley 2000, Wals 2001). This focussed the attention of historians on other living strategies of households and individual workers, besides belonging to a labour movement, like taking in boarders, building up a business or working a plot of land. Within the context of these sets of living strategies the organisational behaviour of workers can be explained better than before.

The IISG (International Institute of Social History) is conducting research on working class living strategies, based on HSN data. Besides Kok 1997, another good examples is a PhD project on autochthonous workers in Rotterdam (1850-1980), which recently started in connection with Leo Lucassens 'pioneer' project on migration mentioned under F. A pilot project demonstrated that if membership lists of militants are available, it is possible to analyse militants and non-militants living strategies, identify differences and offer plausible explanations for the choice to adhere to the movement or not (Heerma van Voss and Vermeulen 2000). The IISG will continue pursuing research along these lines, using the HSN database, supplemented with the censuses.

The opportunities for this kind of research obviously are not restricted to research into the labour movement. The NIOD (Netherlands Institute of War Documentation) holds the membership lists of the NSB (1931-1945), the largest Dutch national-socialist movement. The social base of this movement has given rise to much debate, as was the case elsewhere in Europe (Vellenga 1975, Bergmann 1996). NIOD staff has shown interest in using the HSN database to analyse the life courses of a sample of NSB-adherents.

Altena, B. 1989. *'Een broeinest der anarchie'. Arbeiders, arbeidersbeweging en maatschappelijke ontwikkeling. Vlissingen 1875-1929 (1940)*. Haarlem: Thesis.

Baines, D. & P. Johnson. 1999. In Search of the 'Traditional' Working Class: Social Mobility and Occupational Continuity in Interwar London. *Economic History Review*, 52: 692-713.

Bergmann, J. 1996. Handwerk und Tradition. Zum Charakter und zur Bedeutung des handwerklichen Traditionalismus bis zum Ende der Weimarer Republik, *Zeitschrift für Geschichtswissenschaft* 44 (10): 869-896.

- Bos, D. 2001. *Waarachtige Volksvrienden. De vroege socialistische beweging in Amsterdam 1848-1894* Dissertatie Universiteit van Amsterdam.
- Brooking, T., D. Martin, D. Thomson & H. James. 1999. The ties that bind: persistence in a New World industrial suburb, 1902-22. *Social History* 24: 55-73.
- Geary, D. 1999. Working-Class Identities in Europe 1850s-1930s, *Australian Journal of Politics and History*, 45 (1): 20-34.
- Heerma van Voss, L. & F. Vermeulen. 2000. 'Back to the country. Living Strategies of Twente Textile Workers ca. 1920'. *Working paper HSN*.
- Horrell, S. & D. Oxley. 2000. Work and Prudence: Household Responses to Income Variation in Nineteenth-Century Britain'. *European Review of Economic History*, 4: 27-57.
- Kok, J. 1997. Youth Labor Migration and its Family Setting, The Netherlands 1850-1940, *History of the Family*, 2: 507-526.
- Laarse, R. van der. 1989. *Bevoogding en bevinding. Heren en kerkvolk in een Hollandse provinciestad, Woerden 1780-1930* Dissertatie Universiteit van Amsterdam.
- Leenders, J.M.M. 1991. *Benauwde verdraagzaamheid, hachelijk fatsoen. Families, standen en kerken te Hoorn in het midden van de negentiende eeuw* Dissertatie Universiteit van Amsterdam.
- Olszen, E & H. James. 1999a. Social Mobility and Class Formation. The Worklife Social Mobility of Men in a New Zealand suburb, 1902-1928. *International Review of Social History* 44: 419-449.
- Olszen, E., T. Brooking, B. Heenan, H. James, B. McLennan & C. Griffen. 1999b. Urban society and the opportunity structure in New Zealand, 1902-22: the Caversham Project'. *Social History* 24: 39-54.
- Van der Linden, M & L. Heerma van Voss (eds.). forthcoming (2001). *Class and other Identities. Entries to West European Labour Historiography*. New York/Oxford: Berghahn.
- Vellenga, S.Y.A.. 1975. *Katholiek Zuid Limburg en het fascisme. een onderzoek naar het kiesgedrag van de Limburger in de jaren dertig*. Assen: Van Gorcum.
- Wals, H. 2001. *Makers en stakers. Amsterdamse bouwvakarbeiders en hun bestaansstrategieën in het eerste kwart van de twintigste eeuw*. Dissertatie Universiteit van Amsterdam

### C. Culture, religion and education

One of the key features of Dutch society in the nineteenth and twentieth century is its religious segmentation, usually referred to as the *verzuiling* ('pillarisation') of Dutch society. The Batavian Revolution provided Roman-Catholics, Lutherans, Mennonites, Remonstrants, and even Jews with the same formal rights as the representatives of the former dominant Calvinist Church. From that moment on, an emancipation process took place that was not finished until the second half of the twentieth century.

Despite a nation-building policy of the Dutch government in the first half of the nineteenth century, which tried to ban religion from the public to the private sphere, religious identity became increasingly important (Knippenberg 1999). Religion became the main marker of cultural differences. Several secession movements from the Calvinist mother church enlarged these differences, the 1834 *Afscheiding* and the 1886 *Doleantie* being the most important.

The struggle between the nation-building state and the different denominations was fought predominantly in the education sector (Te Velde and Verhage 1996). The initial dispute about freedom of education – settled in the 1848 Constitution – changed into a dispute about equal financial rights for public and denominational schools. This dispute was settled in 1917, when the principle of equal financing of public and private (primary) schools was inscribed in the Constitution. In fact, that Constitution institutionalised religious diversity and introduced the proportionality principle (each religious group is entitled to its proportional share in state funding) in the political system, which became the main device for the allocation of (national and local) government money in other sectors besides education as well.

At the same time, modernisation and infrastructural integration influenced local and regional cultures, which became less isolated and underwent the unifying power of a national culture (Knippenberg and De Pater 1988). The school dispute and general modernisation processes also increased interest in school education in general, concerning both primary and secondary schools. Higher levels of school attendance and (quality of) education, and lower levels of illiteracy were the result.

The HSN and several (though not all) census data sets, which include data on for instance religion and illiteracy, will offer unique and outstanding opportunities to improve our knowledge in the fields of culture and education described above. For the first time it will be possible to test hypotheses at the individual level, with data that are representative for Dutch society as a whole. Moreover, the data sets will enable the study of regional or social differences in individual behaviour in this field as well. The combination of HSN data with aggregate statistical data will provide the opportunity for multi-level research (see heading G).

It has been hypothesised that around 1800 there were big differences between Calvinists and Roman-Catholics as far as levels of literacy and education were concerned and that these differences faded away during the nineteenth and twentieth centuries as a consequence of emancipation and modernisation (Knippenberg 1986, Boonstra 1993). Similar hypotheses exist with respect to occupations. Since during the Dutch Republic only Calvinists were allowed to work in the field of public service and education, in the beginning of the nineteenth century Roman-Catholics were far underrepresented in these occupational sectors. This underrepresentation diminished during the nineteenth century, but did not disappear. HSN and census data make it possible to describe these emancipation processes for different social classes, occupational sectors, regions, and local contexts.

Not only Catholics, but also Jews were emancipating during the nineteenth and first half of the twentieth century, although this did not give rise to a Jewish pillar (Blom et al. 1995). They increasingly went to public schools, diminishing Jewish illiteracy, and acquired access to occupations that were closed before. However, they differed from other religious groups with respect to involvement in migration processes and they showed a different demographic behaviour. The data sets will improve our knowledge on the persistence and disappearance of these differences, especially when linked with the database of the Digitaal Monument Joodse Gemeenschap, which is also under construction at the IISG and the NIWI in conjunction with the HSN.

Further, there is much debate about the class position of members of the Dutch Reformed Church that left their church during the two great nineteenth century secession movements: the 1834 *Afscheiding* and 1886 *Doleantie* (See for instance Kuiper 1986). HSN data, supported by census information, will for the first time allow to give decisive answers to the many questions that have arisen in that debate by characterising these members by occupation, age, literacy, and local community.

The pillarisation on the part of (orthodox) Protestants and Roman-Catholics expressed itself in various ways. More attention was devoted to the survival of one's religious group, giving rise to high birth rates. The intergenerational transfer of membership of religious groups was strengthened by the promotion of church marriages, and by church resistance to religiously mixed marriages. Households became more uniform denominationally, as servants or lodgers were chosen only from the same denominational group. Social contacts were restricted to members of the same church. Data on fertility, religious composition of marriages, household composition and witnesses present at marriages can be used to test the importance of these mechanisms and to investigate regional and class differences.

Whereas pillarisation increased cultural differences, the spread of literacy and educational participation caused regional, religious, and class differences to slowly fade away. There are many (political, economic, and cultural) theories about developments of and differentiations in illiteracy according to class, occupation, sex, religion, region, and urbanisation (Boonstra 1993). The data sets to be compiled will provide excellent opportunities to test hypotheses derived from these theories. A combination of HSN data with archival sources from secondary schools will even make it possible to determine the impact of secondary education and of different individual characteristics of pupils and their parents on their occupational career (Mandemakers 1996).

A final example of cultural differences concerns name giving. Regional differences and developments in name giving are very interesting topics in the study of names (Gerritzen 1998). These differences could, for instance, be related to dialectal data from other sources. HSN data also allow the exploration of differences in name giving according to class, occupation, and religion in urban and rural contexts. Changes in time could be determined and explained.

- Blom, J.C.H., R.G. Fuks-Mansveld & I. Schöffer (eds.). 1995. *Geschiedenis van de joden in Nederland*. Amsterdam: Balans.
- Boonstra, O.W.A.. 1993. *De waardij van eene vroege opleiding. Een onderzoek naar de implicaties van het alfabetisme op het leven van inwoners van Eindhoven en omliggende gemeenten, 1800-1920*. Dissertatie LU Wageningen.
- Gerritzen, Th.J.M. 1998. *Onderzoek naar een aantal aspecten van naamgeving in Nederland* Dissertatie Universiteit van Amsterdam.
- Knippenberg, H. 1999. National integration and growing ethnicity among Dutch Roman Catholics and Protestants in the nineteenth and early twentieth century. The impact of education. *The Netherlands' Journal of Social Sciences* 35: 37-52.
- Knippenberg, H. & B. de Pater. 1988. *De eenwording van Nederland. Schaalvergroting en integratie sinds 1800*. Nijmegen: SUN.
- Kuiper, D.Th. 1986. De Doleantie en de Nederlandse samenleving. pp. 203-239 in: W. Bakker et al., *De Doleantie van 1886 en haar geschiedenis*. Kampen: Kok.
- Mandemakers, C.A. 1996. *Gymnasiaal en middelbaar onderwijs. Ontwikkeling, structuur, sociale achtergrond en schoolprestaties, Nederland, ca 1800-1968*. Dissertatie Erasmus Universiteit Rotterdam.
- Te Velde, H. & H. Verhage (eds.). 1996. *De eenheid en de delen. Zuilvorming, onderwijs en natievorming in Nederland 1850-1900*. Amsterdam: Het Spinhuis.

#### **D. Households and changing labour relations**

Central research questions within both the NW Posthumus Institute and IISH relate to the interactions between changing labour relations and individual households. In their responses to macro level changes, families themselves created the conditions for further changes (e.g. by emigration or by fertility change). The research programmes that are being developed at the moment focus mainly on the analysis of decision-making processes within families (Baud and Engelen 1994). The family is seen as a dynamic constellation of individuals. Consequently, conflicts between individual and group interests, as well as between generations and sexes, occur. Besides, families have very different strategic 'repertoires' and scopes for future planning, dependent on their socioeconomic position and cultural preferences. The HSN offers excellent material for a further development of this field of research. The censuses offer detailed information on household composition and the structure of the labour force, albeit not in a longitudinal way, to supplement the HSN.

First, our knowledge of the household within the societal context can be considerably enhanced (Wall et al. 1983, Verdon 1998). The nature of intrafamilial relations depends, among other factors, on the stability and duration of cohabitation and the presence of strangers (step family, boarders) in the household. The composition of the household was also part of the strategic repertoire of families. Taking in relatives or boarders, or expelling children was used to increase the welfare of the family as a whole or of some of its individual members. Very little, though, is known about the reasons for variation in household composition. Extended families were, for example, far more common in the East of the Netherlands than elsewhere. Hypotheses on the influence of household production, inheritance and regional customs concerning the living-in of unmarried relatives exist, but have hardly been tested. With the HSN data the types of households in which individuals lived across their life course can be observed. Herewith researchers can go beyond the classic concept of the family life cycle. The integration of the experiences of broken families, permanently single persons and infertile couples into social and demographic history is feasible because HSN breaks with the traditional preference for married couples and complete families. The population registers that are central to the HSN data offer unique data on the structure of households and the timing of changes. Hypotheses on regional variation can be tested, including factors such as migration and religion.

International variations in the relations between households, kinship networks and institutions will be investigated in the context of a comparative research program of, among others, NIDI, Universities in Norway, Belgium and Switzerland, and EHESS (France). The aim of this project is to study coherent long term developments of family forms over the last centuries

in three key European regions - northern, central and southern - within a policy relevant contextualisation. Part of the research agenda is to assess the relative importance of familial versus non-kin solutions with respect to the care for orphans, for the elderly and for unmarried adults. The household data that will become available in the datasets of the HSN and the population censuses will offer an excellent opportunity to investigate this topic.

The internal division of tasks within families relates both to the strategic behaviour of families and to potential conflicts within families. However, little is known about the extent and nature of women's and children's work, which generally elude official recordings (Tilly and Scott 1978, Mitterauer 1992). HSN offers a variety of ways to fill the existing gaps. Civil registers and population registers offer occupational titles for men and women of differing ages. This information can be compared with the occupational censuses as well as placed within the familial context. Secondly, the life courses themselves show how careers were formed, e.g. by transmission of the parental occupation, by apprenticeships or by formal education. A linkage with, for instance, patent registers will reveal female labour activities. And finally, information will come available on differential family strategies. Families may spread their risks by placing children in different sectors of the economy. Also, the timing of leaving home elucidates the choices within labouring families between employing children within the household itself and employing them elsewhere to bring in wages (Oris and Van Poppel forthcoming).

Household behaviour needs to be placed in an economic context that is as complete as possible. The refinement of theories in social and demographic history has been hampered by the exclusive concentration on occupational titles. In this way information on subsidiary activities, on income gained, and the meaning of property for familial decisions is ignored (Smith 1984). Here HSN offers an excellent starting point for a correction of the existing literature. For (part of) the HSN cohorts, the occupational titles of the sample persons and their parents will be appended with information from linked land registers, patent and tax records. Tax records are intact for at least half of all communities in the period 1870-1922. Standard demographic rates on, for instance, child mortality and age at marriage, can be drawn up for income and property groups and compared to rates based on occupational classifications. Also, the meaning of (parental) property for life courses and family strategies can be assessed. A pilot project in this sense is the current research of Jan Kok who combines HSN material with *kadaster* (land register) and notarial records in order to study the effect of property transmission on household formation. The population, housing and occupational censuses offer detailed information on household composition, housing situation and structure of the labour force, which will supplement the HSN.

Baud M. & Th. Engelen (eds.). 1994. *Samen wonen, samen werken? Vijf essays over de geschiedenis van arbeid en gezin*. Hilversum: Verloren.

Mitterauer, M. 1990 *Familie und Arbeitsteilung. Fragestellungen und Zugangsweisen*. Wien/ Köln: Böhlau.

Smith, R.M. (ed.). 1984. *Land, kinship and life-cycle* Cambridge: Cambridge UP.

Tilly, L.A. & J.W. Scott. 1978. *Women, work and family*. New York etc.: Holt, Rhinehart and Winston.

Oris, M. & F. Van Poppel (eds.). forthcoming, *Leaving home in Eurasian perspective*.

Verdon, M. 1998. *Rethinking households. An atomistic perspective on European living arrangements*. London: Routledge.

Wall, R., J. Robin & P. Laslett (eds.). 1983. *Family forms in historic Europe*. Cambridge: Cambridge UP.

## **E. Social mobility and the life course**

Various research schools and institutes (ICS, Amsterdam School for Social Research, NEHA) have expressed their ongoing interest in social mobility studies based on HSN material. These studies range from international comparative analyses of societal 'openness' to innovative studies on specific social groups. Much is known about social mobility and the openness of society in the second half of the twentieth century. Knowledge about earlier time periods, however, is scattered. Data that cover a longer period are scarce, not only in the Netherlands. And where such data exist, they are usually local and results cannot be generalized to the

national level (and migration cannot be taken into account). The HSN and census data sets will provide a unique combination of nationally representative data that stretch over a long period.

The study of career mobility, intergenerational mobility and marriage mobility will be boosted by this national sample spanning large parts of two centuries. The HSN database offers many occupational titles pertaining to different stages of individual lives. Also, comparison of father and son (in law) is feasible. By converting the occupations in international classification schemes, such as the recently developed HISCO, and standardizing the grouping of occupations into social classes, world-wide comparative studies are possible, and are in fact being planned (Van Leeuwen, Maas and Miles, forthcoming). It will become possible to answer the question whether the renewal of the occupational structure took place from one generation to the next or within individual lives. Besides, the wealth of occupational titles in the database allows for the study of urban segregation of occupations, of processes of diffusion of 'new' occupations, and of concentration of occupations in specific areas. As occupational titles are collected at different points in life (including at marriage, or at the marriage of children), it will be possible to reconstruct occupational prestige for different birth cohorts and at different ages for the time period which is beyond the reach of present day social surveys (Liefbroer and Dykstra 2000).

The labour and industrial censuses offer the most detailed systematic lists of occupational titles and industries and their classification since the last decade of the 19<sup>th</sup> century. For instance, in the 1899 census over 6000 occupations and groups of occupations are mentioned. The occupational titles of the 1889 census have been used in several attempts to stratify the Dutch population according to socio-economic status, and recently the first analyses of the socio-economic structure of the labour force in 1899 have been made. As yet, no attempt has been made to link the occupational and industrial information from the censuses for the years until 1947. Only for the post-war years the CBS has made linkage schemes to make the occupational and industrial data comparable over time.

Research to extend the linkage schemes for the pre-WW II period would make it possible to analyse the changes in the composition of the Dutch labour force over a much longer period than has been done so far. This can then be done both at the regional (municipal) level and for specific occupational categories. A linkage scheme for occupational titles and groups from 1889 until 1971 will be a basic source in itself for further research on a variety of topics related to the social and economic history of The Netherlands. It is moreover extremely useful to link an emerging Dutch standard occupational classification system to similar systems that have been developed internationally (Van Leeuwen, Maas and Miles, forthcoming).

Nationally representative data will not only increase knowledge on social mobility at the national and international level, but also allow to study the development of specific social groups. Firstly, starting from the HSN-sample, an analysis can be made of the ways in which the traditional Dutch elites reoriented themselves after the expansion of voting rights, the changing structures of trade and industry and the declining importance of land ownership (Dronkers and Stokman 1984). For this research, the HSN database will be linked to sources on the Dutch nobility. Secondly, there is a strong interest - e.g. within the NW Posthumus Institute and Netherlands Economic History Archive (NEHA) - in research projects concentrating on entrepreneurs. The historiography of this field suffers from teleology, because of the strong bias towards successful entrepreneurs. Recent attempts to include small entrepreneurs and the middling classes in historical research suffer from a relative lack of sources (Crossick and Haupt 1998). Research into the determinants of commercial success faces the problem of how to draw a sample from small or beginning entrepreneurs and how to access sources. HSN can be used as a starting point, by tracing persons who have, at one time in their lives, started their own business. Their life courses will be analysed to explain their eventual success or failure as businessmen and businesswomen. Also, an extensive survey of remaining archives of their businesses can be endeavoured.

Not only elites and economic success but also proletarianization and impoverishment will be the topic of research using the HSN data and the censuses. The development of wage labour in certain periods and areas can be traced by combining occupational titles in a particular stage of the life course with those of the persons and positions held previously. The diverse family backgrounds and career trajectories of wage labourers help to understand their attitudes towards collective action and the formation (or absence) of group consciousness. The intergenerational transmission of life chances and attitudes is also important in the study of the proletarian 'sub-societies' found in impoverished areas or city quarters (Katz 1993). Was the 'deviant' behaviour often ascribed to these groups (criminality, illegitimacy) indeed a 'cultural' trait passed on from one generation to the next?

- Crossick, G. & H.-G. Haupt. 1998. *The Petite Bourgeoisie in Europe 1780-1914*. London and New York: Routledge.
- Dronkers, J. & F.N. Stokman (eds.). 1984. *Nederlandse elites in beeld; recrutering, samenhang en verandering*. Deventer: Van Loghum Slaterus.
- Katz, M.B. (ed.). 1993. *The 'underclass' debate. Views from history*. Princeton: Princeton UP.
- Liefbroer, A.C. & P.A. Dykstra. 2000. *Levenslopen in verandering. Een studie naar ontwikkelingen in de levenslopen van Nederlanders geboren tussen 1900 en 1970*. Den Haag: Sdu Uitgevers.
- Van Leeuwen, M.H.D. I. Maas & A. Miles. forthcoming (2001). *HISCO. Historical International Standard Coding of Occupations* Leuven: Leuven University Press.

## F. Migration

Migration is a hot topic, both in the Social Sciences, and in History. Migration history was originally mainly confined to traditional immigration countries like the USA and Canada, but European historians have become aware of the important role of migration in historical developments on a national (e.g. for the Netherlands Lucassen and Penninx 1997), a European (Page Moch 1992, Canny 1994, Bade 2000) and a global (Lucassen and Lucassen 1997, Hoerder 2001) scale.

An important recent paradigm shift in migration history concerns the unit of analysis, which has become the individual migrant. Instead of the study of aggregate migration streams, historians try to study the collective behaviour of individual migrants. Three considerations have led to this shift of the research designs from an aggregate level to an individual level. First, the increasing awareness that migration in most cases is not just the movement of a person from point A to B, but rather an intricate pattern, often involving return migration. Secondly, the wide variations of integration patterns as migrants are highly selective both in their marriage patterns (endogamy versus exogamy) and in their membership of "ethnic" organisations. In the third

place, in not a few cases reliable aggregate data can only be reached on the basis of the study of individual cases.

In this respect a major difficulty for a more remote past, say the period before the Second World War, is the lack of systematic data for individual migration movements - outside the American passenger lists which have been studied extensively. This is even more true for the period before the First World War, when population registration and travel documents were considered by most governments to be of little importance. This is most sharply reflected in the lack of sources for overland migrations, in particular short distance migrations and migrations which were considered to be "temporal".

Recently, historians are looking for methods to fill this gap. One method is genealogical research (e.g. Rosental 1999), another is the study of population samples on a local (e.g. Fertig 2000) or on a wider geographical level. HSN is one of the very few data sets which offers possibilities in the latter sense. Three types of research offer themselves:

- internal migration of persons born in the Netherlands from the early nineteenth to the early twentieth centuries, as has been shown already in particular by Jan Kok for the province of Utrecht (Kok 1999) and by Hilde Bras for female migrants from the province of Zeeland (Bras 1998 and in her forthcoming doctoral dissertation, to be defended at Utrecht University). One of the questions which will be studied within this larger field is the complex migration pattern of Jews within the Netherlands.

- overseas emigration from the Netherlands to the three main destinations. Of these the USA, for which passenger lists are available, have been studied already extensively (e.g. Galema 1996), but these studies ask to be linked to HSN in order to show the selectivity of emigration. On the contrary, emigration to the Dutch East Indies (1600-1950) and to Germany (1870-1940) have hardly been touched upon (for Germany only Kösters-Kraft 2000). Recently NWO has funded a postdoctoral project which will allow Bosma to study the Indonesian case starting from HSN data.

- immigration to the Netherlands by people born outside the Netherlands by definition is not part of the original HSN set-up. For a full picture of the population of the Netherlands HSN data have to be linked systematically to those of foreign immigrants living in the Netherlands. Apart from the ongoing "pioneer" project of Leo Lucassen (Lucassen 1999) many possibilities offer themselves where the full original census records of 1830 and 1840 (and in a different form 1850, 1860, etc.) have been preserved, like in Amersfoort, Gouda, Rhenen, Utrecht, Rotterdam and Watergraafsmeer.

Although the censuses do not allow for the analysis of individual migration histories, they are an indispensable source of basic information on the gross migration rates at the municipal level. Especially in combination with the data on differential patterns of population growth, natality and mortality as available in the 'Hofstee-files' held at the NIDI (now part of the Historical Data Bank of Dutch Municipalities), a reconstruction of gross migration flows can be made. The censuses also supply information on 'foreigners' who settled in The Netherlands, which will make it possible to investigate which places attracted in-migrants.

Bade, K.J. 2000. *Europa in Bewegung. Migration vom späten 18. Jahrhundert bis zur Gegenwart*. München: Beck.

Bras, H.. 1998. Domestic service, migration and the social status of women at marriage. The case of a Dutch sea province, Zeeland 1820-1935. *Historical Social Research*. 23 (85): 3-19.

Canny, N. (ed.). 1994. *Europeans on the Move. Studies on European Migration 1500-1800*. Oxford: Clarendon.

Fertig, G. 2000. *Lokales Leben, atlantische Welt. Die Entscheidung zur Auswanderung vom Rhein nach Nordamerika im 18. Jahrhundert*. Osnabrück: Rasch.

Galema, A.. 1996. *Frisians to America 1880-1914. With the Baggage of the Fatherland*. Groningen: REGIO-project.

Hoerder, D. forthcoming (2001) *Cultures in Contact. European and World Migrations, 11th Century to 1990s* Durham, N.C.: Duke UP.

Kösters-Kraft, M. 2000. *Wanderarbeit und Grossbaustelle. Niederländer beim Bau des Dortmund-Ems-Kanals 1892-1900*. Osnabrück: Rasch.

- Kok, J. 1999. Migratie als gezinsstrategie in midden-Nederland, 1850-1940. pp. 89-156 in J. Kok et al. *Levensloop en levenslot. Arbeidsstrategieën van gezinnen in de negentiende en twintigste eeuw*. Groningen/Wageningen: Nederlands Agronomisch Instituut.
- Lucassen, J. & L. Lucassen (eds.). 1997. *Migration, Migration History, History*. Bern: Lang.
- Lucassen, J. & R. Penninx. 1997. *Newcomers. Immigrants and their descendants in the Netherlands 1550-1993*. Amsterdam: Spinhuis.
- Lucassen, L. 1999. Een kwestie van een lange adem. De determinanten van het vestigingsproces van immigranten in de Nederlandse samenleving, 1860-1960. *Facta*. 7 (8): 18-21.
- Page Moch, L. 1992. *Moving Europeans. Migration in Western Europe since 1650*. Bloomington: Indiana UP.
- Rosental, P.-A. 1999. *Les sentiers invisibles. Espace, familles et migrations dans la France du 19e siècle*. Paris: ESSHC.

## G. Methodological issues of data-integration across levels of aggregation

The data sets to be produced by the digitization programme will be on different levels of aggregation. The HSN data set is explicitly on the level of the individual and household, and will include personal names. The census data is on a variety of aggregation levels, which will not be the same throughout the period. For the most recent census years, micro-data is available, although the records are anonymous. For the census years before WW II the data are only available in tabular form, although the published censuses of 1889 and 1899 are very detailed. Another distinction between the two data collections is that the HSN data are longitudinal, whereas the census data are transversal (cross-sectional). Finally, the HSN is a sample, while the census covers the whole population of the Netherlands.

From various points of view, ranging from legal to methodological, it is important to study the potential and the dangers of the integration of the various data sources. At least four interrelated research questions can be formulated:

- 4 In which ways can the statistical disclosure of sensitive information in the sources be prevented (or limited)?
- 5 In which ways is it possible to link data from the HSN to that of the censuses (without breaching confidentiality/privacy of respondents)?
- 6 What are the possibilities to use synthetic estimation methods to reconstruct individual records from (low-level) aggregate census tables?
- 7 How can the census data be used to validate the sample data of the HSN?

The first question concerning the statistical disclosure of sensitive data is of relevance both for the censuses and for the HSN, especially for the individual data from the 20<sup>th</sup> century. Internationally the risk of statistical disclosure of sensitive data has recently gained attendance. Expertise in this field has been developed in the CBS and elsewhere, for example in the context of the European Esprit project on Statistical Disclosure Control (e.g. Mokken et al. 1992; Bethlehem, Keller and Pannekoek 1990; Elliot 1996). In the literature on data mining, the issue is also discussed extensively (see, e.g. Agrawal and Srikant 2000). Even if the risk of the breach of confidentiality when using historical data is small in comparison to contemporary data, and is even negligible for early 20<sup>th</sup> and 19<sup>th</sup> century data, it still is of some relevance to take it into account both for the censuses and for the HSN, especially for the late 20<sup>th</sup> century individual data. Building upon the expertise developed at the CBS, procedures will be developed to counter the risk of statistical disclosure when these data are made accessible to the scholarly community.

The second research question of linking the HSN data with the censuses is in a way the opposite of the previous one. Here the issue at stake is how the individual, longitudinal data from the population registers and the cross-sectional, mostly aggregate data can enrich each other. The censuses offer many background variables that are not available in the population registers of the HSN. On the other hand, the detail of individual persons and households of the HSN is not available for most census years. In analyses at the individual level, ecological effects of higher levels (groups, municipalities, regions) may be taken in consideration. Methodologically speaking, this family of techniques, known as multi-level or cross-level analysis, offers solutions for variables operating at different levels that may influence each other (Snijders & Bosker 1999; Goldstein 1994; Hox 1995; Gerland 1996). In other words, multilevel modelling is a family of statistical procedures that try to come to terms with influences that are located on different levels of aggregation. The combined analysis of the data sets will stimulate the application and further development of such techniques, which have rarely been employed in historical research.

The third research question on synthetic estimation methods deals with attempts to reconstruct "synthetic" individual records from aggregate data. This technique is related to the estimation of the filling of a table given the marginal totals. For example, Paass (1989) describes a stochastic modification algorithm used to construct a synthetic sample X from different input sources, the sources being independent samples or summary statistics from an underlying population. The first step in the process is to construct an X as a best fit to the data by a maximum likelihood or minimum cost criterion, and the second step is to generate a sample with

a cost value near the minimum which also has maximum entropy. Rubin (1993) proposes that only "synthetic data" rather than actual micro-data should be released as a method of disclosure risk avoidance. The synthetic data would be generated using multiple imputation. They would look like individual reported data and would have the same multivariate statistical properties. However, with this scheme there would be no possibility of disclosure, as no individual data would be released. Similar techniques may be developed to attempt to reconstruct synthetic individual records from detailed census tables, eg. for those of 1889, 1899 and 1947.

The fourth research question specifically relates to the validation of the sample data of the HSN with the census. This problem area is related to several aspects of the previous ones. The sample of the HSN should be random and unbiased, but the censuses may offer means to check whether this is true. Synthetic estimation may be used to estimate coverage correction factors which can then be applied to the sample files.

Agrawal, Rakesh & Ramakrishnan Srikant. 2000. *Privacy preserving data-mining*. IBM Almaden Research Center (San Jose, CA).

<http://www.almaden.ibm.com/cs/people/srikant/papers/sigmod00.pdf>

Bethlehem, J.G., W.J. Keller & J. Pannekoek. 1990. Disclosure Control of Microdata. *Journal of the American Statistical Association* 85 (409): 38-45.

Elliot, Mark. 1996. Attacks on census confidentiality using the Sample of Anonymised Records: an analysis (Paper for presentation to 3rd International Seminar on Statistical Confidentiality, Bled, Slovenia, 2nd October 1996).

Gerland, Patrick. 1996. Socio-economic data and GIS: datasets, databases, indicators and data integration issues. Paper presented at the UNEP/CGIAR (Consultative Group on International Agricultural Research), Arendal III Workshop on Use of GIS in Agricultural Research Management. Norway, June 17-21, 1996.

Goldstein, Harvey. 1994. 1999R, *Multilevel Statistical Models*. London: Edward Arnold.

Griffin, Mark A. & David A. Hofmann. 1997, Hierarchical Linear Models in Organizational Research: Cross-level Interactions *Research Methods Forum* (2, Summer).

[http://www.aom.pace.edu/rmd/1997\\_forum\\_hlm\\_models.html](http://www.aom.pace.edu/rmd/1997_forum_hlm_models.html)

Hox, J.J. 1995. *Applied multilevel analysis*. Amsterdam: TT-Publicaties.

Mokken, R.J., P. Kooiman, J. Pannekoek & L.C.R.J.. Willenborg, 1992. Disclosure Risks for Microdata. *Statistica Neerlandica* 46 (1): 49-67.

Paass, G. 1989. Stochastic Generation of a Synthetic Sample from Marginal Information *Proceedings of the Bureau of the Census Fifth Annual Research Conference*. Washington, DC: Bureau of the Census: 431-445.

Rubin, D. 1993. Discussion, Statistical Disclosure Limitation. *Journal of Official Statistics* 9 (2): 461-468.

Snijders, Tom & Roel Bosker. 1999. *Multilevel Analysis: An introduction to basic and advanced multilevel modeling*. Beverly Hills, etc.: Sage.