NOMINATIVE DATA IN DEMOGRAPHIC RESEARCH IN THE EAST AND THE WEST

Edited by:
Elena Glavatskaya, Gunnar Thorvaldsen,
Georg Fertig, Mikolaj Szoltysék
NOMINATIVE DATA IN DEMOGRAPHIC RESEARCH IN THE EAST AND THE WEST

Monograph

Ekaterinburg
Ural University Press
2019
The main aims with this book are to compare source materials, databases and research results, as well as creating new opportunities for collaboration in the field of social and population history in the East and the West.

All the contributions are based on nominative source material, mainly censuses and vital records, which have been preserved, scanned, transcribed into databases in order to be used for cross-sectional and longitudinal research.

The chapters in the first part of this book mostly focus on the construction of nominative databases in Germany, Spain and Romania. The chapters in the second and third part are case studies on the relationship between marriage and fertility; mortality and fertility; marriage behavior and religion; urban mortality; migration, etc. made on the Russian, Austrian, Estonian, Hungarian and Norwegian databases.
CONTENTS

Introduction: Demography and Databases in the East and the West.
Gunnar Thorvaldsen, Elena Glavatskaya, Mikolaj Szoltysek ............ 5

Part 1
Demographic Data
Chapter 1. Troublesome Riches: Genealogical data as sources for historical demography in Germany. Georg Fertig ............... 14


Part 2
Family and Fertility
Chapter 4. Marital fertility in Albania during WWI. Siegfried Gruber .... 80


Chapter 6. The Influence of Childhood Mortality on Subsequent Fertility During the Demographic Transition. Mark Gortfelder ........ 121

Chapter 8. Urban and rural families in late 18th century Saint Petersburg province according to the 5th tax revision (revizskie skazki).
Maria Markova .............................................. 156

Part 3
Migration and Mortality


Chapter 11. Infant mortality in the late 19th and early 20th century Urals: macro and micro analyses. Dmitrii Bakharev, Elena Glavatskaya....... 202

List of the contributors ........................................ 220
INTRODUCTION:
DEMOGRAPHY AND DATABASES
IN THE EAST AND THE WEST

Gunnar Thorvaldsen,
Elena Glavatskaya,
Mikolaj Szoltyszek

A conference held on the border between Asia and Europe inspires dual vision: looking eastwards as well as westwards. Our main aims with this conference were comparing source materials, databases and research results, as well as creating new opportunities for collaboration.

In terms of historical sources, a comparison of the quite accurate Chinese Han dynasty census from 2 CE with the heterogeneous descriptions of property in the British Domesday Book a millennium later, makes it clear that the history of population lists is longer in the East than in the West. For military and taxation purposes, the Mongols extended the Chinese enumerations to Russian territory in the thirteenth century. Russia continued this tradition, and the first census-like (tax) lists seen by West Europeans after the end of the Roman Empire were probably Russian in origin. The German ambassador to Moscow, Baron von Herberstein, in the sixteenth century reported that censuses were taken biannually. In the 19th century, bilateral contacts between statisticians and international statistical conferences spread the practice of nominative census taking first to Great Britain in 1841, next to the United States in 1850 and later even further afield. This introduction briefly browses the history of the nominative source material, and how many of these original census manuscripts have been preserved, digitized, transcribed and disseminated as historical microdata – oftentimes through international cooperation (Thorvaldsen 2016). Interspersed are references to the contents of the other articles in this volume.
Highlights from the history of the sources

Unfortunately, nearly all census manuscripts from the first Russian imperial census of 1897 and the Soviet censuses from 1920 onwards were destroyed in order to save archival space. The authorities even destroyed the 1937 census *aggregates* because the results were politically unacceptable. The preserved manuscripts are a few local 1897 questionnaires; the computerized versions for the most recent decades; and significant parts of the 1926–27 Polar census – the most comprehensive census taken anywhere, with details about the economy and qualitative descriptions of the northern indigenous peoples (Anderson, 2011; Glavatskaya, Kljukina-Borovik, 2013). The archival situation is better for the comprehensive collections of Russian tax lists and parish records. Central State Archives in Moscow and St. Petersburg as well as most regional archives hold revizkie skazkie (tax revisions) from the 1720s to 1858, some of them with census-like individual level data about entire households. Covering the same period, as well as the later period from the 1850s until the revolution in 1917, the archives contain parish registers that give vital information about the baptized, the married, the buried and the divorced for most parts of the Russian Empire. Most localities are represented, and there are even records for religious minorities. This practice stopped after the Revolution, but paradoxically the new atheist rulers preserved the parish registers, while destroying the secular censuses. Few of the ministerial records in Eastern Europe and beyond have been transcribed for computer analysis, but fortunately new samples are now being brought forward in the Urals, Altai and Transylvania.

Some countries in-between the East and the West have archived historic population registers and/or a combination of censuses and church registers. Denmark, Iceland and Norway which used to make up the Danish Kingdom, demonstrate the common history of censuses and church registers and their employment for demographic analyses, with the full count 1801 nominative census as a highlight (Thorvaldsen, 2017). The centuries long series of longitudinal source material enabling the construction of detailed population registers from catechismal and other church records makes the situation different and superior in Sweden and Finland.

Efforts to create an overview of the size of the population at a given time and demographic developments occurred in parallel in different countries. This can be explained both ideologically and with practical
needs. Theological opposition traditionally blocked work on population statistics - man should not control God. The basis for such scepticism in the scripture can be questioned, even though there are biblical passages warning about numbering people for military purposes, see for instance Samuel II, chapter 24. Perhaps practical reasons were more important - the priests knew who would have to carry out the work. The bishops’ visitation reports show how they had to press the priests to make them keep the church books properly (Dyrvik, 1983). A further reason for resistance was political, as the proper construction of central statistics requires a strong central state power capable of requiring the necessary demographic data from local and regional leaders. The latter felt their position threatened by transferring control over their subjects to the central level, cf the gentry’s opposition towards holding censuses in the UK in the 18th century.

The ideological element of the early social sciences and population statistics is also important. The Cartesian research method recommends studying each component of the object under study carefully, just like the statistical aggregates were based on the priests’ information about each individual. Descartes prescribed how to compile anatomical details into an overall presentation of the research object, in analogue with the way vital statistics provide aggregate overviews of population developments. Thus, the methodological recommendations were accompanied by biological analogues about how the body parts relate to the whole man like the individual relates to society (Olson, 1993).

Practical concerns indeed triggered the collection of census data and vital statistics. Military needs obviously motivated the taking of male censuses. Data collection was part of the state’s intention to overview human resources prescribed by 18th century leading ideology, mercantilism. The practical goal was to promote business development in each country, for instance to increase revenues for military activities. Population growth became an instrument because the people were seen as an input factor in the economy as well as a criterion to measure whether the mercantilist policy was successful. Towards the end of the 18th century, officials would subordinate the economic objectives to the political objective of promoting the well-being of the population. “The aim was to form administrators or bureaucrats who should be concerned with the welfare aspect … including the moral and physical health of the population.” (Olson, 1993, p. 134). The nominative, full-count 1801 census and the ordinance of 1812 to the priests about church books with fixed headings were instruments for such reforms in the Danish Kingdom (Dyrvik, 1983).
Before the 20th century, Russian administrators had experience with aggregating census data primarily from the 1897 census, which covered the whole empire except Finland, and from censuses run in many cities in 1870s and 1880s. However, on the individual level, they had for nearly two centuries worked on the revizkie skazkie (census like tax revisions), which had been introduced by Peter the Great in 1718 – the 10th and last was organized in 1858. In seven of them, the whole population was in principle included, making it possible to study the household and family structure, as Maria Markova does based on the fifth revision for urban sites in and around St Petersburg in the present volume. The relationship variable in the revisions make it possible to see what persons in a household belonged to the same family and on this basis to classify some households as multi family. The revision also provides data on age, occupation and marital status of the persons listed, including women. Also, the historical demography group at Ural Federal University transcribe and research the Revizkie Skazki from 1852 and 1858 about the Obdorsk1 Samoeds (contemporary Nenets) and Ostiaks (contemporary Khanty). People were obliged to pay taxes for the men listed until the next revision including their serfs, thus also for those who had passed away in the meantime. This was the basis for Gogol’s masterpiece novel Dead Souls. The plan of Chichikov, Gogol’s hero-villain, was to buy up the “dead souls” at reduced rates, acquiring a list of fictitious but legally valid serfs. These he meant to use as security when investing in an estate and alive serfs. We can find these lists in archives all over the previous Russian Empire and scanned images are available on the Internet, for instance from Tobol’sk collection of the Tjumen’ State Archive.

A number of source-critical issues can be raised, both about general and detailed aspects of the church books, the population censuses and other demographic source materials. For instance, when comparing place and date of birth of individuals in the church book with census entries, they often do not match (Thorvaldsen, 1996a, p. 63-64). However, divergences can also be due to the fact that people used a church in the neighbouring parish because the distance there was shorter, or the parish borders were adjusted. The most serious problem may be that the oldest church books lack many entries. This applies mostly to the funeral lists, and to the least extent in the list of marriages. The church

---

1 The 19th century Obdorsk region included the Yamal Peninsula together with the Eastern slope of the Polar Urals, part of the Gyda peninsula and the gulfs of the Ob’ and Taz Rivers with its tributaries.
books recorded the ecclesiastical acts, not primarily aiming to overview population developments. For example in Norway, the roughness of the oldest church books is revealed by formulations such as “Earthed 5 Bodies” or “Buried a Fin wife”. The under-enumeration of the dead is partly due to the fact that people were most keen to pay for a church burial when important persons died (Høgset, 1990). The vocabulary in the church book is usually less precise than in the census, often indicating a rough status category like “cottar” rather than a more specific occupational function such as “fisherman”. Researchers have compensated for this by linking church book and population record data at the individual level (Engelsen, 1983). Cause of death is another type of imprecise or often missing information. Methods to study migration in spite of the many missing migration records in the Norwegian church books are presented in the article by Gunnar Thorvaldsen. Georg Fertig analyses how the rich collections of demographic data based on vital events records and address registers existing throughout Germany can be used in demographic research. Marius Eppel and Oana Sorescu-Iudean present a more specific picture of nominative sources for the study of the Transylvanian clergy.

**Historical databases**

In addition to church records, archives spread across the European continent contain large series of censuses and census-like sources. Many of these have been transcribed digitally and included in the comprehensive Mosaic database (Szołtysek and Gruber, 2016). In the present volume, Siegfried Gruber bases his study of marital fertility in Albania on one Mosaic component, the census taken there during the Austro-Hungarian occupation in 1918. Some countries in-between the East and the West have archived historic population registers in part based on the church registers. Sweden, Finland, Belgium and the Netherlands have paper-based population registers from the eighteenth or nineteenth century onwards. Three Swedish databases contain longitudinal population registers for five Swedish regions, including the capital city of Stockholm between 1876 and 1928. These databases offer details on migration, educational levels, etc. that are seldom found in historical records. Their richness makes digitization resource demanding – with traditional transcription methodology they may not cover Sweden during this or the next century. An impressive bibliography of recent publications based on the linked church records from Swedish regions,
covering individuals from 1680 to 1950, can be found at their website https://www.umu.se/en/centre-for-demographic-and-ageing-research/. With respect to transcription efficiency, the digital techniques developed in cooperation between historical demographers and computer scientists at the Autonomous University in Barcelona, which Joana Pujades Mora and her team present in their article, are most promising. To compensate for regional selectivity in the meantime, the Swedish censuses from 1860 to 1930 are being transcribed, encoded, linked and made available with national coverage. The Historical Sample of the Netherlands contains longitudinal data following a representative group of 78,000 individuals during the period 1812 to 1922. Peter Öri links household lists and vital events in order to build a historical population register, a preliminary version of which he uses here to study marriage patterns in the Hungarian village of Zsámbék from the 18th century to World War II. And Mark Gortfelder finds a partial effect of child mortality on fertility (on stopping, but not spacing of births) by employing the Family Registry of the Estonian Republic 1926–1949, which also contains retrospective data on persons born before 1926. Hilde Sommerseth and Evelien Walhout present a wider scope of the mortality issue in their study of death causes in the mid-Norwegian town Trondheim, which is also an important contribution to the SHIP project.

Elena Glavatskaya and her team members Dmitrii Bakharev, Alexander Bobitskiy, Elizaveta Zabolotnykh and Anastasia Vishnevskaya analyse the main marriage patterns and infant mortality in two articles about late 19th – early 20th century Ekaterinburg, for the latter theme also the rural surroundings. They used aggregate census data along with microdata extracted from several city parish registers (metricheskie knigi), transcribed into the database “Ural Population Project”. Both the majority Russian Orthodox Church parishes and several religious minorities: Catholics, Lutherans, Muslims and Jews are represented. In all, the database contains about 25000 baptism records, 15000 marriage records and 31000 burial records. The analyses show that marriage patterns depend on religious affiliation as well as social status and migration. They also found that while in the city the high infant mortality rate among the Orthodox majority had minimal changes, in Ural villages IMR decreased as a result from the efforts of the zemstvo – a self-government, elected, sub-provincial institution introduced in 1865.

The basis for keeping vital statistics was reformed with the introduction of civil registers, for instance in Russia after the 1917 Revolution and in Norway from 1905. The inappropriate use of such instru-
ments can be seen there during World War II when the Germans in February 1942 attempted to establish a central population register (Søbye, 1998). A longitudinal population register combines cross-sectional data with vital register data of the kind found in the church books. It is usually based on a census as a starting point and is updated with demographic events and migration. The primary aim of the Germans was to control passenger traffic in all municipalities, but the register could also be used to assemble statistics about births, marriages and deaths, something Statistics Norway stressed in its comments on the reform, which they were then able to delay until after the War by blaming quality problems. In connection with the census of 1950, all municipalities updated the population records with information from the population census manuscripts. This census is presently the first full-count, closed historical census in the world to be transcribed (Thorvaldsen, 2018).

If the population register contains all the variables that underpin vital statistics and aggregates from population censuses, it will significantly simplify the collection of basic data. The statistical agency can stop organizing questionnaire-based censuses and instead extract data from the population register by combining it with additional databases. These can be linked to the central population register at the individual level with a unique identity number as the linkage key. This is the reason why Denmark held its last form-based census in 1970. In addition to the registry method saving costs, its main advantage is that with little extra resources, the computer can construct statistical overviews of the population’s composition and the ongoing changes as often as desired. However, it would be desirable to check the registers by sending forms to at least a representative sample of the population. This also provides additional information about potentially missing variables such as housing, commuting and education. Education acquired abroad may be a variable where we will always need form-based updates.

Eighteen European countries used register data at least partly for their enumeration during the 2011 round of census taking, sometimes in combination with data from questionnaires. The purpose was to create census aggregates, although they also constructed nominative lists. In the future, researchers can access both aggregates and (anonymized) census-like micro-data, albeit extracted from registers. The advent of population registers signals the end of the difference between nations focusing on censuses and those focusing on vital register data. The methodology will bring forth new source materials that are more suited to follow over time population segments that are increasingly ge-
ographically mobile, both nationally and internationally, and will produce censuses and vital register data more economically. The downside may be that the population loses the control they had over the information, which they used to provide through the questionnaires.

The need for big data also in historical demographic research inspires the employment of the new information technology methods to rationalize data entry and the combination of sources into longitudinal databases with record linkage. The amounts of available source material are enormous, and even if much has been computerized, even greater collections of census manuscripts, ministerial registers etc are awaiting scanning, transcription and linkage. Just think of the quantities of metricheskie knigi (church books) from the vast Russian Empire during the two centuries before the Revolution. Only a fraction of this has been opened up with modern IT so far, and there are small prospects to get resources to transcribe much of this with traditional methods. Therefore, the developments described in Joana Pujades’ article with crowdsourcing, word spotting and the use of artificial intelligence - machine learning specifically – are most promising as methods to rationalize digitisation decisively. Fortunately, there are now similar methodological developments at several research centres around the world, and they are also being discussed at the international conferences for historical demography and social history.

References


Chapter 1.
TROUBLESOME RICHES: GENEALOGICAL DATA AS SOURCES FOR HISTORICAL DEMOGRAPHY IN GERMANY

Georg Fertig,

Introduction

Historical demography is not the kind of science that can live from models only; its practitioners are always in need of data. Regional variations within and beyond Europe tend to become more interesting to a discipline that has long ago overcome simple dichotomies pitching the west against the rest. In this perspective, Germany is a particularly interesting field for quite a few reasons. First, it is situated in between countries that have much stronger traditions of historical demography: south of Scandinavia, east of France and the Low Countries, mostly lagging behind in terms of scientific output also after what is published on the historical demography of other central European countries. Germany itself has a rather weak tradition of social science history, and most German historians shy away from quantitative methods, leaving a broad scope of historical interpretations open to others. Unfortunately, this applies also to the big questions of historical structures and changes that actually do call for quantitative testing. But then, a strong tradition of regional, contextual, more qualitative research is available—the only drawback is that it is transparent only to those who master the language. It is of course a truism, which applies in every country of the world, that only those who learn the language will ever be able to understand the field.
Second, Germany is a rather large and heterogeneous country which offers lots of possibilities for comparison. Other than most European countries, historic Germany consisted of many religiously (mostly) homogeneous territories, with either Catholic, or Calvinist, or Lutheran denominations. Until 1900, multiple legal systems governed the civil law within the recently unified Empire. Western regions of Germany were, and still are, part of the ‘blue banana’, the economic core of western Europe running from south west England to northern Italy, to which another core population belt running through Saxony towards southern Poland was adjoined (Klüsener, Zagheni 2014). Other regions, however, such as in the north east, were only sparsely populated. Ecotypes ranged from protoindustry to grazing to manorial export farming (and many more). The river Elbe separated systems of free and unfree peasantry and agricultural labour (or so the older literature had it). High variability also applies to more demographic topics such as household systems, demographic turnover, the epidemic transition or the fertility decline. Germany is thus a field where it is feasible to draw meaningful comparisons, deliberately holding some regional characteristics constant (say, religion and law), and studying others as variables of interest (say, ecotype and household system—or vice versa).

Computer Genealogy in Germany

This chapter will focus on a third reason why Germany is an attractive field for international historical demography: its strong tradition of data generating citizen science which is shockingly underexploited by the non-quantitative mainstream academic historians within Germany. I will first present an overview of the databases collected by CompGen, the German association for computer genealogy, and some other genealogical associations. I will skip commercial databases, since their accessibility for quantitative research depends on firm policies which have not yet been documented systematically. Rather, I will focus on a specific type of data, the local heritage books (Ortsfamilienbücher or Ortssippenbücher; literally: local family books or local lineage books), which exist both as printed books and as online databases.

Computer genealogy, as practiced by CompGen, is basically the collaboration of genealogists using modern information science tech-
niques for the purpose of producing open source data which is meant to benefit both non-academic genealogy and academic science. In principle, there exist three styles of genealogical work. The classical approach is descent oriented, often by family name, recently also using genetic techniques. A second approach is source type oriented, focusing on photographing and indexing one particular source (say, grave stones). The third approach is place oriented, bringing all sources for a particular town or area together. The resulting data collections are partly in the responsibility of CompGen, partly in that of other associations or projects. They can be categorized as memorial data, spatial data, and life course data. Moreover, computer genealogy generates software standards which may solve problems also in a demographic context. As to memorial data, it includes collections of grave stones, family advertisements, and private photographies which are typically not in the focus of historical demography.

In the field of spatial, or cross-sectional data, a first asset of CompGen is the historic gazetteer GOV (Geschichtliches Ortsverzeichnis). It covers geographic location, names (including historical names), and most importantly, the historically changing hierarchies of sub- and superordinate administrative and ecclesiastical entities. It has been estimated that about 85% of all settlement names for the time around 1900 (that is quite below units such as parish or town) within Germany are already covered in GOV (Zedlitz, Luttenberger 2014). GOV is continually expanded by a team that uses historical administrative gazetteers mostly from the late 19th and from the 20th centuries. Coverage before the 1840s is still limited.

<table>
<thead>
<tr>
<th>What</th>
<th>When</th>
<th>Where</th>
<th>How many</th>
<th>Whose</th>
</tr>
</thead>
<tbody>
<tr>
<td>GOV: Place and administrative unit identification database</td>
<td>19th c. to present</td>
<td>Germany, Western Europe, Eastern US, Australia</td>
<td>1.1 mil. places</td>
<td>CompGen</td>
</tr>
<tr>
<td>Address books</td>
<td>Mostly mid 19th to 20th c.</td>
<td>German localities</td>
<td>4.4 mil. entries, 469 books, 279 places</td>
<td>CompGen</td>
</tr>
<tr>
<td>Casualty lists</td>
<td>1914-18</td>
<td>Germany</td>
<td>8.5 mil. entries</td>
<td>CompGen</td>
</tr>
</tbody>
</table>

Source: compgen.de.

2 For a more thorough discussion contrasting different modes of genealogical research, see Timm 2016.
Addresses are a second, more small-scale type of spatial data. They are available in printed address books in considerable frequencies since the early or mid 19th century, and they contain not only names and street locations, but also occupations. CompGen teams have digitized several hundred books, covering considerable time spans for some larger places such as Berlin, Aachen, or Dresden. Spatial social inequalities within cities are indeed a currently hotly debated aspect particularly in German educational politics, while academic discussions of social topography have largely been put to rest by German social historians (for a recent overview, see Lenger, 2014).

Nominative lists of casualties published during wars are a third type. For the First World War, CompGen has organised a successful crowdsourcing project where all 8.5 million entries—soldiers who were killed in action, deceased otherwise, wounded, missing, or captured—were digitized. Beyond the date of casualty, the lists typically also contain the birth places. Thus, they can be interpreted in terms of micro-regional consequences of war, beyond offering a statistical base for more conventional questions of military history. Regional health, family, or employment consequences of war casualties remain further paths of enquiry to be studied. A comparable project for casualties in the army of Austria-Hungary is ongoing.

The other and more typical kind of data genealogists collect and refer to are life courses and kinship. The production cycle of such data begins with parish registers, which are indexed or transcribed, and finally interlinked by genealogists to form either family-wide lists of descendants or ancestors, or parish-wide local heritage registers. Earlier projects in academic historical demography have organised the entire process using the working time of doctoral students, which is a costly decision since it can be estimated that even for a small parish, one hundred years of manual family reconstitution will require about a year of full time labour.

Parish registers exist for practically all places. Their availability has been largely enhanced when the Genealogical Society of Utah started to microfilm them during the 1950s. These films have been available in

---

4 The largest group in Germany was directed by the late Walter Rödel. For a bibliography of the studies he directed, see https://www.regionalgeschichte.net/fileadmin/Superportal/Bibliothek/Autoren/HistDemographAK/VerzHist-DemArbeiten_AKMz.pdf.
Latter Day Saints ‘Family Research Centers’ scattered across Germany, and the world; a practice now replaced by online access (partly limited to Family Research Centers). Geographically, most of these films are from the westernmost parts of Germany; they also tend to cover more Protestant than Catholic parishes. In 2008, the Catholic Church has ended its cooperation with the Mormon genealogical societies since genealogy has a strong theological function in Mormon belief, which Catholics do not share. Protestant churches are currently building a system for the paid online distribution of scanned parish registers, Archion, while the archival network Icarus offers a free system, Matricula, for both Catholic and Protestant dioceses.

Demographers cannot work from scanned films, they need transcribed or indexed registers. Fortunately, indexing is a genealogical megatrend. Transcriptions, or indexes, of registers exist within the FamilySearch system and are accessible for free online. Their quality and completeness is, however, questionable and many genealogists suggest to use this material only as a starting point. Ancestry.com also is creating a paid system, where more indexes can be found than are available in FamilySearch. Other transcriptions have been organised by genealogical societies, most importantly in the Rhine land region, and some have been published in print. There is no central pool of transcription or index data; CompGen does not currently own any of them, but keeps track of those that exist within its Wiki system.5

The next step in the production of genealogical data is to link entries. This is a crucial step also for historical demography, where family reconstitution is a core practice. If you want to study life expectancies, it is quite a plausible step to connect the birth dates and death dates of people and then to apply life table methods; the only alternative would be to use projection and simulation techniques such as Populate or Camsim which are technically quite challenging. The same goes for the study of fertility or kinship networks–linked data, not original entries, are where demography starts to be interesting. Linkage can in principle be done entirely through algorithms, with the help of algorithms, or manually (Christen, 2012, Schraagen, 2014). Grounds for linking two or more entries can either be made transparent and reversible, or not. Genealogists typically do not use algorithms, and they also do not write down rules that say under which conditions they decide to link or not.

5 http://wiki-de.genealogy.net/Kategorie:Verkartung
to link entries. They also do not document cases where two or more diverging linkage decisions could be justifiable. To be fair, even advanced projects in historical demography have not done so, either. In countries where patronyms are a standard part of naming conventions, such as Russia and Iceland (but also the Frisian region within Germany), linkage decisions are easier to make; in others, including most of Germany, much less so. Enhancing the transparency and revisability of linkage decisions is a core aim of CompGen longterm software development project, Gedbas4all, a graph data base. The standard format of genealogical computing, Gedcom, rests upon the concepts of individuals, families, and events. The longitudinal data standard of historical demography as far as there is one, the Intermediate Data Structure (IDS), is not much different; it rests upon the concepts of individuals, contexts, and events. Gedbas4all brings in the additional concept of assertion, opening the possibility of a more transparent and source oriented data organization. Developing Gedbas4all could enrich international historical demography strongly.

Linkage data are the main product of genealogical activity. They are produced in the form of descendant lists or more frequently ancestor lists. CompGen hosts an open database system, Gedbas, where genealogies of both types can be uploaded and searched, with 18 million persons. Analysing this kind of material meets two challenges: first, there are many duplicate entries in Gedbas, and second, a population out of which this huge sample has been drawn is impossible to define. In other words, any analysis must explicitly address the question of representativity–which may, however, be doable.

Local heritage books (Ortsfamilienbücher, OFB) are basically family reconstitutions for entire places. They do not suffer from the same weaknesses as genealogical lists. Insofar they cover the entire population of a given place, representativeness is not a real issue. The population at risk should cover every person that was living in the given place during the period of interest; it does so with one qualification: as long as no census type data are integrated, the information we have will refer to every person that was born, married, had children or died in the place. Hence, it is necessary to define clearly for what time period we consider a person as present, and particularly how to deal with truncated and censored data. CompGen keeps track of about 3.700 printed OFB, and offers a database of Online OFB which covers about 700 places and 10 million individuals. As is the case with genealogical lists, the books always have authors, who should be asked personally for any evaluation
of their data. The rest of this chapter will be devoted to discussing this particular type of data source.

<table>
<thead>
<tr>
<th>What</th>
<th>When</th>
<th>Where</th>
<th>How many</th>
<th>Whose</th>
</tr>
</thead>
<tbody>
<tr>
<td>Parish register indexes</td>
<td>16&lt;sup&gt;th&lt;/sup&gt; to 19&lt;sup&gt;th&lt;/sup&gt; c.</td>
<td>Multiple; most frequent in Rhineland</td>
<td>Unknown (about 1.000 places?)</td>
<td>Regional societies, FamilySearch, Ancestry, individual genealogists</td>
</tr>
<tr>
<td>Gedbas ancestry lists</td>
<td>Unlimited</td>
<td>Unlimited</td>
<td>18.0 mil. entries</td>
<td>Authors</td>
</tr>
<tr>
<td>Online-OFB</td>
<td>16&lt;sup&gt;th&lt;/sup&gt; to 19&lt;sup&gt;th&lt;/sup&gt; c.</td>
<td>Multiple places</td>
<td>9.6 mil. entries (for 700 places)</td>
<td>Authors</td>
</tr>
<tr>
<td>Printed OFB</td>
<td>16&lt;sup&gt;th&lt;/sup&gt; to 19&lt;sup&gt;th&lt;/sup&gt; c.</td>
<td>Multiple places</td>
<td>3.700 places</td>
<td>Authors</td>
</tr>
</tbody>
</table>

*Source: compgen.de.*

**Local Heritage Books as a source for historical demography**

Typical Local Heritage Books give a couple of thousand entries, one for each family—that is, a married man together with his wife and children. The entries give birth or baptism, marriage, and death or burial dates and places, as well as additional information and transcripts from parish registers. They are numbered by family entry; numbers point both backwards to the families of origin for husband and wife, and forward to families of procreation for children who married and for those husbands or wives who remarried. Typography varies considerably, as do the organisation and source content of the books.

The OFB are not a new format. First local genealogies of this type were published during the 19<sup>th</sup> and early 20<sup>th</sup> century. After 1933, the term *Sippenbuch*—book of kinship groups or lineages—was established, as were certain conventions for layout and data presentation. Production of *Sippenbücher* was encouraged by the Nazi state, as was popular genealogy in general (Knodel, 1975, Weiss, Münchow 1998, Pinwinkler, 2014, p. 36–41). As with many aspects of the regime, there existed at least two bitterly competing variations of state supported genealogical research: one based in the *Reichssippenamt* (part of the interior ministry), aiming at documenting the entire and particularly urban population in
one giant card index, and another one based in the Reichsnährstand, a corporate organisation of agricultural producers, which focussed on the concepts of local peasant families and their land-family bond. As a percentage, the about 50 books published under the NS regime—i.e., under the Reichsnährstand—are not numerous in comparison to the about 200 that appeared between 1945 and 1975 (when John Knodel analysed this type of source material under a demographic perspective), let alone the about 2,800 that have been compiled since. Still, it is a question what to make of these origins. At the time, the Sippenbücher had a dual function: administrative and ideological. The administrative function—documenting non-Jewish ancestry—had become obsolete in 1945. The ideological function particularly for the Reichsnährstand was to promote a sense of popular community for plain, rural, non-noble people, based on descent,

---

1 Translation: 3728 Hans Michel WÜST (from 3725), Protestant, baptized in Göbrichen 20/8/1651, citizen and blacksmith, 1674–1698 teacher, 1702 member of the local court of justice, buried in Göbrichen 14/1/1733, called ‘the old blacksmith’, married 15/4/1673 Barbara HOFFSEEß (from 1449), born in Öschelbronn 23/12/1649, Protestant, elected midwife in March 1695: ‘1695 ... in March, for the women who are advanced in pregnancy, and are blessed with a fruit of womb, elected by the wives of me, the pastor, of the mayor, of the members of the local court of justice and of the council, and by all the wives here, Barbara, the wife of Hans Michel WÜST here, currently teacher. May God give her his mercy and blessing that everything will work out well’ (entry is very light!); buried in Göbrichen 17/1/1731. 3 children, born and baptized in Göbrichen (except the third one): 1. Hanss Jacob born 25/1/1674 (marriage 3730), 2. Barbara born 31/3/1676 (marriage 755), 3. Hanss Michael baptized in Pforzheim while refugee, godfather 31/10/1699 ‘the smith’s youngest son, unmarried’ (marriage 3731).
and excluding those who did not belong. It is not quite self-evident that this ideological function was ended in 1945 as well. Publication of local heritage books continued after 1945, with the term ‘Sippenbuch’ resurfacing in the mid 1950s. It may hence be worthwhile to consider more deeply if there are any implicit biases of this type of material, even if substantially there is scarcely any difference to the family reconstitutions in the tradition of Henry (1967).

A first limitation is that although potentially OFB can contain all kinds of life course related data at least as notes, their core is about three events documented in the parish registers: birth, marriage, and death. Even within parish registers, there is more serial information than that: persons are often identified using locational and occupational characteristics; godparentship relations are given at all baptism entries; causes of death are often given systematically. Moreover, other person-level sources such as tax lists, census lists, petitions, hypothecary and cadastral registers, inventories and many more exist in many places, and may or may not be integrated in any given OFB. All of this is relevant for some types of questions that historical demographers, historians of the family, social and economic historians might ask. Privileging just birth, marriage, and death certainly gives a reduced view of the social fabric and individual life course. This is, however, a problem shared by much of historical demography rather than that it is embedded in this particular type of genealogies. Conversely, the tendency of genealogical work is to include as much person-related information as possible. In other words, modern genealogy is a combination of prosopography, family and network history rather than pure descent research.

A second limitation is that it might be that some parts of the population are seen as more worthy of inclusion than others. To give an example, Adolf Clarenbach, a Lutheran pastor in the parish of Borgeln as well as a highly respected member of the historical profession, transcribed, interlinked, sorted and published thousands of biographical entries from the Borgeln parish registers. His publications are organised by houses, not by families, starting with the farmholders (in what he published between 1939 and, post mortem, 1954), while data on non-farming house dwellers were only published much later by co-authors (the earliest is Clarenbach 1939, the latest Clarenbach and Rudack 1984). His selective approach went even further, as he explained in 1938 in an article for a publication widely read by pastors: card files of vagrants should be marked ‘V’ and sent to a central card file collection on antisocial elements (Clarenbach, 1938). This was planned at
the time, but not carried out systematically, while 1938 indeed marks a year when ‘antisocials’ were massively put into concentration camps. Indeed, only two cards marked ‘V’ remain in the card file collection he created for Borgeln, both for beggar women who died in Borgeln (LKA Bielefeld, Familienkartei Borgeln, cards Buschkemper and Anonyma). It is quite likely that many more travellers passed by that village over the centuries, occasionally dying or giving birth. Privileging the farm holders and established local families, and ignoring or deleting the less stable parts of the population, is thus certainly a possible source of bias.

A less politically charged aspect of inclusiveness bias regards those children who did not survive their birth or early life. If a genealogist is only interested in those who had descendants themselves, the stillborn or early deaths seem to be less important. Again, the tendency of genealogical research goes clearly against both types of exclusion.

A third potential limitation of genealogical work arises from the fact that there are actually two aspects of it: making source material accessible, and inferring links. For some genealogists, transcribing is for the general good, while inferring is something more private about which everybody has to decide for themselves. Hence, they publish the family entries but not the relations between them, so that their OFB do not contain forward and backward linking.

Table 1

| Titles documented in Genwiki       | 3,606 |
| Books collected at Ludwigshafen library | 3,700 |
| Online-OFB                        | 638   |
| Local data sets in Leipzig database | 8,243 |
| Titles in Leipzig database         | 4,652 |
| ‘Family books’, ‘OSB’, ‘OFB’ in Leipzig database | 1,252 |
| ‘Typical OFB’ in Leipzig database  | 438   |

Sources: compgen.de and Leipzig database.

It is far from clear how many books exist (see Table 1), and where to draw the boundaries between the typical books and comparable material. A research project guided by Volkmar Weiss at the State Archive of Leipzig has, until the late 1990s, collected titles of OFB in a very broad sense, also including unpublished material and projects that never got beyond
the planning stage. The project database contains several thousand title entries until 1997; however, only a smaller part of these titles were acquired by the Leipzig archive, evaluated, and considered a typical OFB.\(^6\)

In the Leipzig project, the evaluation revolved around the questions of whether the entire population was covered, if family entries were forward and backward interlinked, if additional sources beyond the parish registers were added, and if occupational and property information is given. Table 2 and 3 are based on a sub-sample of Weiss’s database, including only those ‘typical’ OFB. I make a distinction between those books printed under National Socialism, those that were available during the peak period of historical demography in the 1970s, when John Knodel started to work from this type of sources, and the many books that were published later. I also added two later sub-samples that I evaluated using Weiss’s criteria: a set of 50 books published by the regional association for genealogy in what is called middle Germany (i.e., Saxony, Thuringia, and Saxony-Anhalt; the *Arbeitsgemeinschaft Mitteldeutsche Familienforschung*) after 2000, and a set of 41 databases from the same area that are published on CompGen’s website online-ofb.de. From table 2, we see a slight tendency to include earlier registers, and also to publish books that do not have the 19\(^{th}\) century included. It is, however, table 3 that tells us most in terms of the quality issues discussed above.

### Table 2

Temporal coverage of OFB, by period and sample

<table>
<thead>
<tr>
<th></th>
<th>N</th>
<th>1699 and earlier</th>
<th>1799 and earlier</th>
<th>1800-70</th>
<th>1870-1900s</th>
</tr>
</thead>
<tbody>
<tr>
<td>NS Period</td>
<td>31</td>
<td>0.58</td>
<td>0.94</td>
<td>0.94</td>
<td>0.15</td>
</tr>
<tr>
<td>Pre Knodel</td>
<td>64</td>
<td>0.66</td>
<td>0.91</td>
<td>0.86</td>
<td>0.81</td>
</tr>
<tr>
<td>Pre 2000</td>
<td>339</td>
<td>0.54</td>
<td>0.67</td>
<td>0.50</td>
<td>0.27</td>
</tr>
<tr>
<td>Central German printed</td>
<td>50</td>
<td>0.94</td>
<td>0.70</td>
<td>0.32</td>
<td>0.97</td>
</tr>
<tr>
<td>Central German online</td>
<td>41</td>
<td>0.90</td>
<td>0.78</td>
<td>0.46</td>
<td>0.15</td>
</tr>
<tr>
<td>All</td>
<td>525</td>
<td>0.62</td>
<td>0.72</td>
<td>0.55</td>
<td>0.34</td>
</tr>
</tbody>
</table>

Sources: compgen.de and Leipzig database.

\(^6\) I owe thanks both to Volkmar Weiss and to his successor Thekla Kluttig for providing me with (slightly diverging) copies of the project database. A printed version of the database is Weiss et al. 1997.
In terms of the first type of limitations discussed, the question whether data beyond birth, marriage, and death from parish registers are included, Table 3 shows ambivalent results. The books published under the *Reichsnährstand* typically were not based on any sources beyond the parish registers, but still they consistently include some occupational and property data. Given the original focus on farmers, this is not surprising. In later printed books, genealogists started to work on additional sources in about a third of the cases, while occupational and property data remained present in most but not all. Recent online databases are different: they contain less occupational data and scarcely any source material beyond parish registers. This may, however, be a technical effect since non parish register data may be held in other database tables than those for the individuals and the couples which form the backbone of the online OFB. Moreover, the function of the online publication is often to provide fellow genealogists with a provisional insight into an ongoing project.

The second type of limitation is not to cover the full population. This was not only a problem, it obviously still is, and particularly was between 1975 and 2000. Even in recent printed works it is not rare that stillborn or those without local residence are left out. The third problem, non-interlinked family entries, seems on the verge of

### Table 3

<table>
<thead>
<tr>
<th>Period</th>
<th>N</th>
<th>Non Parish Register Sources</th>
<th>Occupational and Property Data</th>
<th>Coverage</th>
<th>Interlinkage</th>
</tr>
</thead>
<tbody>
<tr>
<td>NS Period</td>
<td>31</td>
<td>0.03</td>
<td>1.00</td>
<td>0.81</td>
<td>0.74</td>
</tr>
<tr>
<td>Pre Knodel</td>
<td>64</td>
<td>0.28</td>
<td>0.98</td>
<td>0.72</td>
<td>0.80</td>
</tr>
<tr>
<td>Pre 2000</td>
<td>339</td>
<td>0.35</td>
<td>0.74</td>
<td>0.46</td>
<td>0.69</td>
</tr>
<tr>
<td>Central German printed</td>
<td>50</td>
<td>0.36</td>
<td>0.88</td>
<td>0.88</td>
<td>0.98</td>
</tr>
<tr>
<td>Central German online</td>
<td>41</td>
<td>0.02</td>
<td>0.34</td>
<td>0.95</td>
<td>1.00</td>
</tr>
<tr>
<td>All</td>
<td>525</td>
<td>0.30</td>
<td>0.77</td>
<td>0.59</td>
<td>0.76</td>
</tr>
</tbody>
</table>

Sources: compgen.de and Leipzig database.
disappearance. This is certainly a consequence of the shift towards computer genealogy; genealogical database systems organise families in an interlinked way.

The issue of quality control is far from settled with these observations. In the demographic literature, most discussions of the quality of family reconstitutions (Knodel, 1975, Medick, 1996, p. 618, Wrigley et al. 1997, p. 73–118; 574–577) refer to four questions: underregistration in the original sources, the impact of migration on the share of statistically useful families (Henry’s mariages ouverts and mariages fermées), biases introduced by ignoring certain subgroups, and linkage decisions that are logically impossible. Typical methods for testing include manual reconstitution for a sample of entries in order to compare it with the genealogy, and comparing some demographic parameters calculated from the genealogy to the same parameters from other data sets constructed in academic research. Both methods imply that linkage by academic researchers is essentially more reliable than linkage by genealogists. This is of course an extremely self-confident assumption which runs counter to the common observation that errors of judgement do happen among academics at some unknown, but certainly non-zero rate.

**Conclusion and agenda for further research**

In order to harness the massive amount of knowledge embedded in these books, collecting and documenting them is a first step. This seems to be something where libraries, archives, and authors hold the relevant information; genealogical associations have a key role in bringing it together.

A second one is to document the range and quality of the underlying sources: identifying the start and end years of the registers as well as gaps in between. Parish registers were written by pastors; their quality certainly correlates with their individual biographies. For most of Germany, the names, tenures and biographies of pastors are well documented. Documentation of events in the sources will plausibly vary by writer. Some potential biases are easy to identify using simple methods (gender ratio, infant mortality, illegitimacy). A connected issue are underlying behavioural patterns that make it difficult to link entries together. Different intensities of migration lead to different proportions of mariages fermées within the reconstituted population. Naming conventions can influence the recognizabilities of persons, e. g. where husbands take their wives’ names if she is a farm successor, or the usage of
patronymics as a best case scenario. Frequencies of first and last names matter too: it is much easier to reconstitute families where there are many different names, than where half of the families are named Mueller, and all women go by Marie.

A third field is assessing the quality of the record linkage itself. A basic assumption of the Cambridge Groups’ family reconstitution study was that if the underlying sources are reasonably complete, reconstitutions are good enough for demographic analysis. This pragmatic approach can certainly be justified. Knodel, in contrast, explicitly addressed the issue of underlinkage. Birth intervals are obviously too long if some births are linked to the wrong family. But overlinkage can distort the results, too. Genealogists tend to be more satisfied when they find some linkage than when they do not. Simulation of record linkage using different algorithms and comparing these to manual results would help address this problem.

Taken together, this chapter was aimed at introducing historical demographers to the growing pool of local life course and kinship data that awaits analysis in Germany, one of Europe’s most differentiated and central fields for study. Collaboration both internationally and between academic and citizen science (Fertig, 2016) is the ‘future of historical demography’ in Germany.

References


Chapter 2.
THE BAIX LLOBREGAT (BALL) DEMOGRAPHIC DATABASE, BETWEEN HISTORICAL DEMOGRAPHY AND COMPUTER VISION (NINETEENTH – TWENTIETH CENTURIES)¹

Joana Maria Pujades-More,
Alícia Fornés,
Josep Lladós,
Gabriel Brea-Martínez,
Miquel Valls-Fígols

Introduction

Historical studies always need empirical evidence, and the enhancement of historical research in the «Big Data Revolution» (Ruggles, 2014), inspired additional need for large quantities of high quality data. Historical demography has always consistently used considerable amounts of data, depending on the technical developments of the moment (Billari and Zagheni, 2017). Many historical demographic databases have a long tradition, such as the Historical Samples of the Netherlands (HSN), the China Multigenerational Panel Dataset (CMGPD-LN); the North Atlantic Population Project (NAPP), the Demographic Database at Umeå University (CEDAR), the Research Program in Historical Demography at Montreal University (PRDH), the Scanian Economic Demographic Database at Lund University (SEDD), among others (Cf Ruggles et al., 2011; Lee and Campbell, 2010; Mandemakers, 2002; Edvinsson, 2000; Dillon

¹ This research has been supported by the project ‘NETWORKS: Technology and citizen innovation for building historical social networks to understand the demographic past’ (http://dag.cvc.uab.es/xarxes/) funded by Recercaixa program (2017–2019) (PI Joana-Maria Pujadas-Mora and Alicia Fornés). The team includes researchers from the Universitat Autònoma de Barcelona, the Center for Demographic Studies (CED), and the Computer Vision Center (CVC).
et al., 2018; Bengtsson et al. 2012). These kind of data is particularly important to understand the demographic past (fertility, family formation, health, social stratification, social inequality, etc.) and can help to understand the present and to forecast the future. However, the digitization of historical sources is still time-consuming with high staff costs. Therefore, the recent conjunction of Historical Demography and the Computer Sciences promises to shorten the construction time for historical individual level databases and allows the building of bigger and more informative databases (Pujadas-Mora et al., 2016; Hall et al. 2000).

The **Baix Llobregat (BALL) Demographic Database** is an ongoing database project containing individual census data from the Catalan region of **Baix Llobregat** (Spain) during the nineteenth and twentieth centuries. The BALL Database is built within the project ‘**NETWORKS: Technology and citizen innovation for building historical social networks to understand the demographic past**’ directed by Alicia Fornés from the Center for Computer Vision and Joana Maria Pujadas-Mora from the Center for Demographic Studies, both at the Universitat Autònoma de Barcelona, funded by the Recercaixa program (2017–2019). Its webpage is http://dag.cvc.uab.es/xarxes/. The aim of the project is to develop technologies facilitating massive digitalization of demographic sources, and more specifically the **padrones** (local censuses), in order to reconstruct historical ‘social’ networks employing computer vision technology. Such virtual networks can be created thanks to the linkage of nominative records compiled in the local censuses across time and space. Thus, digitized versions of individual and family lifespans are established, and individuals and families can be located spatially.

---

2 For more information about each of those databases: The Historical Samples of the Netherlands (HSN) (https://socialhistory.org/en/hsn/index); China Multigenerational Panel Dataset (CMGPD-LN) (https://www.icpsr.umich.edu/icpsrweb/DSDR/studies/27063); North Atlantic Population Project (NAPP) (https://www.nappdata.org/napp/); The Demographic Database at Umea University (DDBPOPUM)(http://www.cedar.umu.se/english/ddb/databases/popum/); The research Program in Historical Demography at Montreal University (PRDH) (https://www.prdh-igd.com/); The Scanian Economic Demographic Database at Lund University (SEDD) (https://www.ed.lu.se/databases/sedd). This list can be extended with some novel impressive databases as the Historical Population Register (HPR) from the Norwegian Historical Data Centre (http://www.rhd.uit.no/nhdc/hpr.html) or the ones which are created for the International Demographic Unitat the Ural Ferderal University (Russia) (https://urfu.ru/en/research/international-researchcollaboration/international-research-laboratories/international-demographic-unit/). For more information on historical longitudinal databases please see: https://www.ehps-net.eu/databases.
Historical and Geographic context of the Baix Llobregat

Catalonia was one of the first places in Southern Europe to industrialize (Brea-Martínez and Pujadas-Mora, 2018; Martínez-Galarraga and Prat, 2016). The Baix Llobregat region played an important role because since the second half of the nineteenth century (Figure 1), the flourishing Barcelonese cotton industry moved towards the lower parts of the Llobregat River and its delta in search of water for the demanding production (Nadal, 1992). Then, the region changed from an agricultural to a wider occupational and social structure.

The industrial growth first featured textile production and afterwards metallurgy (Carbonell i Porro, 1995). However, in the last decades of the nineteenth century the agricultural crisis in Europe affected prices and incomes. In Catalonia the *phylloxera* vineyard infection worsened the crisis (Garrabou et al., 1991). Although the agrarian crisis also disturbed the economy of the *Llobregat* riverside, its effect was limited, because the region disposed of highly fertile soils and had faced an incipient commercialization of its agricultural production since the eighteenth century. During the nineteenth century, the city of Barcelona used to be provided with fruits and other orchard products from the *Llobregat* and since the start of the twentieth century they exported agricultural commodities abroad (Tribó, 1989).

Thus, the *Baix Llobregat* region represents a highly interesting case study. On the one hand, the region played an important role in Catalan industrialization so that the BALL Database includes observations of a society facing all stages in the industrial revolution: the end of the pre-industrial era and the take-off, implementation, expansion and consoli-
On the other hand, the successful agrarian commercialization in the *Baix Llobregat* since the eighteenth century adds extra elements to the classical model of modern economic and industrial growth, which implies an increasing secondary sector and a declining primary sector (Kuznets and Murphy, 1966; Nadal, 1975; Carreras, 1990; Martínez-Galarraga and Prat, 2016; Brea-Martínez and Pujadas-Mora, 2018).

The BALL Database, therefore, offers analytical opportunities regarding the relation between industrialization and demography. For instance, a classic theory argues that declining fertility was caused by modernization and industrialization and the emergence of nuclear families (Franck and Galor, 2015; Freedman, 1979; Cherlin, 2001). However, our area of study experienced the earlier fertility decline within Catalonia - together with France the earliest fertility declines in Western Europe (Cabré, 1999; Weir, 1993; Coale and Watkins, 1986). Moreover, the *Baix Llobregat* region was traditionally featured by the strong presence of stem families. Nowadays, this area is still shaped by strong family ties (Reher, 2004; Fauve-Chamoux, 2009; Borderías and Ferrer, 2017; Esping-Andersen, 1999). Thus, the BALL Database provides an interesting and complete socioeconomic and socialhistorical “lab” for studying the demographic, familial and individual responses to a transforming world.

**Sources: The *Padrones* (Local censuses)**

Local censuses (*Padrones* in Spanish) were taken regularly in Spain since the nineteenth century. They were a result of administrative centralization and the efforts of the liberal state to increase population and wealth (Porter, 1995; Wolf, 1989). Institutionalization of state statistics was a response to a desire for quantifiable material within a framework of epistemological development of scientific objectiveness and impersonal knowledge (Porter, 1995). This was a common although not simultaneous process in nineteenth century Europe. Spanish local censuses were compulsory after 1823 (García Pérez, 2007; Reher and Valero-Lobo, 1995). Thus, they were carried out before the first modern national census (1857) or the definitive implementation of the civil register (1871).

The state progressively implemented local censuses throughout the country. Several decrees informed municipal officials of the obligation to take them. Thus, the Royal Decree of March 14th 1857
demanded all the registers to be nominative and simultaneous. Only after the enactment of the Municipal Law of August 20th 1870 was the taking of local censuses fixed to 5-year intervals. Finally, in the Municipal Statute of March 8th 1924, the formats of local censuses were standardized with respect to the type of variables that should be recorded (García Pérez, 2007; García Ruipérez, 2012). In this way, local censuses showed the sociodemographic features of each inhabitant in a particular household in an urban center or in the countryside. For each person, the register included first names and surnames, age or birth date, civil status and occupation and the family or working relationship with the household head. In some periods, the information is also available about the individuals’ literacy and income. Most importantly, the local censuses contain the only data preserved at the individual level in Spain since national census manuscripts used to be destroyed once the population number was estimated and the main variables aggregated.

Table 1

Number and time range of local censuses in BALL Demographic Database, by municipality.

<table>
<thead>
<tr>
<th>Municipality</th>
<th>Number of Local Censuses</th>
<th>Period</th>
<th>Population 1857</th>
<th>Population 1950</th>
<th>Individual records</th>
</tr>
</thead>
<tbody>
<tr>
<td>Begues</td>
<td>10</td>
<td>1854/1928</td>
<td>799</td>
<td>968</td>
<td>9,658</td>
</tr>
<tr>
<td>Castellví de Rosanes</td>
<td>7</td>
<td>1857/1950</td>
<td>323</td>
<td>268</td>
<td>2,100</td>
</tr>
<tr>
<td>Collbató</td>
<td>32</td>
<td>1852/1950</td>
<td>865</td>
<td>416</td>
<td>12,836</td>
</tr>
<tr>
<td>Corbera de Llobregat</td>
<td>18</td>
<td>1857/1950</td>
<td>885</td>
<td>1,397</td>
<td>15,124</td>
</tr>
<tr>
<td>El Papiol</td>
<td>14</td>
<td>1875/1950</td>
<td>1,100</td>
<td>1,159</td>
<td>14,256</td>
</tr>
<tr>
<td>Molins de Rei</td>
<td>14</td>
<td>1852/1955</td>
<td>3,002</td>
<td>8,024</td>
<td>55,566</td>
</tr>
<tr>
<td>Sant Feliu de Llobregat</td>
<td>20</td>
<td>1828/1955</td>
<td>2,484</td>
<td>7,327</td>
<td>83,528</td>
</tr>
<tr>
<td>Santa Coloma de Cervelló</td>
<td>7</td>
<td>1900/1950</td>
<td>211</td>
<td>1,227</td>
<td>7,537</td>
</tr>
<tr>
<td>Torrelles de Llobregat</td>
<td>30</td>
<td>1842/1950</td>
<td>496</td>
<td>732</td>
<td>21,062</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td><strong>152</strong></td>
<td><strong>1828/1950</strong></td>
<td><strong>10,165</strong></td>
<td><strong>21,518</strong></td>
<td><strong>221,167</strong></td>
</tr>
</tbody>
</table>

Source: Authors’ own elaboration (BALL Demographic Database, version July 2018).

The BALL Demographic Database until now contains nine municipalities (Figure 2): Begues, Castellví de Rosanes, Collbató, Corbera de Llobregat, El Papiol, Molins de Rei, Sant Feliu de Llobregat, Santa Coloma de Cervelló and Torrelles de Llobregat, altogether representing a total number of 152 local censuses comprising
the period 1828 to 1950 (Table 1). The medium and small size of the communities registered in the BALL local censuses is far from an inconvenience for the study of networks. It is in reality a strong feature, making it easier to assess the mechanisms that individuals and households used at the micro level in order to adapt to and interact with socioeconomic changes at the macro level (Coleman, 1986).

Figure 2. Map of BALL Demographic Database (Baix Llobregat county).
Source: Authors’ own elaboration (BALL Demographic Database)

The pre-1880 local censuses were taken before the standardization of census formats. They are basically nominative lists of the inhabitants gathered in the census (Figure 3). For this reason, the number of variables registered usually includes only first names, surnames, ages and civil statuses.
From 1881 the number of variables registered in local censuses increased to around 15 in each census (Figure 4). In this way, the recorded variables were first names, two surnames, occupations, marital status, age and/or birthdate, address (street and house number), birthplace as well as time of residence. From 1906 onwards, other variables are the relationship to the household head, literacy (reading and/or writing), and fiscal contributions (see Annex 1). Nevertheless, in spite of the rich detail in these sources, the local censuses masculinize the data, leaving
out many women’s occupations. Female work was usually underrecorded or registered with labels such as *su sexo* (tasks of her sex/gender) or *sus labores* (her own tasks) (Borderías, 2012). The only exception to this female invisibility in the local censuses is found in those from 1936 during the Spanish II Republic (1931–1939), which mostly registered women’s occupations as well as which factories or enterprises employed the wage earners.
Type of database

As mentioned above, the aim of the project is to create historical ‘social’ networks using census data. Therefore, we use a social network model to transcribe the nominative census records. This data model also allows using the powerful techniques of graph based data analytics for querying the database. This representation is scalable, allowing not only to integrate and link data from other censuses but also other demographic sources like birth, marriage, death records. Figure 5 illustrates how the census data is structured according to a social network or graph representation, with the time dimension as a stack of linked graphs. Each individual graph is a static representation of the population at time $t$ (a particular census) where the nodes are observations of people. The corresponding households where they were registered and the graph edges represent relations (genealogic or other relevant affinities deduced from the source documents like occupations, household neighbours etc). In dynamic, time varying graphs, the life courses of individuals are constructed by linking the corresponding record of their observation at time $t_n$ with the observation at time $t_{n+1}$.
Technically, we implemented this in a relational database according to the data model in figure 5. Current database functions allow us to implement graph structures in a native format, while providing query methods to analyse the graphs (community detection, record linkage, centrality-based node detection etc). As future work, we plan to migrate to such a format. The proposed design is inspired by the Intermediate Data Structure (IDS), a standard proposed by the EHPS Network (Alter & Mandemakers, 2014). This should ensure the interoperability among data sets, and be the basis for the implementation of data analytics.

The design proposed in Figure 6 has the central entity Person, containing the information that is permanent over time. Each person has different Observations, belonging to the corresponding censuses that have been analyzed. Like in the census, observations are grouped in Households. A census record is considered an Event, which is a class with different subclasses (according to the different event types: census record, marriage record, birth record or death record). This structure allows future users to reconstruct the life course of a person based on the extraction of his/her events.
Figure 6. Data Model of the database. Source: Authors' own elaboration (in unified modeling language)
This design is flexible enough to incorporate other event types, not necessarily related to demographic data, but relevant in the life course. Finally, the database architecture contains entities for storing the images from which the information is extracted and the image segments that contain the relevant information as cropped by the computer vision algorithms.

The transcribed data are transferred into a digital format compatible with different formats (Excel, SPSS, STATA, R, etc), downloadable from a crowdsourcing platform used also for collaborative transcription of the original sources.

**Transcription:**

**Crowdsourcing platform, videogame, computer vision & volunteers**

The transcription of the manuscripts can be performed either manually or through computer vision techniques. Given that paper degradation and the high variability in handwriting styles in historical manuscripts impose many difficulties, the existing handwriting recognition techniques still need more development before we shall trust automatic transcription completely. For this reason, the local censuses have been transcribed by a community using a web-based crowdsourcing platform, and validated through a game-sourcing mobile application, both assisted with computer vision techniques.

Since manual transcription is tedious and time-consuming, we have developed a crowdsourcing platform for data entry (Figure 7). The idea of crowdsourcing is to split the work into many micro-tasks and ask contributions from a large group of people, especially from the online community. Thus, the task is shared among many users and finished in less time. The graphical user interface of our web-based platform offers a data entry tool with a user friendly environment, integrating both the original source and the data entry form in the same view.
Figure 7. Crowdsourcing platform: http://dagapp.cvc.uab.es/PadronsXarxes/.

Source: Authors’ own elaboration
We also use the transcribed data to teach the computer vision algorithms how to interpret further manuscripts. Concretely, our handwriting recognizer, based on deep learning, uses document images with their corresponding transcriptions to improve the transcription system and to adapt itself to every new handwriting style. As a result, these algorithms can be used to speed up the transcription in two different scenarios: first, to assist the transcription via information transfer, and second to perform an automatic transcription with manual validation.

In the first scenario, we use the redundancy in censuses to automatically transfer repeated information from one census to the next. As was stated before, local censuses in Spain were recorded in intervals of a few years and the nominative information about individuals recorded within households, was quite stable. This redundancy is used to assist the transcriber and speed up the transcription. Once one census has been manually transcribed, the redundant information (names, surnames and addresses) is transferred to the next one, so it is only necessary to update the changes manually: adding new members or deleting those who left or died in each household. For this purpose, household records from consecutive censuses are automatically aligned using the street address. Then, the individuals are automatically located in the image of the census manuscript using visual word search, «word spotting». Concretely, the individual names and surnames from a census are searched in the corresponding home’s records in the next census (figure 8). Since the process is based on a focused search, the accuracy is high. This redundancy has been used to assist the transcription of the 1886 census, once the 1881 census was transcribed (Mas et al., 2016), which led to a 70% reduction in the transcription time.

In the second scenario, the transcription is performed automatically, followed by manual validation. First, snippets corresponding to word images are segmented from the manuscripts. Then, these word images are clustered to find high frequency words that can be jointly transcribed using a small percentage of representative instances. Afterwards, the software sends the words in each cluster to the deep learning based transcription system, which provides the most plausible transcriptions for each word image (Figure 9). Finally, clusters that contain words with the same transcription can be transcribed at once. In this way, we avoid the validation of every single word, speeding up the transcription while maintaining high quality performance.
Figure 8. Architecture of the proposed system in first scenario.  
*Source: Mas et al., 2016*

Figure 9. Architecture of the proposed system in second scenario.  
*Source: Chen et al., 2018*
The manual validation has been done through gamesourcing, understood as crowdsourcing via gamification which consists in the application of game-design elements and principles in a non-game context. Concretely, two Android games have been developed (Chen et al., 2018). The first game is designed to validate the word clustering algorithm. It shows some instances of a given cluster, and asks the user to confirm that those words are the same, thus belonging to the cluster. The second game validates the output of the transcription algorithm. When the player selects one word, the system shows the most probable transcriptions according to the algorithm, and the user selects the correct answer among these possibilities (Figure 10). Experiments demonstrate that the transcription effort can be significantly reduced, and that user engagement is higher than with the traditional crowdsourcing web-based application.

Figure 10: Images of the two games. Source: Chen et al., 2018
Thanks to the coordination and collaboration with the Regional Archive of the *Baix Llobregat*, local study groups, town halls and local archives, it was possible to gather groups of transcribers, volunteering in each municipality (figure 11). Collaboration of volunteers in gathering data was already used in the “Cambridge Group for the History of Population and Social Structure” (Wrigley & Schofield, 1989). Until July 2018, 119 transcribers participated (58 females and 61 males), whose mean age is around 60 (Table 2). The transcribers had diverse cultural backgrounds, but enthusiasm about local history and genealogy in common.

Figure 11. Pictures of some groups of transcribers. Source: Authors’ own elaboration.
In order to enhance transcription quality and continuously engage the volunteers, we have performed several informative sessions for presenting the NETWORKS project and how the online application of transcription works as well as raising people’s awareness about the local censuses and their importance. These informative sessions succeed to engage and train the volunteers, who afterwards received authorization for conducting online transcription.

Table 2

<table>
<thead>
<tr>
<th>Municipality</th>
<th>Men</th>
<th>Women</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Begues</td>
<td>3</td>
<td>4</td>
<td>7</td>
</tr>
<tr>
<td>Castellví de Rosanes</td>
<td>3</td>
<td>6</td>
<td>9</td>
</tr>
<tr>
<td>Collbató</td>
<td>3</td>
<td>2</td>
<td>5</td>
</tr>
<tr>
<td>Corbera de Llobregat</td>
<td>4</td>
<td>1</td>
<td>5</td>
</tr>
<tr>
<td>El Papiol</td>
<td>4</td>
<td>4</td>
<td>8</td>
</tr>
<tr>
<td>Molins de Rei</td>
<td>4</td>
<td>4</td>
<td>8</td>
</tr>
<tr>
<td>Sant Feliu de Llobregat</td>
<td>27</td>
<td>31</td>
<td>58</td>
</tr>
<tr>
<td>Santa Coloma de Cervelló</td>
<td>8</td>
<td>3</td>
<td>11</td>
</tr>
<tr>
<td>Torrelles de Llobregat</td>
<td>5</td>
<td>3</td>
<td>8</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td>61</td>
<td>58</td>
<td>119</td>
</tr>
</tbody>
</table>

*Source: Authors’ own elaboration.*

In summary, these technological developments allow citizens and archivists to participate in the extraction of demographic information through web-based crowdsourcing platforms and gamesourcing applications, which incorporate handwriting recognition algorithms to assist the transcribers. This collaboration can be understood as a way of popularizing science for the public and helps to install critical thinking. We shall develop new user experiences in the near future, like geoprojections on interactive maps. Thus, the project facilitates the consumption and dissemination of the historical knowledge in an illustrative and pedagogical way.

**Quality checks**

Having online applications for the purpose of transcription enhances supervision and quality control in real time, as well as interplay with...
the transcribers. In addition, functions for correction are also implemented, as are periodical controls in order to ensure better transcription:

– Pre-transcription:

Recommendations for transcription were delivered to all transcribers in face-to-face training sessions, in order to ensure a common and systematic approach to transcription.

– Control during the transcription:

The data entry tool has compulsory fields for the most important variables, blocking the submission of an incomplete transcription (figure 12). Additionally, in order to avoid leaving out any household member, the total number of household members has to be declared when beginning transcription of every new household. The application includes a “Help” menu with different tools such as dictionaries, lists of names as well as geographic and occupational guides to help the transcribers when in doubt. Moreover, all the transcribers can continuously address their queries via email or social networks to enhance the communication between researchers, technicians and transcribers. Every transcriber had a reviewer in charge of verifying transcription quality and solving doubtful cases. Finally, we released statistics about the transcription and database progress to the community of transcribers in order to dynamise the transcription process.

Figure 12. Control during the transcription. Source: Authors’ own elaboration

– Post-transcription control:

Once the transcription of a given local census was finished, several analyses are conducted regarding the frequencies of different variable values for identifying likely inconsistencies.
Harmonization and codification of data

One central characteristic of the BALL Demographic Database is that instead of transcribing data in a harmonized way, the volunteers have been told to transcribe all records literally in order to avoid inconsistent interpretations of the nominative data. In this way, once the sources have been transcribed it is necessary to apply linguistic harmonization. The same individual appears with different names and surnames, or apparent variants because they were recorded originally with different spelling, or abbreviations. One combination of a common first name and surname might refer to several individuals (Goiser and Christen 2006; Herzog et al., 2007, Schürer, 2007). These questions need particular attention in the case of the Catalan language, which was not standardized until 1913. Also in occupation titles, locations, relationships with the household head or even marital status there is vast variability. Apart from reducing such variability, harmonization is important to make variables and outcomes more accessible for analysis and comparisons at the national or international level. For this reason, we recode the variables in each local census in order to remove phonetic variations caused by the different dialect of Catalan and influences from other languages which favored many written variations (Peytaví 2010; Rubió and Lizondo, 1997). For instance, we found the surname Ferrer (Smith) written as Ferré, Farré, Farrer, and so on.

Linguistic harmonization helps overcome these obstacles and clusters anthroponyms in order to compile dictionaries of names and surnames (Bloothooft, 1998; Christen, 2012). Thus, nominative data have been harmonized according to language criteria to facilitate the record linkage that identifies the same individual in different censuses (Jordà, 2016; Jordà et al., 2013). In addition, places have been geolocated, and occupations have been coded using the Historical International Classification of Occupations (HISCO) (van Leeuwen et al., 2002).

Once all occupations were codified, we ranked and classified them according to sociooccupational position by means of HISCLASS (van Leeuwen and Maas, 2011) and HISCAM (Lambert et al., 2013), respectively. HISCLASS differentiates individuals in consonance with the social group to which they belonged according to dimensions like manual/non-manual division, skill level, degree of supervision and economic sector, which gives 12 different classes going from unskilled rural workers at the bottom to higher managers and professionals at the top. HISCAM is a different occupational stratification scale based
on the Cambridge Social Interaction and Stratification scheme, using marriage data from Belgium, Britain, Canada, France, Germany, the Netherlands, and Sweden (Prandy, 2000). The main idea behind this scaling is that individuals who interact more (in terms of occupational and social relations) are closer in terms of social position, assuming that these interactions represent the occupational stratification structure. The result is a ranking of occupations (theoretically from 0 to 99), showing not only similar social standing but also differences between occupations. Finally, we have grouped each occupation into economic sectors, following an adaptation of HISCO for the Catalan historical labor market (Pujadas-Mora et al., 2014).

As it was mentioned before, the BALL Database, in July 2018, is composed of 9 municipalities (figure 2) and 152 local censuses that correspond to a total of 221,667 individual observations (see table 1). The harmonization of its names, occupations and places uses the reference tables created during the harmonization of the Barcelona Historical Marriage Database (Jordà, 2016). In this way, we rely on a reference table for the names (female and male names) with 27,434 unique records, the surnames with 101,238 unique records, the occupations with 24,399 unique records and the place names with 47,556 unique records, which facilitate the work of linguistic normalization and harmonization.

Towards longitudinal data:
The Sant Feliu life course database

Among potential derivations from the BALL Demographic Database is the possibility to create longitudinal databases of individuals and households by linking local censuses. For this, we have used one of the most populated towns in Baix Llobregat, one with a large time coverage, which is also the capital of the county, namely Sant Feliu de Llobregat, to create a sub-database. Thus, the longitudinal database of Sant Feliu is based on reconstruction of individual life-courses using local censuses. The town of Sant Feliu de Llobregat was one of the most important in the region, in economic and administrative terms, being the judicial district capital, with the arrival of new economic activities such as textile and metallurgical industries since the second half of the nineteenth century and the railway station in 1855. The database contains the information in all the 15 censuses recorded in Sant Feliu from 1828 to 1940. This information has benefited from computer-assisted data transcription through crowdsourcing in which 58 volunteers have col-
laborated (27 men and 31 women) for a period of 2 years. As explained above, the local censuses in Spain (including Sant Feliu) were taken at short intervals.

The Longitudinal Database of Sant Feliu contains 59,084 observations of individuals, which increased chronologically, mainly from the 1920 onwards when Sant Feliu de Llobregat started having important immigration. The individuals are distributed in 12,748 households across the entire period with a mean number of 4.6 individuals per household, levels that decreased over time from 5.2 persons per household in 1828 to 4 in 1940 (table 3). The nominative data has been harmonized, addresses have been geolocated, and occupations have been encoded.

### Table 3

<table>
<thead>
<tr>
<th>Years</th>
<th>Individuals</th>
<th>Households</th>
<th>Individuals per Household</th>
</tr>
</thead>
<tbody>
<tr>
<td>1828</td>
<td>2,209</td>
<td>426</td>
<td>5.2</td>
</tr>
<tr>
<td>1833</td>
<td>1,470</td>
<td>313</td>
<td>4.7</td>
</tr>
<tr>
<td>1839</td>
<td>1,946</td>
<td>377</td>
<td>5.2</td>
</tr>
<tr>
<td>1857</td>
<td>2,472</td>
<td>533</td>
<td>4.6</td>
</tr>
<tr>
<td>1878</td>
<td>2,762</td>
<td>610</td>
<td>4.5</td>
</tr>
<tr>
<td>1881</td>
<td>3,005</td>
<td>598</td>
<td>5.0</td>
</tr>
<tr>
<td>1889</td>
<td>3,118</td>
<td>644</td>
<td>4.8</td>
</tr>
<tr>
<td>1906</td>
<td>3,606</td>
<td>805</td>
<td>4.5</td>
</tr>
<tr>
<td>1910</td>
<td>3,809</td>
<td>866</td>
<td>4.4</td>
</tr>
<tr>
<td>1915</td>
<td>4,330</td>
<td>936</td>
<td>4.6</td>
</tr>
<tr>
<td>1920</td>
<td>4,353</td>
<td>918</td>
<td>4.7</td>
</tr>
<tr>
<td>1924</td>
<td>5,575</td>
<td>1081</td>
<td>5.2</td>
</tr>
<tr>
<td>1930</td>
<td>6,392</td>
<td>1459</td>
<td>4.4</td>
</tr>
<tr>
<td>1936</td>
<td>7,023</td>
<td>1458</td>
<td>4.8</td>
</tr>
<tr>
<td>1940</td>
<td>6,727</td>
<td>1675</td>
<td>4.0</td>
</tr>
<tr>
<td>Total</td>
<td>59,084</td>
<td>12,748</td>
<td>4.6</td>
</tr>
</tbody>
</table>

*Source: Authors' own elaboration.*
Harmonization and codification significantly reduced the number of spelling variations, erroneous or misleading data. From the initial 1,976 given names transcribed literally, the harmonization dropped the number to 1,033 variants. Since Spanish individuals usually receive two surnames (paternal and maternal), the surname harmonization was multiplied by two. The 4,269 different paternal and 5,641 maternal surnames, became 2,767 and 3,312 after harmonization, respectively. Finally, at the geographic level, the 3,743 different addresses observed were harmonized to 1,552 different geographic points. Nevertheless, as the database includes local censuses spread over a wide period, not all contain the same information. For instance, local censuses in Sant Feliu prior to 1906 did not have information on the specificity of relationship with the household head, i.e. who were spouses, children and/or other relatives. We have used other variables such as surnames, marital status and age in order to reconstruct family relationships within the same household. Finally, the procedure of harmonizing and coding the occupations resulted in 1,611 different occupational titles that we simplified into 292 HISCO codes.

Figure 13. Life course example. Bartomeu Carcereny Amigo (1874-1940).
*Source: Authors’ own elaboration.*
After standardization, nominative record linkage was performed. The procedure uses deterministic (two string distances, the Bag distance and the Levenshtein distance) and probabilistic criteria, together with a language model that avoid similarities in pronunciation to be considered as the same surname (e.g. Piera & Riera, Pons & Pous, etc.) to and penalized typographic errors (Villavicencio, Jordà, and Pujadas-Mora, 2015). This allowed us to follow 10,405 individuals at least in two different censuses. In the best cases, we reconstructed up to 12 different observations for individuals, which signifies assessing the entire life-cycle of a person. For instance, consider the life course of Bartomeu Carcereny Amigó (figure 13). This individual was a native of Sant Feliu and born on 1874. From then on, the longitudinal database allowed us to follow him in 11 different enumerations across his life until the last observation found in 1940, when he was 66 years old.

As can be seen from this example, one possibility of analysis with longitudinal data, is the work career of individuals. The first two observations of Bartomeu were without occupation. Around 14 years old, we found Batomeu’s first occupational observation as locksmith apprentice, which he continued until the first decade of the twentieth century. However, from the 1920s until the 1940s the four last observations of Bartomeu informs us of two different intermittent occupations, mechanic and turner. Thus, as can be observed through a single case, the longitudinal database of Sant Feliu informs us not only about changes at the individual micro level, but also how this individual reacted to macro level transformations such as industrialization, given he turned from an artisan position to an occupation that emerged with a new mode of production.

**Popularizing science: Onomastics browser & demographic visualizations**

The NETWORKS project has an important commitment to citizen science that goes beyond their participation in census transcription. The project also aimed at reaching the public by contributing to society through the creation of a browser constructed to display onomastics and interactive demography (figure 14). In May 2017 the onomastics application of the Sant Feliu database was launched: http://158.109.8.76/xarxes/. This also shows the genealogical tree (with a varying number of generations) for each individual in the town.
Figure 14. Onomastic browser. *Source: Authors’ own elaboration.*
Moreover, it includes interactive visualizations of population pyramids and the town’s most common names and surnames during the nineteenth and twentieth centuries. Additionally, we linked the web site with Wikipedia to provide the historical context of the database and biographical information about individuals recorded in the censuses. Wikipedia articles are also presented through an app (figure 15).

Another mobile app is being developed to enrich the census information with pictures, kept in historical archives or family archives.

Figure 15. Wikipedia connection. Source: Díaz, 2018

Concluding remarks

Citizen science, crowdsourcing, historical demography, computer vision and gamification are the cornerstones of the *Ball demographic database*. Formally, the Ball database is a relational database containing census data at the individual level from the Catalan county of *Baix Llobregat* in Spain for the nineteenth and twentieth centuries with more than 221,667 individual observations built by different interdisciplinary projects in Digital Humanities. These projects are a joint venture of the Center for Demographic Studies and the Computer Vision Center, both at the Universitat Autònoma de Barcelona. Hence, this database is a proof that activating volunteer citizens through an *ad hoc* online crowdsourcing platform as a data entry tool and games for touch-screen devices, and Handwritten Text Recognition (HTR) techniques into data collection of primary sources can meaningfully reduce the time for building individual-level historical demographic databases. This integration has been possible thanks to recent advances in Handwriting Recognition, the expansion of information technologies, the popularization of handheld devices, the assimilation of internet in everyday life and the massive digitalization of historical documents. However, the current state-of-the-art of Handwriting Recognition still requires some human intervention.
Thus, two scenarios have been designed in BALL database to speed up transcription using computer vision algorithms, based on deep learning techniques. First, assisted transcription by information transfer from one document to another and second, automatic transcription with manual validation. These scenarios have proved to reduce progressively the human effort. In the first scenario, the system benefits from the redundant information (names, surnames and addresses) shared by consecutive censuses only needing the manual transcription (through a crowdsourcing platform) of one census, whose information is transferred to the next census. Using visual word search, so-called word spotting, each transcribed household in the first census is linked with the image of the same household a few years later, assisting the manual updating of information from one census to the next one. Moreover, the use of gamesourcing experiences for the validation of automatic transcription carried out by handwriting recognition algorithms is a twist in the integration of computer vision methods and keeping the human intervention. At the same time, the engagement of citizens through gamification appears to be higher, as can be expected, which has a beneficial effect on the total number of transcribed words. Overall, we have set the stage for effective semi-automatical processing of large document collections in order to create databases in a faster and more effective way as part of the Big Data revolution.

Additionally, the BALL database is more than a database for demographic and computer vision research. The possibility of browsing this database through names and surnames using an open access and user friendly webpage boost greatly their use, mainly for the public at large, who is usually not familiar with this kind of sources. Queries can be also done through genealogical trees, that have been generated using record linkage techniques, facilitating extremely the search task for many users, who are looking for their ancestors. At the same time, visualizations of the population’s evolution and onomastics through graphics are offered as tools to popularize demography among the general audience. In this way, citizens who are a fundamental part in the building of the database can also take advantage of the final result through the above-mentioned webpage, which is an important asset in terms of knowledge transfer to society. Besides, in spite of pursuing scientific objectives, these citizen-centered projects have an important potential social impact in terms of literacy or shortening the technological gap in terms of social inclusion and cohesion.

*Source: Authors’ own elaboration.*
Chapter 2

References

Alter, George and Mandemakers, Kees. (2014). The Intermediate Data Structure (IDS) for longitudinal historical microdata, version 4. Historical life course studies. 1


Pujadas-Mora et al.


Chapter 3.
INTEGRATING NOMINATIVE DATA ON THE ORTHODOX AND GREEK CATHOLIC CLERGY IN MODERN TRANSYLVANIA: AN OUTLINE

Marius Eppel,
Oana Sorescu-Iudean,

Writing in the late 1830s, the British traveller John Paget described the Romanian clergy near the former Roman-era settlement of Densuş, in the County of Haţeg in Transylvania, noting that “except for a somewhat greater neatness of person, and the long black beard which hung down to his breast, the Wallack priest was in no way distinguished from the humblest of his flock.” Perhaps surprisingly for a traveller used to the ranks of the Anglican Church, the figures of the Romanian (likely Orthodox) ecclesiastical hierarchy were in many respects similar to those whom they shepherded: “With just enough education to read the service of the church, just enough wealth to make them sympathize with the poor, and just enough religion to enable them to console them in their afflictions, these men exercise a greater power over the simple peasant than the most cunning Jesuit, the most wealthy Episcopalian, or the most rigid Calvinist” (Paget, 1839, p. 202). Despite the characteristic traveller’s lens through which Paget viewed the Romanian clergy in Transylvania, as well as most of the affairs in Hungary at the time (Bökös, 2017), the account was deemed, even by contemporary Hungarians, to be quite “effective” in describing the area’s characteristics (Popova-Nowak, 2008, p. 215). It is therefore likely that the image of the clergy as an historical actor who wielded a great deal of influence over the communities of faithful was not far from the truth. Nevertheless, the extent to which the placement of the Romanian clergy on the social-economic ladder in the province conforms to the same image remains an open question.

Within the specific milieus of the composite state of the Habsburg Empire, and after 1867, Dualist Hungary, social-occupational groups such as the clergy, who could mobilize the great, critical mass of the

---

1 The study was supported through the grant CNCS – UEFISCDI, project no. PN-III-P4-ID-PCE-2016-0661.
peasantry towards more secular purposes, along with the entrenchment of nationalism, were of crucial importance (Hitchins, 1994, p. 169–170). Despite the flourishing of a consistent strand of ecclesiastical historiography in the past two decades, the Romanian clergy in Transylvania remain largely an unknown quantity in many respects. Leaving aside various narrative accounts, contemporary press articles, or other textual sources, the middle clergy – the parish priests – have not been studied comprehensively from the perspective of their inner structural roles within the mosaic of nations and confessions present in the area. What is more, apart from fragmentary studies, they have rarely been dealt with as a self-standing category, which straddled the ranks of the elite and those of the great majority of the rural population.

This paper proposes to outline the development of a database that will enable historians to obtain a much-needed systematic view of this social and professional category in the multi-denominational and multi-ethnical state of Dualist Hungary, focusing primarily on the basic, lowest level, that of the individual mid-level clergyman and the parish, the two building blocks of the national-ecclesiastical construct. As many other recent projects seeking to implement a digital approach to historical sources concerning a social-professional group have evidenced, a substantive account of the way in which the database has been envisaged and designed is direly necessary in order to open it towards a wider category of scholars and to place it firmly within the realm of traditional historical endeavours. The present study will therefore adopt a fourfold structure: it will firstly provide a brief overview of the literature concerning the Romanian middle clergy during modernity from a structural perspective, noting this group’s entanglements with other fields of history apart from the ecclesiastical sphere; secondly, it will situate the database into the recently flourishing landscape of digital initiatives focused on similar categories in other regions or timeframes; thirdly, it will outline the “universe” of the database (Mandemakers, Dillon, 2004), focusing on the core sources used, their characteristics, and the extent to which they can be integrated into a common data structure; finally, it will offer a preliminary sketch of the database, focusing on the main entities – individuals, offices, and parishes.

The Romanian Middle Clergy as Historical Actor: A Brief Overview

Throughout the nineteenth and early twentieth centuries, and especially after the establishment of Dualism, competing national demands
made themselves increasingly felt in Transylvania as in many other composite territories of the Dual Monarchy. Since early modernity, Transylvania had been home to some nine ethnic groups (Romanians, Hungarians, Germans, Slovaks, Ruthenians, Jews, Croats, Serbs, and Gypsies), resulting in a conglomerate which overlapped partially with a diverse landscape of confessions (Holom, Sorescu-Iudean, and Hărăguș, 2018, p. 331–333). The Romanians in the area had either remained part of the Greek Orthodox denomination, or had adhered to Greek Catholicism, a Church united with Rome, but which had nevertheless maintained its confessional and national distinction (Brusanowski, 2017, p. 187–197). Both Romanian Churches have been the object of consistent research endeavours, which have adopted various perspectives. Given the limits of the current paper, we will single out only a few exemplary studies that showcase the approaches employed, as well as the important structural positions occupied by the middle clergy within Transylvanian society and the Hungarian state, after 1867.

Essential for the understanding of the frameworks and evolutions of the two Romanian churches are first and foremost the general ecclesiastical histories, which mainly focus on the activity of the highest ranks of the clergy, on institutional development, and on the process of confessional definition. For the Greek Catholic Church, the ties established with the Holy See are of prime importance in understanding most of these aspects (Sima, 2013). Worth noting, though they mainly tackle eighteenth- or early nineteenth-century developments are the recent histories of the Greek Catholic Church (Dumitran, 2007; Nedici, 2013) as well the older but still authoritative work by Genț (1912), which provides a detailed image of the various waves of administrative reorganisation during the nineteenth century. What might be termed traditional ecclesiastical histories, though written from an avowedly confessional perspective, are still the primary reference point for researchers dealing with the evolution of the Orthodox Church in Transylvania (Floca, 1990; Păcurariu, 2013). More recently, the works by Brusanowski have shed light on the institutional growth and definition of the Orthodox Church and its confessional educational system (2005, 2007). The structural parallels between the two Romanian Churches during the modern period have also been briefly analysed in a recent study (Cârja, 2007a). A greater proliferation was witnessed by case studies of various breadths, ranging from examinations of ecclesiastical jurisdictions such as vicarages (Andrei, 2006) or deaneries (Covaci, 2007) to single communities (Pantșuru, 2013). As part of these well-circumscribed endeavours, historians have also focused on publishing the plentiful source material
produced by various ecclesiastical structures (Suciu et al., 2011). However, by far the most abundant strand of research has targeted individual clerical figures, generally of higher ranks, either by adopting a biographical approach (Cârja, 2007b) or by employing a prosopographic framework through which the lives and activities of certain sub-groups of clergymen were recovered as part of the current historiography focused on elites (Popa-Andrei, Covaci, and Sima, 2013). This latter type of analysis has also been spurred by the current tendency to offer an in-depth view of those historical figures who played a direct part in the events marking the end of the First World War and the establishment of the Greater Romanian state in 1918 (Dăncilă-Ineian, Judean, 2018). Issues that have a direct bearing on the structural development of the two Romanian Churches such as the conversion of entire communities from one denomination to the other have also been explored, but only for select deaneries (Drăgoi, 2011). Most of the abovementioned studies have also dealt at least cursorily with questions stemming from the sphere of social history, such as the perpetuation of the clerical profession both intra and inter-generationally, the recruitment of priests (Popa-Andrei, Covaci, Bedecean, et al., 2014; Popa-Andrei, Covaci, 2013) the social and economic status of priests within various communities or milieus, their national activism, etc. Recently, researchers have also dealt with issues such as the pension system implemented for the Greek Catholic clergy in the latter half of the nineteenth century (Covaci, 2012).

Nevertheless, the research landscape remains deeply fragmented, and case studies, though detailed, cannot encompass broader jurisdictional forms, especially those wider than the diocese. As in the case of other national ecclesiastical institutions, such as for instance the Anglican Church, no definitive aggregate figures exist for the number of clergymen activating in Transylvania from the mid-nineteenth to the early twentieth century (Burns, Fincham, Taylor, 2004, p. 729). Analyses that rely on aggregate data for one diocese or another also necessarily eschew individual-level approaches, outside of brief narrative intermezlos. These in turn can only exemplify subjectively selected tendencies, mainly adhering to commonplaces already noted in historiography, rather than help to clarify the origins and results of broader patterns in the social and geographic mobility of priests or the economic and geographic characteristics of parishes. The salience of this group has also been emphasized by works examining the numerous entanglements between the clergy and the Romanian national movement (Hitchins, 1994, p. 169–220), the banking system (Balog, 2014), and the Romanians’
main cultural association in Transylvania (Popovici, 2018). It can therefore be stated that obtaining a clear image of those who composed this group, their attributes, as well as those of the frameworks within which they exerted their offices is paramount to understanding not only the clergy, but also the institutional history of Transylvania, in its broadest sense.

The Confessional Infrastructures Database within the Current Digital Landscape²

Certainly, the initiative to map out professional categories or educationally-circumscribed groups within a certain territory, especially during the nineteenth and twentieth centuries, is not a wholly novel enterprise. Researchers have attempted to flesh out the staff of the Imperial court in early modern Vienna by means of an extensive relational database³ (Kubisk-Scharl, Pölzl, 2013), or to document the medieval and early modern graduates of faculties throughout the Holy Roman Empire⁴ (Gubler, Schwinges, 2018), by implementing digital systems to deal with wide-ranging prosopographic accounts. The records produced by various churches within Europe have likewise not escaped the digital gaze within the past decade, as, compared to sources produced by other types of institutional actors, and depending on the historical period, they offer an incomparable wealth of information. What is more, the structural uniformity characterising the Churches present throughout Europe results in similarly structured records, that either allow for the creation of structured datasets, or are especially promising for various social-historical endeavours.⁵ An example of the latter type of approach is the Consistory Database, which contains depo-

² The database as well as a comprehensive user manual and source documentation will be available starting June 2019 at http://confinfrahist.granturi.ubbcluj.ro/. For a brief review of the project itself, see http://www.elite-research.eu/Socio-professional%20trajectories.html.
⁵ We are excluding from this brief overview perhaps the most wide-ranging databases dealing with Church-produced records, namely the historical population databases integrating information from parish registers. For an overview of the majority of existing population databases, see https://ehps-net.eu/databases (accessed December 2018).
tions offered in late medieval trials held before the London Consistory, and deserves to be highlighted among such endeavours.\(^6\) What is more, the Church’s potential to produce reliable statistical records has long been recognized by both early modern states and contemporary researchers, as evidenced by projects such as that of the Statistical Accounts of Scotland Online, which recently centralized in an accessible document portal the accounts produced by eighteenth and nineteenth century Scottish parish ministers at the initiative of Sir John Sinclair.\(^7\)

The focus on the material side of ecclesiastical infrastructures has also recently come to the fore of digital research, yet again in the English milieu, with projects such as that making use of the churchwardens’ accounts of England and Wales, soon to be released online.\(^8\) Finally, the closest methodological “relative” of our current endeavour is the “Clergy of the Church of England Database” (CCEd), which provides a digital infrastructure populated with the careers of all clergymen in England from 1540 to 1835.

While our current endeavour shares many common traits with those briefly noted above, the specificity of the region, timeframe, and sources under question lend it a distinct character. First, it focuses on both the clergy as individual historical actors, attempting to reconstruct their educational and professional pathways, as well as on the confessional infrastructures in which they functioned. As opposed to previous projects, it integrates data from two different denominational strands – the Orthodox and the Greek Catholic – which will aid future researchers to bridge the denominational divide in the study of the Romanian ecclesiastical culture, an issue that is currently difficult to overcome due to the breadth of records and the differences in practices. Secondly, as will be detailed further, it seeks to provide a platform to integrate a significant variety of historical sources, both narrative and structured. In this sense, a primary goal is to remain open-ended in a certain sense, allowing future researchers to complement the broad image resulting from present data entry efforts with smaller-scale but more in-depth records tied to entities already present in the database. What is more, the current project goes beyond the recording, standardisation, and linkage of clerical careers in modern Transylvania, by delving into the con-


fessional infrastructures themselves and integrating data related to the material patrimony of parishes, ecclesiastical and school buildings, and other ecclesiastical administrative structures (diocese-level schools, chapters, etc.). Finally, like the initiative targeting the churchwardens’ accounts of England and Wales, it seeks to provide timestamped images of ecclesiastical organisation and jurisdiction, which will be tied to the secular administrative organisation in the area, as revealed in the censuses of 1850–1910.

The “Universe” of the Database: Sources and Challenges

This section will briefly chart the “universe” of the database, as per the recommendations outlined by Mandemakers and Dillon (2004). A fuller account of all these issues will be offered on the database website, well before its public release, as well as in a further study documenting the structure of the database upon the completion of the project.

It should firstly be noted that our endeavour focuses on the Greek Catholic and Orthodox clergymen from two major ecclesiastical administrative units in Transylvania: the Greek Catholic diocese of Alba Iulia and Făgăraș, and the Orthodox diocese of Sibiu. These two units were the most important for each confession, seeing as they housed the decision-making centres of the Greek Catholic and Orthodox Churches in Transylvania. The diocese of Alba Iulia and Făgăraș (previously Făgăraș) was established as such in 1850 and was raised to the rank of Greek Catholic metropolitan see or archdiocese in 1855. Likewise, the diocese of Sibiu was among the oldest Orthodox dioceses in Transylvania, and became an autonomous metropolitan see in 1864. We have also hypothesised, based on previous surveys, that due to their centrality and rank within the entirety of the ecclesiastical structures of the two Churches, the two dioceses generally also produced reliable records and allowed the highest ranks of the clergy to transmit information to the lowest ranks with more ease and accuracy (Sima, Covaci, 2015).

Throughout the nineteenth and early twentieth century, both Romanian Churches in Transylvania produced a type of record generally referred to in secondary literature as schematismae, or yearbooks. These sources were not a specific product of modern Transylvanian ecclesiastical milieus, but rather a kind of record generated by the Catholic or United Catholic Church throughout Europe since early modernity. Church schematismas mapped out the entirety of the positions within the ecclesiastical hierarchy for a certain unit such as a diocese, from the
highest to the lowest level, and listed the individuals occupying each office at a certain point in time. Such sources were produced for instance in the historical dioceses from nowadays Ukraine during the 1840s and 1850s\(^9\), in most German bishoprics since the middle ages\(^{10}\), in historical Polish\(^{11}\) or Hungarian dioceses.\(^{12}\) Almost structurally identical sources were also produced by the imperial Habsburg court, initially only for the court staff in Vienna, and later on for the entire administrative apparatus in its provinces (Kubiska-Scharl, Pölzl, 2013). Despite the perceived discursive split characteristic to the process of confessionralization in East-Central Europe, the Protestant and Reformed Churches also adopted this mechanism in order to maintain a clear image of their infrastructures and staff, producing highly similar records.\(^{13}\) Despite the regional specificities of the institutions producing them and the resulting professional attributes of the persons listed therein (clerical officials, military officials, state officials, etc.), these types of sources are widespread throughout Europe, and have close equivalents even on the British Isles, in the so-called *libri cleri*, registers of appointments, or licensing books (Burns, Fincham, Taylor, 2004, p. 730). We may therefore state that, as far as the potential to create internationally-compatible datasets and to conduct transnational analyses of the parish clergy is concerned, these types of sources are extremely well-suited.

The ecclesiastical *schematism* or yearbooks produced by the two Romanian Churches became more widespread in Transylvania starting from the middle of the nineteenth century, most likely tying into the administrative reforms that led to the establishment of the two archdioceses. Certainly, these sources were not the first attempts by the Church or the state to obtain a clear image of ecclesiastical infrastruct-
tures within this territory. Starting from the 1730s-1740s, the Habsburg authorities would request that both the Greek Catholic and Orthodox submit conscriptions of church property, and of the number of families or households in their respective communities, as an instrument to aid in the definition of confessional policy in the province (Dumitran, 2007, p. 168–220). Many of these conscriptions were edited throughout the nineteenth and early twentieth centuries, with varying degrees of accuracy, while their findings would constitute points of contention within the competing national narratives of the late nineteenth century (Câmpeanu, 2003, p. 116–117). Nevertheless, due to significant variance in the quantity and completeness of information provided by these eighteenth-century sources as compared to their nineteenth-century equivalents, the current project will eschew their data entry and integration at this time. What is more, as many eighteenth-century conscriptions were carried out at the behest of or by imperial-appointed authorities in Transylvania, significant doubts have been expressed regarding their accuracy and completeness. Certainly, this does not wholly eliminate the possibility of deepening the current data entry efforts into these sources in the future, as faulty data can be corrected or complemented by appealing to other serial sources stemming from local ecclesiastical archives.

Church schematisms were initially published yearly, as part of calendars. In the case of the Orthodox Church, this rhythm of publication would survive. Up to date, the present project has entered the schematisms for the Orthodox archdiocese of Sibiu from 1882 to 1914, totalling over 73 000 lines. The Greek Catholic schematisms were only published periodically, roughly every decade or more often, thus resulting in cross-sectional views of the state of the confessional infrastructures and staffing. However, on bicentennial anniversaries of the establishment of a certain diocese, the Greek Catholic Church published detailed histories of the staffing and administrative organisation of the respective ecclesiastical unit. This was the case for the Schematism for the Archdiocese of Alba Iulia and Făgăraș of 1900, which included all known clergymen who had activated in the parishes part of the deaneries in the diocese since 1733 and up to 1900. At the end of this schematic account, small-scale prosopographic accounts of those clergymen who were active in the diocese in 1900 were also added. What is more,

14 The Greek Catholic schematisms have been digitised by the Library of the Romanian Academy in Cluj-Napoca and are available at http://www.biblio.acad-cluj.ro/?q=node/4. Further digitized schematisms are also available online at http://byzantinohungarica.hu/node/141 (both links accessed December 2018).
Greek Catholic schematisms were prefaced by detailed biographical accounts of the major figures in the higher ranks of the clergy. Presently, this schematism has already been entered as raw data and is undergoing standardisation. For the Orthodox archdiocese of Sibiu, the schematisms were complemented by so-called “Soul Protocols” (*Protocoalele sufletelor*), which contained information on the clergymen ordained and appointed in the parishes of the diocese, including the dates of ordainment. These types of sources were available at the ecclesiastical archives in Sibiu. Presently, the project team has managed to enter the information from these protocols for the years between 1857 and 1868. What is more, on certain years, each entry for a certain parish also contains demographic information, listing the total number of Orthodox individuals, of deceased individuals, or births (all by gender), as well as the number of marriages contracted. Information on the historical collection of data for the Greek Catholic schematisms is available in ecclesiastical circulars, issued by the respective diocese roughly one year prior to the publication of the schematism itself.\(^{15}\)

As far as the estimation of source reliability is concerned, nineteenth century schematisms fare much better than their eighteenth-century correspondents. As has been noted, the anniversary accounts list the entirety of the clerical staff in the diocese for a period of two centuries and note all administrative changes in the diocese (i.e. to which deanery a certain parish belonged, at a certain point in time, and how this hierarchical relationship evolved). What is more, each parish entry contains information on the confessional make-up of the community: the figures for each confession outside of the primary, Greek Catholic one, are taken from the census conducted by the Hungarian state in the same year. It is therefore possible to compare state-derived figures with those provided by ecclesiastical authorities, and to ascertain whether any areas showed significant discrepancies. The series of Orthodox soul protocols, though incomplete in their recording, also allow the comparison of these figures, at least for the Orthodox population in Transylvanian communities. Given the high stakes of a community or area’s confessional structure in a multi-denominational and multi-ethnical region, where the state began to enact de-nationalising legislation towards the later decades of the nineteenth century, the ability to compare state-driven recording initiatives with data stemming from the communities

\(^{15}\) Serviciul Județean Alba al Arhivelor Naționale (hereafter abbreviated as SJANA), *Fund Mitropolia Română Unită Blaj, Administrația Centrală Arhidiecezană*, folder 1/1910, ff. 177v-178v.
themselves (recorded by parish priests) is quite significant.

These types of sources – schematisms and soul protocols – represent the current core of the database initiative and will form the backbone to which other kinds of serial records will be appended in the future. The choice of these core-records was a natural one, as they provide a complete image of the staffing and infrastructures of the two Romanian Churches in Transylvania.\(^\text{16}\) The next sources that we plan on introducing into the database “universe” consist in registers containing information on the priests ordained at the theological academy of Blaj (the centre of the archdiocese of Alba Iulia and Făgăraș) starting with 1869\(^\text{17}\), on the widows and orphans of priests in the same diocese and their financial support\(^\text{18}\), and on the situation of vacant parishes in the diocese in the early twentieth century.\(^\text{19}\)

### Database Structure: A Cursory Overview

As the database design process is not complete at the time, and the general user interface has not yet been created, this section will only provide a cursory overview of the main principles behind the process of construction, with reference to concrete cases and source contents.

The concept of the database is a quite simple one, inspired by two already extant database models: the source-oriented Historical Population Database of Transylvania\(^\text{20}\) and the research-oriented Historical Data Grinder.\(^\text{21}\) Like the latter, we adopted an entity-attribute-value model, but owing to the experience of introducing a digital solution to the field of traditional historical scholarship, we decided on keep-

---

\(^{16}\) Selected images from the source material, singling out and explaining their structural elements at variable level will be available on the project website by the end of 2019.

\(^{17}\) SJANA, Fund Academia Teologică Greco-Catolică Blaj, folder J1/1869-1948, Protocolul preoților ordinați în Arhidieceza gr.cat. de Alba-Iulia și Făgăraș începând cu anul 1869.

\(^{18}\) SJANA, Fund Mitropolia Română Unită Blaj, Arhiva Generală, folder 13/1890-1922, Conspectul preoțelor văduve și al orfanilor de preoți din Archidiecesa gr. cat. de Alba-Iulia și Făgăraș.

\(^{19}\) SJANA, Fund Mitropolia Română Unită Blaj, Administrația Centrală Arhidiocesană, folder 201/1907-1916, Protocolul parohiilor vacante în Arhidiocesa gr.cat. de Alba Iulia și Făgăraș.


ing closer to the original source structure as a means of increasing this instruments’ accessibility in the future. The database is presently structured into several main and several secondary tables: a source table, separate tables for each kind of ecclesiastical administrative unit (diocese, deanery, vicarage, parish), a table containing individuals, a table containing all ecclesiastical offices mentioned in the sources, a table for what we have termed “ecclesiastical administrative structures”, which contains information on ecclesiastical courts, chancelleries, and the like, and a table linking individuals to offices and to certain ecclesiastical administrative units at certain points in time. Source-faithful spreadsheets were created over the past year and half of project implementation, have been broken up into their respective constitutive entities at the highest possible level of granularity, and are currently being standardised. The lowest level of ecclesiastical structure – the parish – is also linked to a table of localities, extracted from the censuses of 1850-1910 (each census counts as one source, so that the timestamps of ecclesiastical sources and those of secular sources correspond): the parish of Aiud (with id 1, of deanery id 1, of diocese id 1) is therefore linked to the locality no. 1 of a certain constituency and county, as per the census hierarchy. This was not always a straightforward process, as parish as well as place names were not only presented in several languages (Romanian, Hungarian, Latin), depending on the source, but could also change over time. Despite the increased time allocated to data entry and standardisation involved by creating direct correspondences between ecclesiastical and lay administrative units over time (exclusively for the two dioceses under study), we believe that this will aid in future projects seeking to analyse the entanglements between the Church and the lay authorities, shedding light for instance on the factors determining the involvement of parish priests in certain temporal matters, such as elections, national associations, etc. Moreover, the allocation of source timestamps that will be attributed to a certain make-up of the ecclesiastical administration of the two dioceses will enable researchers to examine the evolution of this infrastructure over time, as well as cross-sectionally. As these snapshots of confessional infrastructures had previously been disparate, only tenuously tied in individual studies of deaneries or communities or in cumbersome sources, the longitudinal history of parishes or deaneries could not be described fully.

The individual table will contain the typical fields encountered in serial sources (first name, last name, nickname, title, noble particle, gender, etc.), and will be linked both to the offices table, as well as to vari-
ous educational institutions or to various ecclesiastical administrative structures (diocese level courts, etc.). Individual names will be the final fields to undergo standardisation and will appear in the three major languages encountered in the sources (Romanian, Hungarian, and Latin).

The office table has been the first standardised, and exposes the main difficulty encountered in building the conceptual framework of the database, namely the differences in terminology between the official vocabularies of the two Romanian Churches. These differences mainly stemmed from the confessional poles towards which the ecclesiastical hierarchies gravitated: in the case of the Greek Catholic Church, Roman Catholic, Latin-inspired terminology was usually employed, while the Greek Orthodox Church maintained more of its Byzantine (Greek) lexicon. Given that the great majority of offices were structurally equivalent, we created a column of standard office titles in English, with columns listing the Latin and Romanian equivalents. Thus, the Orthodox “presbiter” and the Greek Catholic “protopop” designated the same official, namely the head of a deanery (“tract presbiterial”, “protopresbiteriat”, “protopopiat”, “district protopopesc”): both versions were standardised with “dean/decanus/protopop”, in order to avoid unnecessary proliferation of confession-specific variants. Similarly, we decided to standardize so-called office statuses: many of the priests or the teachers at confessional schools were listed “pensionat” (retired), “ambulant” (travelling, meaning teaching at several schools in neighbouring parishes), “interimar” (temporary or substitute). We dealt with offices that were singled out as being exerted in a certain ecclesiastical administrative unit, such as the deanery (“administrator protopopesc”) or the parish (“administrator parohial”) by linking a particular office to the precise administrative unit in which it was exerted: thus, a temporary or substitute parish administrator was standardized as an individual exerting the office of administrator, with the attribute “temporary” in a certain parish. Likewise, a deanery administrator was linked to the office of administrator (full) in a certain deanery. This eliminated the great majority of office descriptions listed in the schematics, rising as high as 700, to only some 40 standard offices, with other attributes. The standard office table also contains meta-variables created in order to help researchers who are perhaps not entirely familiar with both ecclesiastical hierarchies and to aid in the study of broader groups of clergymen: offices exerted at diocese level were coded as “upper level”, those in the parish “low-level”, and those in structures that bridged these units, such
as defenders or prosecutors in ecclesiastical courts as “mid-level”.  

**Conclusion**

Ecclesiastical history in Transylvania has a well-established tradition, and many works have been published on the Greek Catholic and Orthodox churches in modern times. Researchers have examined the Church-State relations, how the network of confessional schools supported by these two institutions contributed to cultural development of the Romanians in Hungary, the involvement of the clergy in the efforts of the First World War, and many other tangent topics. However, very rarely have these research endeavours resulted in structured datasets that could be integrated and harmonized in order to paint a broader picture. The reasons for this are manifold, but its outcome is relevant for the entire field: fragmented analyses still predominate, comparisons across broad time periods and geographical areas are not possible, and broader topics are difficult to tackle. We argued that the creation of the database is therefore direly necessary, if only in order to overcome these issues. As long as researchers continue to work on their own, with little aid from or contributions to digital infrastructures, it will be difficult to make broader arguments.

Our decision to start our data entry and database design endeavour with the most obvious choice of source material, namely the schematizations produced by the two Romanian Churches in Transylvania, will hopefully allow us to lay a strong and reliable foundation in digitally-focused ecclesiastical history in this part of Europe. Seeing as they cover all aspects of ecclesiastical structure, from the complex hierarchy of administrative and jurisdictional units, to the material underpinnings of parishes such as church and school buildings, and finally to the most often discussed individual staffing of positions, these sources can rightfully be said to constitute the central node in the network of records produced by ecclesiastical institutions. Any and all further sources that mention priests or parishes from the area under study, regardless of the institutional context in which they were produced or their structures, will be linkable at least to a certain degree to the data currently under definition, as long as they fulfil one major condition. The information they contain referring to either of the main entities in the database (in-

---

22 A complete overview of the constructed variables and standardisation system implemented will be available on the website, with links to the relevant secondary literature that aided in identifying these structural equivalencies.
dividuals or ecclesiastical units) must allow for the identification of correspondences with particular entities in the standard database: priests with very common names, for whom no other contextual information regarding education or parish location is known, will be difficult to link with a high degree of certainty. Likewise, the lack of a clear source timestamp will generally encumber if not wholly prevent linkage of such information to the data systematized in our infrastructure. Nevertheless, even more tenuous ties between database entities and various subjects identified in other researcher-produced datasets can in the future be verified and attributed a higher likelihood of identity, should this be the case.

To conclude, we hope that this will be a useful instrument for social historians not only in Transylvania, but also beyond. We think that the data entered will have immense linkage potential, for instance to the HPDT, but also to other datasets and contexts where priests make their appearance. After we will have included data from other sources, such as records pertaining to priests’ income, to their families, and to their education, a lot more insight will be gained from this. What is more, due to the high level of standardization and the clear explanation of the database structure, we hope to make it useful and usable first of all for social historians and specialists in ecclesiastical history of Transylvania who are not experienced in using digital tools in their analysis.

References


77


Chapter 3


Chapter 4.
MARITAL FERTILITY IN ALBANIA DURING WWI

Siegfried Gruber

Historic background

Albania was a part of the Ottoman Empire until 1912, when it proclaimed its independence from the Ottoman Empire. In the First Balkan War (1912) the neighbouring states of the European parts of the Ottoman Empire, Greece, Bulgaria, Serbia, and Montenegro, declared war on the Ottoman Empire and conquered these European territories except the small part which still belongs to Turkey. Montenegro, Serbia, and Greece wanted to partition the Albanian territory among them and only with the help of the Great Powers (Austria-Hungary and Italy) in 1913 an independent Albania was internationally recognised. During WWI the neighbouring states tried again to get hold of Albanian territories and with neither an international commission nor an Albanian central government exercising control over the whole country, the situation deteriorated rapidly. Encouraged by the successful occupation of Albanian territory by Greece and Italy in southern Albania, the Serbs and Montenegrins invaded northern Albania in early 1915 and in January 1916, the Austrians drove the Serbs and Montenegrins out of their new holdings to occupy Northern and Central Albania as far south as the Vjosa River, which served as a northern border to the Italian holdings.

Fertility in Albania

Albania was, apart from Kosovo, the last country in Europe to enter the demographic transition. After WWII Albania had the highest fertility
in Europe with a total fertility rate (TFR) of about 6.0 children, which
even rose to almost 7.0 children in 1960 (Aassve, Gjonca and Mencarini,
2006: 8) and a subsequent decline to 2.2 children in 2002, still the second
highest in Europe – only in Kosovo fertility was higher (Gjonca, Aassve
and Mencarini, 2008: 261; Falkingham and Gjonça, 2001). My own re-
search has shown that despite a reported average number of 5.2 children
per father (Coon, 1950: 23) the highest average number of children living
in the own household was 2.6 children for ever-married men at an age of
55 years and 2.7 children for ever-married women at an age of 45 years
according to the population census of 1918 (Gruber, 2001: 9). A thorough
analysis of fertility patterns is still missing, while some aspects like urban
fertility patterns (Gruber, 2004) and the influence of migration on fertility
(Gruber, 2011) have already been dealt with.

In the beginning of the 20th century Albania became a major
research area for ethnographers and travellers, who published also
about the marriage patterns in this country. Fertility patterns were
analysed only later in the 20th century, based on fieldwork in 1929/30.
The American anthropologist Carleton S. Coon noted that “in north-
er Albania (Malsia e Gegnisë) girls are married as soon as they
come to sexual maturity and begin bearing children as soon as they
are biologically able. There is no time of peace” (Coon, 1950: 27).
There is not much known about fertility in Albania in the first half of
the 20th century because registration of vital events was quite un-
derdeveloped within the Ottoman Empire: Albania introduced national
vital statistics only in the 1920ies (Rothenbacher, 2013: 168) and it
took quite some time to reach full coverage. Christians (Catholics
in the North and Orthodox in the South) had church books for reg-
istering vital events, but their analysis is still in its infancy. No such
registration was required for Muslims, who constituted the majority
within Albania (about 70 percent). As a consequence, most publica-
tions start with data from 1950 onwards (Gjonca, Aassve and Men-

Statistical information during Communist times was regarded
as a state secret (Falkingham and Gjonça, 2001: 309) and so the first
major publication about the fertility transition in Albania was pub-
lished only in 2001 (Falkingham and Gjonça, 2001). These are the
reasons why Albania appears on the maps of the Princeton European
Fertility Project only about the status in 1960. Overall fertility in
1960 was highest in Albania and neighbouring Kosovo (Coale and
Population characteristics in Albania

The borders of Albania were drawn at the expense of Albania and therefore many Albanians lived outside its borders while ethnic minorities were rather small within its borders. Muslims constituted the majority in Albania (about 70 percent), while there was a Catholic majority in the Northwest of Albania and an Orthodox majority in Southern Albania (mostly outside the territory of the census of 1918).

The educational level of the population of at least 15 years in Albania was still quite low in 1918: only 2.5 percent of the male population in rural areas were literate. The corresponding figure for women was 0.1 percent. The city population had higher shares of literate people: 29.5 percent of men and 6.1 percent of women.

Albania’s economy was still dominated by agriculture: Only 12.6 percent of the population registered in the census of 1918 lived in the six cities having at least 3,800 inhabitants. The largest city, Shkodra, had 23,000 inhabitants, while Tirana, which became after WWI the capital, had only 10,000 inhabitants (Seiner, 1922: 7). Half of the urban population in 1918 were engaged in the production and sales sector. These sectors were dominated by small enterprises and shops in the bazar as traditional crafts dominated while industrial production was only in its very beginnings.

The overwhelming majority of the population lived in the countryside and depended on agricultural activities. The mountainous parts of the country were dominated by animal husbandry while cultivation of land dominated in the valleys and lowlands. In the most fertile areas large estates existed while in the mountains small plots and pastures dominated.

Patriarchal culture and household structures

The Albanian society in the beginning of the 20th century was still highly influenced by a patriarchal culture and so had not been much challenged by the Ottoman Empire. Especially the mountainous parts of the country were only loosely integrated into this empire and the new nation state had troubles in establishing its administration and juridical system in a society which had been shaped by customary law until that time, too. In the beginning of the 20th century Albania (especially Northern Albania) became of major interest to ethnographers as a site

---

1 For more information on the influence of customary law on marriage behaviour see Kera and Papa, 2003: 31–44.
of alleged extreme backwardness and patriarchal values. A comparative analysis of patriarchal features in households (e.g. male household heads and patrilocal marriage) and demography (female age at marriage and preference of sons over daughters) of 266 historical European societies showed that rural Albania in 1918 ranked highest among them (Szoltyszek, Klüsener, Poniat, Gruber, 2017: 241).

More than half of the rural population lived in multiple family households (using the typology proposed by Hammel and Laslett, 1974), while a quarter lived in simple family households.

**Age at marriage**

Most quantitative information about marriage patterns in Albania around 1900 is based on the analysis of the population census of 1918 at the start of the 20th century. The average age at marriage (SMAM: singulate mean age at marriage, see Hajnal, 1953) was 18.0 years among women, and 26.6 years among men according to the census. There were considerable differences among rural and urban Albania, ranging from 17.8 years for women and 26.1 for men in rural Albania and 19.1 for women and 30.1 for men in urban Albania, in Shkodra age at marriage was highest in Albania, 20.0 for women and 34.0 for men (Gruber, 2017: 144). The male age at marriage is therefore higher than postulated for the “Eastern European marriage pattern” (Hajnal, 1965: 101).

The mean age difference between spouses was 9.3 years (also caused by uneven sex ratios) and therefore very high according to European standards. The older the husband the larger the age difference to his wife was (Gruber, 2017: 146). This age gap was even similar to the maximum in an analysis of 77 lesser developed countries in the time period 1950 to 2005 (Carmichael, 2011: 426). In the Southern Albanian region of Mallakastër, 4.2 percent of married men were married to more than one wife (Nicholson, 2006: 48). Overall in 1918 we find 4.4 percent of all married men living with more than one spouse. Polygamy among Muslims was most widespread in the eastern regions of Albania and least common among the urban population (Gruber, 2012: 106).

**The Population Census of 1918**

In the absence of available vital registration data, a pertinent source for this kind of fertility research is the population census conducted by the Austro-Hungarian army in 1918 in Albania (see Nicholson, 1999).
This census is the first for Albania (although not covering the whole territory of present-day Albania) in which the original data is still available on the level of the persons recorded, and it is of high quality given the circumstances under which it was taken (Gruber, 2007: 257). It is still widely unknown, and thus in a demographic atlas of Albania data from 1926 is considered the earliest population data (Bërxfholi, 2003). Gjonça mentions only the preliminary census of 1916, and gives the credit for the first general census conducted in Albania to the 1923 census (Gjonça, 2001: 38f.).

The Austro-Hungarian army occupied the majority of the territory of the newly created independent Albanian state (except parts in the south and southeast), and established a new administration in 1916. Officers of the Austro-Hungarian army collected the census data with the assistance of Albanian officers (Seiner, 1922: 3). The census-takers were instructed to make sure that no persons were excluded from the count, such as female children (Seiner, 1922: 4). These efforts appear to have been successful, since the census counted almost the same number of men and women, whereas in censuses of other countries in the region, there was always a clear male majority in the population (for Serbia see Sundhaussen, 1989: 80).

The research project, “The 1918 Albanian Population Census: Data Entry and Basic Analyses,” based at the University of Graz and funded by the Austrian Science Fund (2000–2003), sought to convert the data into machine-readable form. The census data of 1918 is a rich source for a variety of questions related to studies about population structure and behaviour. Up to now, the data of 285 villages and 6 cities have been entered in a database, which contains 140,611 persons.

**The Child-Woman-Ratio and challenges to its use**

Since the census data does not provide information about the number of children ever born, we need to use the information about the number of children who were alive at the time of the census. Information about the mortality in Albania in the beginning of the 20th century is missing, too. In this paper I use therefore the Child-Woman-Ratio (CWR) to analyse the fertility level in Albania during WWI. It is the number of children under the age of five years divided by the number of women in the childbearing

---

2 [http://www-gewi.uni-graz.at/suedost/seiner/index.html].
3 [http://www-gewi.uni-graz.at/suedost/seiner/availability.html].
age, multiplied by 1000. It is an indirect measure of fertility and biased downwards, because children who died before the census was taken are not included in the calculation (Pullum, 2004: 423).

The first challenge to its use is caused by the date of the census: the census date was March 1st, 1918. The number of children under the age of five years would therefore consist of the months January and February of 1918, the years 1917, 1916, 1915, 1914, and the months of March to December of 1913. The census enumeration form included a column for the date of birth and a column for the age in years. Most people reported their age (92.2 percent) and not their date of birth (only 14.6 percent reported at least the year of birth). Only 8.2 percent of one-year-old infants were reported with their exact birth date. It is therefore not possible to use the months of birth in 1913 with less than one percent of these children having reported their birth date. The analysis is therefore based upon the children born in the calendar years 1913–1917, which correspond to the ages 1-5 years. The census instructions did not specify whether ages should be recorded in completed years or started years. An analysis of people with reported ages and birth years shows that reported ages were generally the difference between the birth year and 1918.

The second challenge is caused by the quality of the age reporting, which is affected by the low level of education respective literacy and to a large extent missing vital registration system. Many people were not reporting exact ages in the census, but have rounded them to multiples of 5 and 10. In an analysis of 115 historic populations within Europe, the highest levels of age heaping were found in areas of present-day Belarus, southern Romania, and Albania in 1918 (Szoltysek, Poniat, and Gruber, 2018: 17). This age heaping was pronounced on final digits 5 and especially 0, therefore we shall use in our analyses age-groups centered around these final digits (e.g. 18–22 years) instead of the generally used age groups in demographic analyses. This will minimize distortions caused by approximated ages, but still cannot eliminate them. All computed ages in this analysis must therefore be treated as approximate ages.

The next challenge is the possibility of missing girls in the census. Coon, who collected data in northern Albania during the autumn, winter and spring of 1929–30, gives a sex ratio of 163.0 at birth (Coon, 1950: 26f.) meaning that there were 163 male to 100 female births. These figures were obviously exaggerated because of a patriarchal bias, although the analysis of census data for the sex ratio of children aged one and younger shows also an uneven sex ratio of 121.5 (Kera and Pandelejmoni, 2008: 132). This sex ratio implies either that there was actually
a major surplus of male births in Albania, or that the mortality rate was higher among female babies, or, more probably, that very young female children were not reported entirely in the census. Seiner’s published results for the whole population of Albania showed an almost balanced sex ratio of 100.6 men to 100 women (Seiner, 1922: 8) which is in contrast to other Balkan countries. The sex ratio at birth should be about 105 (Hobbs, 2004: 133f.), therefore sex ratios below 100 and above 105 will be adjusted to 105 in this contribution.

The final challenge is posed by calculating age-specific fertility rates: there are children which cannot be assigned to their mother, because they do not live in the same household or because the mother is already dead. In addition, there are women who are too old be the mothers of their children. Children without mothers and of mothers, who are more than 52 years older than the reported child will be distributed proportionally to women of ages 15 to 52 for age-specific calculations.

Married women contributed the most important part to general fertility, because births out of wedlock were extremely rare in Albania. As a consequence I shall concentrate in the analyses on married women and specifically on married women whose husband was present at the time of the census. Absent spouses (mainly for occupational reasons – either seasonal or for longer time periods) reduce marital fertility and we want to eliminate this effect in this analysis. I shall apply age standardization in analysing sub-groups of women to compensate for possible different age structures. As standard I shall use the age structure of all married women with spouses present.

Table 1

<table>
<thead>
<tr>
<th>Data used in the analysis</th>
<th>Cases</th>
</tr>
</thead>
<tbody>
<tr>
<td>Overall in database</td>
<td>140,611</td>
</tr>
<tr>
<td>Children 1–5 years</td>
<td>20,997</td>
</tr>
<tr>
<td>Women 15–50 years</td>
<td>37,726</td>
</tr>
<tr>
<td>Women 15–50 years, married, both spouses present</td>
<td>22,132</td>
</tr>
</tbody>
</table>

Source: Gruber, Kaser, Kera, and Pandelejmoni, 2018.
Our first analysis will be to compare the marital fertility in Albania during WWI to the marital fertility of other countries in South-eastern Europe before WWI, because published census data allows to calculate CWR for married women. More elaborated measures like period or cohort total fertility rates are available only from 1950 onwards (Falkingham and Gjonça, 2001: 312). The measured results show that rural Albanian marital fertility (green bars in figure 1) during WWI was lower than in neighbouring countries one or two decades earlier. Especially in Romania in 1899 rural marital fertility was considerably higher. Urban marital fertility (red bars in figure 1) was in all these countries lower than rural marital fertility, only in Albania both were at the same level (without adjustment urban marital fertility was even a bit higher than rural marital fertility). and Albania during WWI had the highest level together with Romania in 1899. Urban marital fertility in Serbia in 1900 was a bit lower and in Bulgaria in 1910 clearly lower than in Albania during WWI. Greek marital fertility was very high among these countries and the published data does not distinguish between urban and rural populations. Therefore we can assume that rural marital fertility must have been similar to the one in Romania in 1899. The adjustment of the Albanian data (see above) does not much influence the results. In all the other countries in figure 1 sex ratios of children up to five years did not exceed a level of 106, therefore no adjustment was needed.

This result can be caused by two different reasons: first, WWI reduced rural marital fertility in Albania, as was generally the case in countries affected by WWI. The second reason could be that pre-WWI rural marital fertility was already lower than in Serbia and Romania around 1900. We would need data about pre-WWI fertility to know which of these two possibilities is more likely.

---

4 These were the last censuses conducted before WWI which have data for calculating CWR and which are available. CWR for married women are slightly inflated because all children are assigned to these married women, although a small fraction of the children belonged to unmarried or widowed mothers.
Differences in marital fertility in Albania during WWI

We have seen in figure 1 that marital fertility in Albania during WWI was almost the same for rural and urban women. This is in contrast to general fertility, which was higher for rural women because of a lower age at marriage for rural women. There was some variation between rural regions ranging from about 750 in Zhuri to 1010 in Berat. Most regions had quite similar values, while the two regions with the highest levels were situated in the south of the census territory. Marital
fertility levels in the six cities with available census data (the data for the city of Berat has been destroyed) show a similar variation from a low of 670 in Durres (a port city) to a high of 917 in Shkodra, the largest city in the census territory.

Religious differences in marital fertility in Albania during WWI were almost negligible in rural as well as in urban areas. There were so few literate women, therefore the analysis of the effect of literacy on marital fertility during WWI is based on the literacy of the husband. Women of literate men had higher fertility than women of illiterate men, which was true for rural as well as urban women.

The next analysis will concentrate on the influence of occupational groups on marital fertility during WWI. This analysis is once again based on the occupational group of the husband because so few women had a reported occupation and those with a reported occupation were doing generally domestic work. The occupational groups are based on the major groups of HISCO (van Leeuwen, Maas, and Miles, 2002), but groups 0/1, 2, and 3 are combined into “white collar workers”. Men without reported occupation of their own are assigned to the occupational group of the household head in case this was a kin related person. In rural areas the majority of men were engaged in agricultural work and therefore all other groups are put into one group. In rural areas women of men engaged in agriculture had slightly higher fertility than men engaged in non-agricultural work. In an urban environment women married to men in the white collar sector had the highest fertility with women married to men in the sales sector were following in the second place. The other four occupational groups had very similar fertility.

Table 2

<table>
<thead>
<tr>
<th>Marital fertility in Albania during WWI, by subgroup</th>
</tr>
</thead>
<tbody>
<tr>
<td>By region</td>
</tr>
<tr>
<td>Rural</td>
</tr>
<tr>
<td>Kruja</td>
</tr>
<tr>
<td>Puka</td>
</tr>
<tr>
<td>Shkodra</td>
</tr>
<tr>
<td>Tirana North</td>
</tr>
<tr>
<td>Zhuri</td>
</tr>
<tr>
<td>Urban</td>
</tr>
</tbody>
</table>

89
<table>
<thead>
<tr>
<th></th>
<th>Rural</th>
<th>Urban</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gora</td>
<td>897</td>
<td></td>
</tr>
<tr>
<td>Tirana South</td>
<td>927</td>
<td></td>
</tr>
<tr>
<td>Berat</td>
<td>1010</td>
<td></td>
</tr>
</tbody>
</table>

By city

<table>
<thead>
<tr>
<th>City</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Kruja</td>
<td>779</td>
</tr>
<tr>
<td>Shkodra</td>
<td>917</td>
</tr>
<tr>
<td>Durres</td>
<td>672</td>
</tr>
<tr>
<td>Elbasan</td>
<td>840</td>
</tr>
<tr>
<td>Kavaja</td>
<td>786</td>
</tr>
<tr>
<td>Tirana</td>
<td>804</td>
</tr>
</tbody>
</table>

By religion

<table>
<thead>
<tr>
<th>Religion</th>
<th>Rural</th>
<th>Urban</th>
</tr>
</thead>
<tbody>
<tr>
<td>Muslim</td>
<td>821</td>
<td>822</td>
</tr>
<tr>
<td>Catholic</td>
<td>867</td>
<td>843</td>
</tr>
<tr>
<td>Orthodox</td>
<td>867</td>
<td>830</td>
</tr>
</tbody>
</table>

By literacy

<table>
<thead>
<tr>
<th>Literacy</th>
<th>Rural</th>
<th>Urban</th>
</tr>
</thead>
<tbody>
<tr>
<td>Illiterate</td>
<td>825</td>
<td>784</td>
</tr>
<tr>
<td>Literate</td>
<td>1012</td>
<td>934</td>
</tr>
</tbody>
</table>

By occupational group

<table>
<thead>
<tr>
<th>Occupational Group</th>
<th>Rural</th>
<th>Urban</th>
</tr>
</thead>
<tbody>
<tr>
<td>White collar</td>
<td>965</td>
<td></td>
</tr>
<tr>
<td>Sales</td>
<td>892</td>
<td></td>
</tr>
<tr>
<td>Service</td>
<td>812</td>
<td></td>
</tr>
<tr>
<td>Agriculture</td>
<td>828</td>
<td>825</td>
</tr>
<tr>
<td>Production</td>
<td>808</td>
<td></td>
</tr>
<tr>
<td>Others</td>
<td>808</td>
<td></td>
</tr>
<tr>
<td>Non-agriculture</td>
<td>780</td>
<td></td>
</tr>
<tr>
<td>Overall</td>
<td>827</td>
<td>818</td>
</tr>
</tbody>
</table>

Source: Gruber, Kaser, Kera, and Pandelejmoni, 2018.
Age-specific marital fertility in Albania during WWI

Finally we shall have a look at age-specific marital fertility in Albania during WWI in comparison to the age-specific marital fertility of other societies. Albanian fertility during WWI was clearly pre-transitional and therefore we can compare it to natural fertility schedules. The concept of natural fertility was introduced by Louis Henry (Henry, 1961a) and refers to “the fertility which exists or has existed in the absence of deliberate fertility control” (Henry, 1961b: 81). Fertility control exists when the behavior of the couple depends on the number of children already born and will change when the maximum of desired children is born (Henry, 1961b: 81).

Hutterite women (married between 1921 and 1930), often seen as the society with highest reported marital fertility in mankind and having the highest fertility in Henry’s study, too (Henry, 1961b: 84), is used as a comparative case. In figure 2 the fertility curves do not show absolute levels of fertility, but relative values with the highest age-specific value of the three curves set at 100. The purpose is to compare marital fertility over the life course and we shall here concentrate only on rural populations. Hutterite marital fertility (a rural population anyhow) shows a steady reduction from the highest levels at age 15–19 and a sharp drop after age 40. This pattern fits quite well to female fecundity. In contrast rural Albanian women during WWI had rather low fertility around age 20 and their highest fertility levels in their mid-30ies\(^5\). Therefore these two fertility patterns are very different from each other. All other societies in Henry’s study have the maximum of marital fertility in the age group 20–24 years (Henry, 1961b: 84). The same is the case with some comparative data from neighbouring countries: Bulgaria in 1905/06, Dalmatia in 1900, and Serbia in 1900 (Gruber, 2019: 170–172).

In searching for a similar pattern to the Albanian one, the pattern based upon rural Norway in the 1865 census is a good candidate. Marital fertility was also quite low around age 20 and the highest fertility level was in the age group 30–34 years. The pattern of Albania during WWI is very similar, but several years later in life. This is somehow astonishing, because Albania and the Hutterites in the early 20\(^{th}\) century were societies with low female ages at first marriage, while Norway in 1865 had a high female age at first marriage.

\(^{5}\) Albanian age groups are centered around final digits 0 and 5 and therefore the curve in figure 2 is 2.5 years moved to the right.
This result is in complete contrast to the above-mentioned citation by Coon, who wrote that “in northern Albania (Malsia e Gegnisë) girls are married as soon as they come to sexual maturity and begin bearing children as soon as they are biologically able. There is no time of peace” (Coon, 1950: 27). Coon’s remark is in line with a general reasoning that early marriage is connected to the desire of having many children and especially to have a male heir as soon as possible in life. But the data of the census of 1918 does not support this concept.

Figure 2: Age-specific CWR of married women in rural Albania 1918 compared. Source: Hutterites 1921–30: Wetherell, 2001: 593.


Conclusion

The Albanian population census of 1918 can be used for the Child-Woman-Ratio to calculate an indirect measure of fertility. There are some challenges to the application of this method, mainly because of
some data quality problems. Rural marital fertility in Albania during WWI was lower than in neighbouring countries one or two decades earlier, while urban marital fertility was higher than in most of the neighbouring countries. Marital fertility was almost the same for rural and urban areas, while there was some variation between rural areas and between cities. Religion had no obvious influence on marital fertility while literacy of the husband increased marital fertility. Agriculturalists had the highest rural fertility while people in the white collar sector had the highest urban fertility. Age-specific marital fertility showed its peak only around age 35, which is quite unexpected and does not fit the course of age-specific fertility of Hutterite women. Rural women in Norway 1865 on the other hand had a quite similar pattern to the Albanian ones. The reason for the very low fertility of young married women is still an open question. Possible hypotheses are: (1) This is an effect of WWI. (2) This is an effect of the very young age at marriage and the bad situation of young married women (much work and not enough food), so that their fecundity was affected.

Acknowledgement

Research for this publication has been supported by funds of the Oesterreichische Nationalbank (Oesterreichische Nationalbank, Anniversary Fund, project number: 17494).

Published Sources


References


Chapter 4


**Data**


The Digital Archive (The National Archive), Norwegian Historical Data Centre (University of Tromsø) and the Minnesota Population Center. *National Sample of the 1865 Census of Norway, Version 2.0*. Tromsø, Norway: University of Tromsø, 2008.
Chapter 5.
ENTERING THE REPRODUCTIVE
PHASE OF LIFE: FIRST MARRIAGES
IN ZSÁMBÉK, HUNGARY
(1720–1945)¹

Péter Őri

Introduction

Family is a social institution ensuring the demographic reproduction of populations. It starts with the creation of a stable partnership, in the past almost exclusively a marriage according to prevailing norms. Thus, this ritual generally began the demographic reproduction in historical societies, at least in Europe where the first marriage and first birth were close in time (Lundh, Kurosu, 2014). Consequently, marriage customs had a direct impact on population development. The frequency of marriages and the time spent in marriage significantly affected fertility. The percentage ever married was closely related with the level of fertility, whereas the timing of marriages affected marital fertility by determining the length of women’s reproductive phase together with the frequency and timing of widowhood and remarriages.

Besides, marriage can be considered one of the key turning points in the life course. Even if it is not identical with becoming adult or leaving the parental household for a new with more independence, it is a considerable step in this direction.

In this paper we analyse the frequency and timing of first marriages (and to some extent the characteristics of first births) in the Hungarian village (Zsámbék) between the 18th century and WWII. Our goal is to better understand the determinants of first marriage at the individual level and in the longer run. In order to achieve this goal, we use longitudinal family reconstitution data from the Roman Catholic parish registers of the village, and analyse these data using event history models (Cox regression; Cleves, Gould, Gutierrez, 2004).

¹ This paper has been based on research supported by the Hungarian National Research, Development And Innovation Office (NKFI, project no. 113100, Fertility Transition from Micro- Perspective in 19th–20th Century Hungary).
Former research on marriage

Research on marriage and household formation have been a central field of family history and historical demography since the 1960s. Therefore, the literature on the topic is extremely rich, and a survey cannot be the task here. Besides some general statements on international literature, we shall concentrate on marriage customs and Hungarian research findings.2

Research on household structure and family history has been dominated by the works of John Hajnal and Peter Laslett since the 1960s (Hajnal, 1965, 1982, Laslett, 1972, Laslett, 1983, 1988). Regarding marriage, mostly the mean age at first marriage and the share of the never married were studied using population censuses, parish records and vital statistics as data sources. These researchers considered the western marriage pattern (late marriage and a high proportion of final celibacy) a unique model in history worldwide. This model reflects a certain individual and rational decision making with respect to marriage and reproduction, playing a key role in moderating childbearing. In other words, these norms functioned as preindustrial birth control. Such demographic behaviour and the norms and values behind them had far reaching consequences as they were linked to the development of capitalism or modernity.

Even though modified and refined (Laslett, 1983), the model’s static approach and simplifying classification remained while being strongly criticised.3 Because of the creation of the ‘West-East’ model, a great number of case studies and micro-level analyses were carried out in the last decades, revealing the complexity of European marriage customs and household formation rules. Naturally, proving the existence of different models weakens the simplifying binary Hajnal-model. Stressing economic and social factors in marriage and household formation

---

2 In this respect, we quote the work of Tamás Faragó (2003) which provides a very useful survey of the international research up until the millennium. Also see: Oris – Ochiai 2002. For more recent summaries see: Gruber, Szoltysek, 2012 and Szoltysek, 2012. A newer attempt to survey the research field is Őri, Pakot, 2014. The two volumes of basic importance published recently also provide us with detailed summaries of research on family and household history (Lundh, Kurosu et al., 2014, Szoltysek, 2015).

3 See for instance Fertig 2003: a theoretic criticism and the use of event history models in analysing first marriage in 19th century Germany. A detailed picture of the regional models of marriage and household formation can be found in Szoltysek, 2015, p. 41–85. Also see: Lundh, Kurosu, 2014c.
offers a more complex model than the explanation based on mechanic geographic division which dominated Hajnal’s and Laslett’s works. In addition, a dynamic approach to the understanding of how households worked and how their structure changed over time or during the household life-cycle, got more stress during the last two or three decades. Thus, the timing of marriage, migration prior to or related directly to marriage, the social and economic determinants of marriage, the effect of household composition on the chances of marrying were increasingly emphasized besides comparisons by region or culture. This process was accompanied by the emergence of new historical sources and new statistical techniques for demographic analysis.

The creation of large, individual level and longitudinal databases and the use of event history analysis based on population registers or family reconstructions have brought new results while intercultural comparisons also got a new impetus. In this respect the EurAsia project’s volume on marriage is particularly worth considering here (Lundh, Kurosu et al., 2014). It shows that economic and social factors, as well as household composition impacted the incidence of marriage and its timing, as well as outmigration. At the same time, it has become clear that pure differences e.g. in the timing of marriage between regions and cultures can hardly be interpreted without understanding the whole demographic system of a given population, or the role of marriage within it. Thus, the timing of marriage might have differed a lot in Europe and Asia, but the timing of reproduction (the age at first childbirth) proved surprisingly similar. It can be explained by the different meaning of marriage: in Europe it was the beginning of the reproductive phase in the life of couples, while in Asia it was only a step towards reproduction, resulting in early marriage (Lundh, Kurosu, 2014b, p. 443).

At the same time, the use of more traditional methods and sources related to regions that have been hardly studied before also brought exciting new results. In this respect, we have to mention the work of Mikolaj Szoltyszek (2015), which reinterpreted this regional approach of marriage and household formation showing new subtypes and boarder lines and modifying the predecessors’ simplifying picture. The Mosaic database (Szoltyszek, Gruber, 2016) related to continental Europe might also bring significant new results for a vast region that was much less studied so far.

The Hungarian research literature on the topic is less rich but still too large to be detailed here. Various authors have emphasized that historical Hungary – characterized by geographic, economic, social and
cultural heterogeneity – cannot be classified by a dichotomous model as suggested by Hajnal some fifty years ago (Hajnal, 1965, 1982). This is in particular true for household formation rules, but to some extent for marriage customs too. Age at first marriage in Hungary was indeed low in the 18th century, but ethno-cultural factors have played much less role in this regard than socioeconomic factors, local ecotypes or the type of settlement (urban versus rural) (Faragó, 1998, 2003; Őri, 2009, 2016). Later in the 19th century, age at first marriage increased among males (in 1869 around 26), while that of females remained low (in 1869 around 21) (Őri, Pakot, 2014, p. 20). In the first half of the 20th century males married on average at 26–27, while females married at about 23 (Csernák, 1997, Őri, 2016). After WWII, in the era of socialist modernization, the trend of age at marriage reversed and decreased until the 1980s (Csernák, 1997). To sum up, early and general marriage and forming complex, multigenerational households cannot be regarded as characterizing Hungary as a whole, regardless of sex, socio-occupational status, place of residence or time period. Particularly females and landowning peasants married young, and the latter remained longer in the parental households, but other social groups (urban populations, rural artisans, servants, farm hands and intellectuals) married later, with oftentimes considerable differences (Faragó, 2000, Őri, 2016). Keeping these differences and time changes in mind, several scholars have argued that Hungary had some sort of medium position between the ‘eastern’ and ‘western’ marriage patterns and household formation models (Andorka, Faragó, 1983; Faragó, 2003).

**Historical sources in Hungary suitable for studying marriage**

The sources of historical demography resemble those in other European countries. Household lists, population censuses and ecclesiastic records of marriages together with civil marriage records are suitable for analysing the timing and the frequency of marriages or remarriages. Longitudinal family reconstitution databases created by using parish records and civil registrations can reveal changes in marriage customs and make possible the analysis of the determinants of marriage with event history models.

Cross-sectional census-like sources appeared in greater numbers in the 18th century. Roman Catholic *Status Animarum* should have been composed by the local parsons every year. Similarly, Protestant clergymen created such lists for their communities from the 18th century
onwards. However, most of the enumerations were not made originally or were lost in later periods. In most cases, we have household lists only for one or some consecutive years. These sources allow the calculation of *singulate mean age at first marriage* – *SMAM* (Henry, Blum, 1988, p. 34) and the rate of the never married in different age groups.\(^4\) Longer series have been preserved only exceptionally, among which the lists of the Roman Catholic Zsámébk village can be considered the richest regarding content and length, especially complete from the 1795 to 1867. Then, besides the longitudinal analysis of marriage, we can study other relevant topics: service, the leaving of parental household, mobility within the villages or out-migration from the community, post-nuptial residence, etc. The database from this source is still under construction, with links to the parish records. Such a complex database containing both demographic events and household composition will allow statistical analysis of diverse demographic phenomena in a rich context. So far, we can use only the data of the parish register (family reconstitution data) in the present analysis.

The nominative manuscripts of later full count censuses did not survive except for the censuses 1850, 1857 and 1869 from which much was preserved, plus the representative 25% sample of the census 1970. These sources make the analysis of marriage possible in a rich context (timing, celibacy rate, determinants of marriage, the frequency of service). From the census 1869 a representative sample has been created within the Mosaic-project.\(^5\) Aggregate level data for the counties were published by the Hungarian Statistical Office from 1870 onwards, and the unpublished material at the settlement level can be found in the Hungarian National Archive from 1880 onwards. These summary tables make possible the calculation of SMAM or celibacy rates for each age group for every settlement by decade.

The other possible sources for studying marriage are the parish records containing the vital events from the middle of the 18th century and the civil registers from 1895 onwards. Besides the marriage registers with age, place of birth or residence, profession, religion, parents and witnesses, the data linked between the three types of registers in a family reconstitution database allow the analysis of individual demographic life courses in the context of the sources. Not only basic

---

\(^4\) For the 18th century Tamás Faragó has recently created a larger digitalized database from local ecclesiastic household lists. See: Faragó, 2016.

\(^5\) About Mosaic see: Szołtysek, Gruber, 2016. About the Hungarian sample and the results of analysis of household structure see: Őri, Pakot, 2014.
marriage differentials can be calculated and analysed, but also the determinants of marriage or the setting of the marriage in the life-course can be studied by using multivariate statistical methods such as event history analysis (Kurosu, Lundh, 2014, p. 66). Using purely parish registers for this purpose means the lack of household level contextual data in the analysis, but the exceptional combination of the registration of demographic events and continuous series of household lists provide us with this kind of information. Moreover, in this case the problem of migration can be controlled since the continuous (yearly) enumeration of household members makes the timing of migration calculable. For Zsámbék such analysis will become possible but – as mentioned above, database building and data linkage is not ready yet. In this paper, we exclusively use family reconstitution data for analysing marriage customs where we can calculate some variables of family context (the life and death and ageing of parents or birth order) but we have inadequate data on the composition of families or households, or on the co-residence of family members.

The community under study

Zsámbék is close to Budapest (circa 30 kilometres westward from the capital) in a small basin surrounded by low hills. It was a local centre with rural economy, its legal status as “market town” with three to four thousand inhabitants⁶) reflected the local importance. The status of market town meant that the settlement was owned by a landlord (the royal chamber in this case), its inhabitants lived under the landlord’s jurisdiction but enjoyed some autonomy even in the feudal era prior to 1848. They used their plots relatively freely, they paid taxes once a year, and the inheritance of plots and marriage and to some extent migration were also free. After the abolishment of serfdom in 1848, former serfs became the owners of their rented plots.

Roman Catholic German settlers repopulated the village after the Ottoman wars in the early 18th century. It remained mainly German (although the share of Hungarians increased in the period studied) and it was Roman Catholic even in the first part of the 20th century. The Hungarian population living in the settlement were also Roman Catholics,

---

⁶ The population size of the settlement was around 2500 inhabitants at the end of the 18th century. The number increased to 3500 by the middle of the 19th century, reached 4000 at the turn of the century and was around 4500 according to the census 1941 (Őri, 2014, p. 219).
and a considerable Jewish community also settled there. The population mostly lived on agriculture (more than 60% even in the inter-war period), specializing in wheat production, wine production (especially in the 19th century) and animal husbandry, and the village possessed a considerable part covered by forests. In 1900, the agricultural population consisted of equal parts of smallholders and agricultural workers, most of whom were day labourers). The relative closeness of Budapest was a decisive factor: the capital offered an easily available market for agricultural products, had a perpetually increasing labour-force demand and thus remained the most important destination for out-migrants. But the region was full of smaller or larger villages inhabited by Germans and geographic mobility to these neighbouring villages (related mainly to marriage) was also significant.

Population growth was more or less continuous until WWII, in spite of strong out-migration and high infant and child mortality even at the turn of the century. The main factor behind the population growth was high fertility, which also characterized other Roman Catholic German communities in the region. Besides favourable opportunities for out-migrants to find life partners and jobs in the surrounding area, the stem inheritance system prevailing among German settlers (Husz, 2002) might have supported high and uncontrolled marital fertility together with the custom of early and common marriage.

Goals of the analysis, data and methods

In this paper we study the timing and intensity of first marriages in Zsámbék over the period of 1720–1945. We regard first marriage as a key momentum in the individual life course, closely connected with entering adulthood and the reproductive phase both among males and females. Thus, we do not consider the marriage act itself but concentrate on its determinants and relationship to childbearing in and before marriage. We would like to explore the long-term changes in the age at first marriage, the role of socio-economic status and to some extent the effects of family context (the ageing and death of parents or birth order of the individual concerned) and migration. In this way, we hope to better understand the circumstances under which decisions related to marrying were made, instead of the pure registration of early or late marriage.

Serfs on their rented but inheritable plots owned by the Royal Chamber, who became the owners of their rented plots after the abolishment of serfdom in 1848.
Using the genealogical book (*Ortsfamilienbuch*)\(^8\) of the Roman Catholics living in Zsámbék has facilitated the process of family reconstitution (Gallina, Jelli, 2002). Based on the Roman Catholic parish registers of the village, the book contains linked demographic data for all Roman Catholic parishioners, both German and Hungarian. The book has been turned into a complete family reconstitution database, the data of which have been checked and corrected as well as completed with missing information on socio-economic status.

In our analysis we deal with males and females separately considering that they differed not only in the timing and frequency of marriage but also in post-nuptial residence (females mostly left the parental household when marrying while some married men stayed with their parents). Because of the nature of our data and the differing age at first marriage, the samples of males and females considered here are not entirely comparable. First of all, individuals born in Zsámbék with no further information were omitted. Since we do not know their date of leaving, we do not know until when they belonged to the population at risk of marrying. We followed unmarried individuals after the age of 18 in the case of males and after 15 in the case of females. We concentrated on the length of time until marriage or other event which could close the observation. This was if they died unmarried, in the case of females alternatively their first birth out-of-wedlock if we had no further information on them. When someone gave birth to an illegitimate child, it was certain that she had been present until the event was recorded by the priest. Given no further information, she might have out-migrated after the birth. In these cases, we closed the observation one year after giving birth. Thus, the sample of females is a little larger and makes the combined study of marriages and first births possible. Besides the entire sample of males and females that we consider, we analyse separately the sample of those who married in the village. In the latter case, we obviously focus on the timing of first marriage, whereas in the former, the frequency of marriages also counted to some extent. First marriages above the age of fifty are not considered, we close the observation at that age if one of the above mentioned events had not taken place earlier.

First, we present some statistics on ages at first marriage and first birth in and outside marriage, together with the distribution of women at marriage or the end of observation by their status with respect to first pregnancy and childbirth. Then we use event history analysis to

\[^{8}\] The description and evaluation of this kind of sources can be found in Imhof, 1976 and Knodel, Shorter, 1976. A well-known example for the use of the source is Knodel, 2002.
study the timing and the determinants of first marriage.\(^9\) We study the duration of time since the 18\(^{th}\) or 15\(^{th}\) birthday up to marriage or the end of observation with Cox continuous-time proportional hazard models (Cleves, Gould, Gutierrez, 2004, p. 121–157). Our co-variates are the period of observation, place of birth (to include the role of migration), marriage number of the spouse in the models of marrying persons, socio-economic status by HISCLASS categories (van Leeuwen, Maas, 2011), fathers’ and mothers’ ageing and deaths, and birth order related to the same sex. Giving illegitimate birth was also a co-variate in the case of women.

**Hypotheses**

Based on former research, we suppose that the likelihood of first marriages will decrease over time as the age at first marriage increases. Migration before marriage may also lessen the likelihood by delaying marriage, and similarly if a man married a widow. On the contrary, if a woman married a widower who might have been older than was generally the case of first marrying persons, it may have affected to a smaller extent the age of the bride at first marriage (in such cases marriages with larger age difference were usual). We expect that profession or social status had considerable impact on the timing and frequency of marriages, but the effect’s direction is hardly predictable. Persons with higher social status had a better position in the marriage market, rented or real land property also resulted in a better position and required the future spouse’s labour force. Since male heirs might live with the parents after marriage, their marriages were likely to be early. Daughters of landowners left the parental household after marrying, thus their marriages were also early and general. At the same time, non-heirs might remain in the parental household longer, enter into service or out-migrate, events probably delaying their marriage. The landless were in a weaker position in the marriage market, their households were less stable, since they probably left the parental household earlier and more frequently, migrating with a higher probability than the more well-off. Artisans were also more mobile which might have delayed their marriage but this could have less effect on their daughters’ age at marriage.\(^{10}\)

\(^9\) About the use of family reconstitution data in event history analysis see: Gutman, Alter, 1992.

\(^{10}\) About the factors affecting the timing and frequency of marriages see Lundh, Kurosu, 2014a.
We expect that the richer married earlier, but these differences among men may be lessened by impacts of different direction while age differences among women must have been smaller than among men. The ageing of parents probably promoted the children’s marriage since they also grew older and retirement among landowners could ease the heirs’ marriage. The same was true for the death of parents, opening the way towards heritage, causing the dissolution of the parental household or entering widowhood, thus needing to renew the labour force, which the marriage of a child could fulfil. On the other hand, parental death might result in increased poverty, a worse position in marriage market, or the widowed parent might request the work and company of children. Moreover, a competition can be imagined between widowed parents and unmarried children with respect to marriage. Birth order could also affect marriage not through the timing of marriages but through leaving their parental household, or the non-marrying of younger children (especially the non-heirs) if elderly parents needed them at home. These possible effects often seem contradictory, making it difficult to estimate the results of the multivariate statistical analysis in advance.

**First marriages in Zsámbék**

*Descriptive statistics*

Between the 18th century German in-migration and WWII we can study more than five thousand first marriages (Table 1) taking place in the village, and included in the local Roman Catholic parish register with an exact date. Men married on average around the age of 25 while women at the age of 21–22. In this respect the custom of early marriage (the so-called “Eastern” marriage pattern) seems to have existed in the case of both sexes. Moreover, the changes over time do not appear considerable. In the first third of the 19th century age at first marriage rather decreased, from 1840 onwards the mean age of men began increasing and it rose above 25 after 1880. Among women we witness stagnation after 1840, and a considerable increase can be observed only after 1900. An increasing trend among men fits into the picture we have of 19th century Hungary (Faragó, 2000), while decreasing ages in the beginning

---

11 Hajnal himself drew the border line of Western marriage pattern at 26 (men) and 21 (women) years of age (Hajnal 1982, 452). In this respect Zsámbék (25 and 21.5 years of age respectively) can be regarded as a transitory model but we would like to avoid here this kind of detailed classification.
Chapter 5

of the century cannot be interpreted so easily. The first two decades of the century were characterised by war and inflation, but were favourable for food producers, both noble landowners and their tenants, which might explain the decreasing ages at first marriage. After the Napoleonic wars the economic decline and after 1840 the new crisis can be a good explanation of increasing ages at first marriage. We also know that in Zsámbék in the first 30 years of the 19th century, the household structure became increasingly complex. Thus, leaving the parental home for non-heirs became difficult, and the situation changed only after the cholera epidemic in 1831 and during the growing labour force demand from the 1830s (Husz, 2002). Thus, the longer stay in the parental household and the more complex household structure went together with a somewhat lower age at first marriage, but this problem must be studied by involving the entire data set of the local Status Animarum.

Table 1

<table>
<thead>
<tr>
<th>Decade</th>
<th>Men</th>
<th>Women</th>
<th>Women in marriages with child</th>
<th>Women’s age at 1st pregnancy in marriage</th>
<th>Women’s age at illegitimate births</th>
</tr>
</thead>
<tbody>
<tr>
<td>-1799</td>
<td>24.8</td>
<td>21.9</td>
<td>21.2</td>
<td>22.0</td>
<td>31.9</td>
</tr>
<tr>
<td>1800-1819</td>
<td>23.4</td>
<td>21.1</td>
<td>20.0</td>
<td>20.7</td>
<td>22.0</td>
</tr>
<tr>
<td>1820-1839</td>
<td>23.9</td>
<td>21.9</td>
<td>21.1</td>
<td>21.7</td>
<td>21.3</td>
</tr>
<tr>
<td>1840-1859</td>
<td>24.6</td>
<td>21.5</td>
<td>20.8</td>
<td>20.3</td>
<td>22.1</td>
</tr>
<tr>
<td>1860-1879</td>
<td>24.9</td>
<td>21.4</td>
<td>20.8</td>
<td>20.9</td>
<td>21.8</td>
</tr>
<tr>
<td>1880-1899</td>
<td>25.4</td>
<td>21.4</td>
<td>20.9</td>
<td>20.5</td>
<td>21.9</td>
</tr>
<tr>
<td>1900-1919</td>
<td>25.5</td>
<td>22.4</td>
<td>21.9</td>
<td>21.1</td>
<td>19.7</td>
</tr>
<tr>
<td>1920-</td>
<td>25.3</td>
<td>22.1</td>
<td>21.6</td>
<td>20.6</td>
<td>20.1</td>
</tr>
<tr>
<td>Mean</td>
<td>24.8</td>
<td>21.5</td>
<td>21.1</td>
<td>21.1</td>
<td>21.6</td>
</tr>
<tr>
<td>N</td>
<td>5004</td>
<td>5526</td>
<td>3975</td>
<td>3975</td>
<td>633</td>
</tr>
</tbody>
</table>

Source: Family reconstitution database of Zsámbék (Gallina-Jelli 2002 and parish registers), own calculation

Considering marriages where at least one child was born, the dynamics of ages at first marriages are very similar to those of all marriag-
es, but the ages themselves were a bit lower. These women remained in the village after the wedding and most married local bridegrooms. This result confirms our assumption that migration related to marriage somewhat delayed the event.

First pregnancies in marriages or just before marriages when the father was known and identical with the later husband were similar to first marriages; the age difference was minimal. Before 1840 conception immediately followed the wedding, after this date the events followed each other closely. From 1880 onwards, first conceptions seem to have preceded first marriages, clearly showing some kind of trial marriage and changing norms in the village. Ages at first illegitimate birth show that in the 18th century illegitimate birth was characteristic of women who did not marry or married much later than the others, whereas in the 19th and 20th centuries the women giving birth outside marriage were young.

<table>
<thead>
<tr>
<th>The distribution of women at the end of observation</th>
</tr>
</thead>
<tbody>
<tr>
<td>All women</td>
</tr>
<tr>
<td>Childless, non pregnant</td>
</tr>
<tr>
<td>With child</td>
</tr>
<tr>
<td>Pregnant</td>
</tr>
<tr>
<td>With child, pregnant</td>
</tr>
<tr>
<td>Total</td>
</tr>
<tr>
<td>N</td>
</tr>
</tbody>
</table>

12 The number of first marriages with at least one child is considerably lower than that of all marriages because of the exogamous marriages. In these cases, the wedding and the registration of the marriage took place in the village of the bride although the new couple began their common life in the bridegroom’s place of residence. Thus, one part of the first marrying women did not give birth in their former place of residence.
### Women getting married

<table>
<thead>
<tr>
<th></th>
<th>1799</th>
<th>1800-1819</th>
<th>1820-1839</th>
<th>1840-1859</th>
<th>1860-1879</th>
<th>1880-1899</th>
<th>1900-1919</th>
<th>1920-1939</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Childless, non pregnant</td>
<td>84.0</td>
<td>84.9</td>
<td>75.4</td>
<td>76.6</td>
<td>68.0</td>
<td>68.2</td>
<td>67.8</td>
<td>66.4</td>
<td>74.7</td>
</tr>
<tr>
<td>With child</td>
<td>0.0</td>
<td>0.2</td>
<td>0.5</td>
<td>0.8</td>
<td>2.9</td>
<td>2.1</td>
<td>2.7</td>
<td>1.6</td>
<td>1.2</td>
</tr>
<tr>
<td>Pregnant</td>
<td>16.0</td>
<td>14.9</td>
<td>24.0</td>
<td>22.0</td>
<td>27.0</td>
<td>28.7</td>
<td>28.3</td>
<td>31.5</td>
<td>23.5</td>
</tr>
<tr>
<td>With child, pregnant</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.5</td>
<td>2.0</td>
<td>1.0</td>
<td>1.3</td>
<td>0.5</td>
<td>0.6</td>
</tr>
<tr>
<td>Total</td>
<td>100.0</td>
<td>100.0</td>
<td>100.0</td>
<td>100.0</td>
<td>100.0</td>
<td>100.0</td>
<td>100.0</td>
<td>100.0</td>
<td>100.0</td>
</tr>
<tr>
<td>N</td>
<td>1288</td>
<td>498</td>
<td>558</td>
<td>595</td>
<td>588</td>
<td>487</td>
<td>630</td>
<td>882</td>
<td>5526</td>
</tr>
</tbody>
</table>

### Women who died unmarried or outmigrated after giving birth outside marriage

<table>
<thead>
<tr>
<th></th>
<th>1799</th>
<th>1800-1819</th>
<th>1820-1839</th>
<th>1840-1859</th>
<th>1860-1879</th>
<th>1880-1899</th>
<th>1900-1919</th>
<th>1920-1939</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Childless, non pregnant</td>
<td>93.1</td>
<td>85.2</td>
<td>64.5</td>
<td>66.7</td>
<td>58.4</td>
<td>60.8</td>
<td>63.2</td>
<td>67.4</td>
<td>72.9</td>
</tr>
<tr>
<td>With child</td>
<td>6.9</td>
<td>14.8</td>
<td>35.5</td>
<td>33.3</td>
<td>41.6</td>
<td>39.2</td>
<td>36.8</td>
<td>32.6</td>
<td>27.1</td>
</tr>
<tr>
<td>Total</td>
<td>100.0</td>
<td>100.0</td>
<td>100.0</td>
<td>100.0</td>
<td>100.0</td>
<td>100.0</td>
<td>100.0</td>
<td>100.0</td>
<td>100.0</td>
</tr>
<tr>
<td>N</td>
<td>449</td>
<td>81</td>
<td>93</td>
<td>117</td>
<td>166</td>
<td>194</td>
<td>234</td>
<td>288</td>
<td>1622</td>
</tr>
</tbody>
</table>

*Source: Family reconstitution database of Számbék (Gallina-Jelli 2002 and parish registers), own calculation*

Table 2 displays the distribution of women at the end of observation (marriage, death or out-migration after giving an illegitimate birth). The share of pregnant brides almost doubled during the studied period. Since most married after conception, the same holds true for marrying women. The percentage with an illegitimate child at the end of the observation (at marriage or before out-migration) also rose until the end of the 19th century, when a slight decrease followed in the first decades of the 20th century. Altogether mothers formed a
modest proportion of marrying women (the father of their illegitimate child was unknown, but naturally he could be the later bridegroom), their majority left the village after giving birth unmarried. The share of those who gave birth and were pregnant before marriage (here the father and the bridegroom could also be the same person) was also insignificant, although the dynamics of changes were the same as in the case of premarital pregnancy and childbearing. In those cases, where we closed the observation before marriage, its cause was mainly the death of the observed individual, but during the 19th century the role of illegitimate childbearing and out-migration continuously increased as mortality decreased and the frequency of illegitimacy rose. To sum up, we can conclude that almost one third of the women getting married were pregnant by the end of the 19th century. The rate of illegitimacy also increased, peaking between 1860 and the end of the Great War when about 2–3% of the marrying women had at least one illegitimate child (whether with the later husband, we do not know). Based on this, we may assume the existence of trial marriage (pre-nuptial sexuality in order to get certainty about the fecundity of the planned marriage) and changing norms which became more tolerant by the 20th century. First marriage was the entry into the reproductive phase of life as it was elsewhere in Europe (Lundh–Kurosu 2014b). First marriage and first birth were very close in the life course of women, sexual relations were not excluded before marriage as the fecundity of marriages was crucial. Childbirth outside wedlock was not rare either, it did not exclude later marriage (assumingly the father married the mother with a common illegitimate child in most cases) but a considerable part of the unmarried mothers left the community and their further life course remains hidden because of the nature of the source we used.

**Event history analysis**

We shall now analyse the impact of several covariates on the timing and – in part – the frequency of first marriages. For each sex we created two models: people who married in the village and the complete sample (marrying persons, those who died unmarried and women who gave birth to an illegitimate child and had no further event in the parish book). In each case we measure the impact of a covariate as compared to a reference category (1). Results are displayed as hazard ratios, values above 1 mean higher likelihood (first of all earlier marriage) while values under 1 show lower likelihood (basically later marriage). Thus,
an increasing hazard ratio means a higher likelihood to marry, and in other words a lower age at marriage.

Among men the two samples show no significant differences (Table 3). The likelihood of first marriages decreased over time (men married later), only in the first two decades of the 19th century some increase appeared among those getting married (earlier marriages). We have already witnessed this decrease in the marriage age when studying descriptive statistics. In the second model we cannot see this change, maybe due to rising mortality counterbalancing the effect of earlier marriages as more young men died unmarried in this period. This decreasing trend of the hazard ratios stopped only after 1920 when they started to increase modestly. This means that after the war period, there was a slight decrease in the age at first marriage.

Place of birth mattered little; prenuptial migration had no significant effect on the timing of marriage. A different place of birth naturally could mean migration at any age before marriage. It would have been better to include only moves directly preceding the wedding, but our family reconstitution data do not make it possible.

If a first marrying man chose a remarrying widow, this factor significantly delayed the age at marriage. In these cases, the wives were probably older than the grooms or than first marrying brides. Thus, men marrying first time to widows were themselves also relatively older as compared to other men getting married first time.

Socio-occupational status had a strong and significant effect on first marriages: landowners (landowners’ heirs) married the earliest while day labourers and artisans married later. Landowners and their heirs were attractive partners in the marriage market, they could more easily cover the wedding expenses and there were fewer obstacles for their early marriage in the stem family system. Artisans were more mobile, day labourers were poorer and they did not need early marriage to maintain family property as landowners did.

Ageing parents (after 50th birthday) and their death significantly increased the likelihood of marriage. As expected above, a father’s death could open the way to inheritance, independence and/or household headship for male children. Similarly, reaching age 50 might have meant the fathers’ retirement, triggering the marriage of sons. The ageing and death of mothers oftentimes resulted in the dissolution of the parental household which also accelerated the sons’ marriages.

Finally, birth order did not affect considerably the timing of first marriage. Maybe a more complex model where the marital status and
number of brothers and sisters are included will show larger effects in future studies. Another fact that we need to take into account is that the percentage of the missing cases is the highest when the analysis includes family relationship variables. In many instances we do not know the birth date of parents, their age and/or death date, etc., making our results uncertain.

Table 3

The likelihood of first marriage among men, Zsámbék, 1720-1945

<table>
<thead>
<tr>
<th></th>
<th>Men getting married</th>
<th>All men</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Hazard ratio</td>
<td>%</td>
</tr>
<tr>
<td>Periods</td>
<td></td>
<td></td>
</tr>
<tr>
<td>-1799</td>
<td>1</td>
<td>20.3</td>
</tr>
<tr>
<td>1800-1819</td>
<td>1.151</td>
<td>* 8.8</td>
</tr>
<tr>
<td>1820-1839</td>
<td>0.881</td>
<td>† 10.2</td>
</tr>
<tr>
<td>1840-1859</td>
<td>0.762</td>
<td>*** 11.3</td>
</tr>
<tr>
<td>1860-1879</td>
<td>0.714</td>
<td>*** 11.3</td>
</tr>
<tr>
<td>1880-1899</td>
<td>0.595</td>
<td>*** 10.2</td>
</tr>
<tr>
<td>1900-1919</td>
<td>0.539</td>
<td>*** 12.9</td>
</tr>
<tr>
<td>1920-0.613</td>
<td>*** 15.0</td>
<td>0.394</td>
</tr>
<tr>
<td>Place of birth</td>
<td></td>
<td></td>
</tr>
<tr>
<td>local</td>
<td>1</td>
<td>72.6</td>
</tr>
<tr>
<td>foreigner</td>
<td>0.906</td>
<td>† 15.5</td>
</tr>
<tr>
<td>missing</td>
<td>0.689</td>
<td>*** 12.0</td>
</tr>
<tr>
<td>Marriage number of wife</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>2</td>
<td>0.672</td>
<td>*** 4.8</td>
</tr>
<tr>
<td>3</td>
<td>0.595</td>
<td>* 0.3</td>
</tr>
<tr>
<td>4</td>
<td>1.875</td>
<td>0.1</td>
</tr>
<tr>
<td>SES</td>
<td></td>
<td></td>
</tr>
<tr>
<td>landowner</td>
<td>1</td>
<td>23.0</td>
</tr>
<tr>
<td>day-labourer, worker</td>
<td>0.718</td>
<td>*** 38.8</td>
</tr>
<tr>
<td>artisan</td>
<td>0.686</td>
<td>*** 16.0</td>
</tr>
</tbody>
</table>
## Chapter 5

<table>
<thead>
<tr>
<th></th>
<th>Men getting married</th>
<th>All men</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Hazard ratio</td>
<td>%</td>
<td>Hazard ratio</td>
<td>%</td>
</tr>
<tr>
<td>intellectual</td>
<td>0.537</td>
<td>***</td>
<td>1.6</td>
<td>0.629</td>
</tr>
<tr>
<td>missing</td>
<td>0.695</td>
<td>***</td>
<td>20.7</td>
<td>0.449</td>
</tr>
<tr>
<td>Fathers</td>
<td>&lt;50</td>
<td>1</td>
<td>46.8</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>&gt;50</td>
<td>1.266</td>
<td>***</td>
<td>11.4</td>
</tr>
<tr>
<td></td>
<td>dead</td>
<td>1.336</td>
<td>***</td>
<td>5.7</td>
</tr>
<tr>
<td></td>
<td>missing</td>
<td>0.925</td>
<td>†</td>
<td>36.0</td>
</tr>
<tr>
<td>Mothers</td>
<td>&lt;50</td>
<td>1</td>
<td>47.9</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>&gt;50</td>
<td>1.430</td>
<td>***</td>
<td>9.4</td>
</tr>
<tr>
<td></td>
<td>dead</td>
<td>1.311</td>
<td>**</td>
<td>4.4</td>
</tr>
<tr>
<td></td>
<td>missing</td>
<td>1.081</td>
<td>†</td>
<td>38.4</td>
</tr>
<tr>
<td>Birth order</td>
<td>first born</td>
<td>1</td>
<td>36.3</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>other</td>
<td>0.965</td>
<td>11.6</td>
<td>0.954</td>
</tr>
<tr>
<td></td>
<td>last born</td>
<td>0.995</td>
<td>16.9</td>
<td>1.029</td>
</tr>
<tr>
<td></td>
<td>missing</td>
<td>0.748</td>
<td>***</td>
<td>35.3</td>
</tr>
<tr>
<td>Number of</td>
<td>observations</td>
<td>7127</td>
<td></td>
<td>8141</td>
</tr>
<tr>
<td>Number of</td>
<td>marriages</td>
<td>5003</td>
<td></td>
<td>5003</td>
</tr>
<tr>
<td>Number of</td>
<td>persons</td>
<td>5003</td>
<td></td>
<td>5467</td>
</tr>
<tr>
<td>Time at risk (months)</td>
<td>408286</td>
<td></td>
<td>514475</td>
<td></td>
</tr>
<tr>
<td>LR chi2 (22)</td>
<td>637.46</td>
<td></td>
<td>1506.82</td>
<td></td>
</tr>
<tr>
<td>Log likelihood</td>
<td>-37352.76</td>
<td></td>
<td>-38108.5</td>
<td></td>
</tr>
<tr>
<td>Prob &gt; chi2</td>
<td>0.000</td>
<td></td>
<td>0.000</td>
<td></td>
</tr>
</tbody>
</table>

Note: *** p<0.001   ** p<0.01   * p<0.05   † p<0.1

Source: Family reconstitution database of Zsámbék (Gallina-Jelli 2002 and parish registers), own calculation
In the case of women, the picture is slightly different compared to men. Ages at and intensity of first marriages did not change until 1820, but afterwards the decrease in likelihoods is continuous until 1920 in the case of all women in the sample. Although brides’ ages at first marriage were rather unstable (sample 1), the same increase can be seen after 1920 but it was not without precedents. At the same time, differences in the ages at first marriage were not large as we see in Table 1.

Place of birth did not count in the case of those getting married, but in the other model in-migrants (independently of the date of the move) married later and/or with lower intensity. We may assume that in part social position (servants), childbearing outside marriage and out-migration could cause this lower likelihood. Women of foreign origin might have been overrepresented among these persons.

If a bridegroom was a widower, it had a strong impact on age at marriage: remarrying and consequently older husbands had a little bit older wives, the difference in the hazard ratios is considerably large. To marry to second time remarrying men meant almost half of the likelihood of a first marriage as compared to other cases where both partners were never married before.

The SES of the fathers shows the same effect as the bridegrooms’ profession: landowners’ daughters married the earliest. They were more attractive in the marriage market, and their marriage was based on strategic decisions about status reproduction. They married primarily to landowners’ sons, and as they were mutually prone to marry earlier their intentions strengthened each other.

As mentioned above, childbirth outside marriage did not make later marriage impossible. If someone married after this event, it accelerated her marriage with a likelihood almost 50% higher. But if we include non-marrying women too, this effect reversed. In reality a birth out-of-wedlock could significantly decrease the chance to marry in the village if the mother did not manage to wed the father of the child.

Fathers’ and mothers’ ageing and death had a different effect on daughters’ marriage as compared to the sons. The likelihood of marriage decreased significantly after these events which can be explained by the different consequences of marriage among men and women. Men – at least in part – could remain in the parental household after the marriage, and the parents’ ageing and death could open the way towards inheritance, household headship and marriage. Women almost always left the parental household when marrying. Old and/or widowed parents’ presence might increase the need of their work at home, result-
ing in later marriage. Naturally, the number and marital status of the siblings also mattered, but is not considered here.

As for birth order, the differences are significant, and the two models are different in this case too. Among marrying women, first born daughters married earliest. But considering all women in the sample we witness a totally different effect: daughters of higher birth order show higher likelihood – they married more frequently. Some first born daughters remained single longer, were more likely to die unmarried than others and more often gave birth to children outside marriage – probably not independent factors. They had to care for their old and widowed parents, younger brothers and sisters, etc. Those who married did so early, but others could remain unmarried more easily than their younger sisters.

Table 4

The likelihood of first marriage among women, Zsámbék, 1720–1945

<table>
<thead>
<tr>
<th>Periods</th>
<th>Women getting married</th>
<th>All women</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Hazard ratio</td>
<td>%</td>
</tr>
<tr>
<td>-1799</td>
<td>1</td>
<td>16.3</td>
</tr>
<tr>
<td>1800-1819</td>
<td>1.052</td>
<td>8.3</td>
</tr>
<tr>
<td>1820-1839</td>
<td>0.815 **</td>
<td>10.2</td>
</tr>
<tr>
<td>1840-1859</td>
<td>0.849 **</td>
<td>11.8</td>
</tr>
<tr>
<td>1860-1879</td>
<td>0.939</td>
<td>12.0</td>
</tr>
<tr>
<td>1880-1899</td>
<td>0.850 **</td>
<td>10.7</td>
</tr>
<tr>
<td>1900-1919</td>
<td>0.685 ***</td>
<td>13.0</td>
</tr>
<tr>
<td>1920-0.753 ***</td>
<td>17.8</td>
<td>0.665 ***</td>
</tr>
<tr>
<td>Place of birth</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Local</td>
<td>1</td>
<td>84.4</td>
</tr>
<tr>
<td>Foreigner</td>
<td>0.976</td>
<td>6.8</td>
</tr>
<tr>
<td>Missing</td>
<td>0.940</td>
<td>8.8</td>
</tr>
<tr>
<td>Marriage number of husband</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>1</td>
<td>88.2</td>
</tr>
<tr>
<td>2</td>
<td>0.586 ***</td>
<td>8.5</td>
</tr>
<tr>
<td>3</td>
<td>0.364 ***</td>
<td>0.7</td>
</tr>
<tr>
<td>4</td>
<td>0.532</td>
<td>0.0</td>
</tr>
<tr>
<td>missing</td>
<td>0.769 **</td>
<td>2.6</td>
</tr>
<tr>
<td></td>
<td>Women getting married</td>
<td>All women</td>
</tr>
<tr>
<td>--------------------------------</td>
<td>-----------------------</td>
<td>-----------</td>
</tr>
<tr>
<td></td>
<td>Hazard ratio</td>
<td>%</td>
</tr>
<tr>
<td>Father’ SES</td>
<td>landowner</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>day-labour- worker</td>
<td>0.727</td>
</tr>
<tr>
<td></td>
<td>artisan</td>
<td>0.776</td>
</tr>
<tr>
<td></td>
<td>intellectual</td>
<td>0.955</td>
</tr>
<tr>
<td></td>
<td>missing</td>
<td>0.789</td>
</tr>
<tr>
<td>State of woman</td>
<td>no birth</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>illegitimate birth</td>
<td>1.473</td>
</tr>
<tr>
<td>Fathers</td>
<td>&lt;50</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>&gt;50</td>
<td>0.727</td>
</tr>
<tr>
<td></td>
<td>dead</td>
<td>0.892</td>
</tr>
<tr>
<td></td>
<td>missing</td>
<td>0.824</td>
</tr>
<tr>
<td>Mothers</td>
<td>&lt;50</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>&gt;50</td>
<td>0.713</td>
</tr>
<tr>
<td></td>
<td>dead</td>
<td>0.801</td>
</tr>
<tr>
<td></td>
<td>missing</td>
<td>0.981</td>
</tr>
<tr>
<td>Birth order</td>
<td>first born</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>other</td>
<td>0.914</td>
</tr>
<tr>
<td></td>
<td>last born</td>
<td>0.915</td>
</tr>
<tr>
<td></td>
<td>missing</td>
<td>0.802</td>
</tr>
<tr>
<td>Number of observations</td>
<td>9575</td>
<td></td>
</tr>
<tr>
<td>Number of marriages</td>
<td>5524</td>
<td></td>
</tr>
<tr>
<td>Number of persons</td>
<td>5527</td>
<td></td>
</tr>
<tr>
<td>Time at risk (months)</td>
<td>453481</td>
<td></td>
</tr>
<tr>
<td>LR chi2 (22)</td>
<td>496.57</td>
<td></td>
</tr>
<tr>
<td>Log likelihood</td>
<td>-41891.7</td>
<td></td>
</tr>
<tr>
<td>Prob &gt; chi2</td>
<td>0.000</td>
<td></td>
</tr>
</tbody>
</table>

Note: *** p<0.001 ** p<0.01 * p<0.05 † p<0.1

Source: Family reconstitution database of Zsámbék (Gallina-Jelli 2002 and parish registers), own calculation
Chapter 5

Conclusion

First marriages in Zsámbék village, originally inhabited by Roman Catholic German settlers, took place at a young age and thus fit well into the model of the “Eastern” marriage pattern. However, the pattern of early marriage may conceal important differences, e.g. differences over time or by socio-economic status. Age at marriage increased in the 19th century especially among men, and probably not independently of economic conditions. Early marriage was characteristic especially among women and the sons and daughters of landowners (either serfs or free smallholders after the abolition of serfdom). First birth followed first marriage closely, so first marriage was literally the beginning of the reproductive phase of life similarly to in other European populations (Lundh, Kurosu, 2014b). Childbearing in marriage was not the only form of the entering reproductive period of life, however, and the rate of illegitimate births increased permanently; one part of those births was followed by marriage, while still unmarried mothers left the village. First conception often preceded first marriage, a phenomenon increasing in the period studied which permits us to assume normative trial marriage.

Our event history analysis confirmed the findings from the descriptive statistical analysis. It displayed the changes over time and the differences by SES. At the same time, it helped to explore the different meaning and consequences of marriage by sex: women had different roles in the parental households, and after marriage they left it in all probability while men could remain there. Consequently, the ageing and death of the parents had different effect on the marriage chances of men and women. Men benefited with greater likelihood in these cases, while the latter had smaller chances. Similarly, there were considerable differences by sex in birth order, which had no significant impact on men’s marriage, while the specific role of first-born daughters in the parental home seems to be confirmed by the analysis.

Considering subsequent life events related to first marriage and using event history analysis helped to reveal some details about first marriages in our village in Central Hungary. We demonstrated that the picture is much more complex than what we could draw with the help of indicators like age at first marriage and the share of never married. However, in order to get a more detailed picture with more results and satisfactory explanations, we have to continue this local study by refining our models and co-variates as well as by adding further traditional sources of social history.
References


Chapter 5


Chapter 6.
THE INFLUENCE OF CHILDHOOD MORTALITY ON SUBSEQUENT FERTILITY DURING THE DEMOGRAPHIC TRANSITION

Mark Gortfelder

Introduction

The first proponents of Demographic Transition Theory argued that a traditional society with high mortality had to compensate it with high fertility. “Any society having to face the heavy mortality characteristic of the premodern era must have high fertility to survive. All such societies are therefore ingeniously arranged to obtain the required births. Their religious doctrines, moral codes, laws, education, community customs, marriage habits, and family organizations are all focused toward maintaining high fertility,” claimed Notestein (1945) in his classic article. Clearly under the traditional mortality conditions family planning as it is practiced in the 20th century was not possible.

The first theorists took it for granted that the start of the infant and child mortality decline caused the existing, traditional reproduction system to collapse because it led to a population explosion with average yearly population growth of 1–3%. Such a growth could not be sustained so a correction had to be made to the prevailing fertility behaviour, which lead to a sustained decrease of the average number of births per woman (Davis, 1945). Because of this push, fertility became a part of the calculus of the conscious choice (Coale, 1973).

This classic formulation that the mortality decline preceded the fertility decline, however, was not supported by the results of the Princeton project that set out to prove the theoretical postulations (Van de Welle, 1986). This caused some debate among demographers. This article uses micro-level data from the Family Registry of the Estonian Republic to look at this central relationship at the level of individual couples during the time of the demographic transition.

Four research questions are proposed. First, did experiencing child mortality increase the likelihood of another birth, indicating a replacement response as standard demographic theory would expect? Second, by making a distinction between women born 1860–79 and those born
1880-99 we ask whether the replacement effect become more pronounced as fertility limitation spread among the population and became more consistent? Also, can we see a difference between urban women that were the pioneers in modernizing their reproductive life, as opposed to rural women? Third, did child mortality accelerate the timing of next birth? And fourth, did this effect grow in importance for women born later and for those living in urban areas?

The mortality-fertility connection during the transition

Much to the surprise of the scholars involved, the Princeton European Fertility Project did not support the intuitive and theoretically established understanding of the fertility transition as a necessary response by families to the infant and child mortality decline that started from the 19th century. Knodel (1974), for example, found that in half of the administrative units of Germany marital fertility began to decrease earlier than infant mortality. Lesthaeghe (1977) had to conclude that infant mortality was not an influence on the marital fertility decline. Inconclusive or negative conclusion were also found from research done in other countries (Demeny, 1968; Livi-Bacci, 1971, 1977; Teitelbaum, 1984). “At the end of this quest”, the concluding volume of the project summed it up, “we cannot report that the historical evidence confirms that the declines of infant mortality led to the decline of fertility” (Van de Walle, 1986, p. 233).

This led to some division among demographers with many seeing the classical formulation of the mortality-fertility relationship as proven false (van de Kaa, 1996). Others, however, have explained the surprising results (not only of the mortality-fertility synergy) of the Princeton project by the level of analysis and the faults of the indices used (Guinnane, Okun, Trussell, 1994; Reher, 1999). Mainly, the Princeton project used too high levels of aggregation to pick up the change in the behaviour of a considerable minority (20%) and used data from censuses that created a cross-sectional observation. Also, they used infant mortality as the mortality indicator, while historical research has shown that it was in fact child mortality (1–4-year-olds) that started to improve earlier.¹ Even Ansley Coale still held on to this idea in the concluding volume that traditional societies developed customs that promoted moderately high fertility to safeguard against extinction and that fertility did react to the revolutionary decline of mortality (Coale, 1986).

¹ The Princeton project in its concluding volume (Van de Welle 1986: 206–209) makes an overview and addresses these arguments.
Chesnais emphasized that the decline of mortality was very gradual and started already from the 18th century with the decrease of the virulence of certain infectious diseases. He also concluded that although at a local level there have been cases when fertility decline precedes that of early-life mortality, there are no cases, where this has been true at the national level and that on average, European mortality decline preceded fertility decline by 40 years (Chesnais, 1992, p. 140–148). Recent scholarship using macro-data has supported the classical understanding of the importance of decreases in early-life mortality for the fertility decline (Sánchez-Barricarte, 2017, 2018).

After the Princeton Project a wave of individual-level data gathering and analyses has developed that aims to bring a more detailed look into the period of the transition. For the most part, however, child survival has been a controlling variable and the real focus was on other research questions (Alter, Oris, Tyurin, 2007; Bengtsson, Dribe, 2006; Kolk, 2011; Van Poppel et al., 2012; Van Bavel, 2003, 2004), which means that some issues remained unresolved. An emphasis on the changing causal effect is missing and mostly the survival status of only the last child has been included.

As early-life mortality became rarer and parents adjusted their fertility accordingly, the effect on the risk of having another birth should have increased and this we can interpret as a strengthening of rational decision-making regarding reproductive life. This result has been found in some studies that have looked at the change of this relationship during the transition (Knodel, 1982; Reher, Sanz-Gimeno, 2007; Reher et al., 2017; Reher, Sandström, 2015). Reher and Sanz-Gimeno (2007) have also looked at urban/rural differences based on occupation. All of these studies have been conducted on small datasets from one or a few localities, not national data, which raises questions about representativeness.

The results of the Princeton project have suggested that more thought should be given to the complex relationship of mortality and fertility. The impact of early-life mortality on subsequent fertility of the mother is layered (Palloni, Rafalimanana, 1999; Knodel, 1982). First there is a pure physiological effect, if the child that died was an infant. An infant death means an end to breast-feeding, which leads to a quicker return to fecundity. The effect could also be there if breast-feeding ends without infant death in societal contexts, where mothers stopped breastfeeding early, which was not the case in Estonia (Rammul, 1928-1938). This physiological effect does not influence the hazard for next birth, but the timing of the next birth, and is powerful only when family
limitation is weak or non-existent.

The second mechanism, replacement, refers to deliberate attempts to “replace” an existing and desired child who dies, in order to meet parity-specific (net) fertility goals. The replacement strategy raises overall fertility and should be most evident if parity-specific family limitation is exercised more consistently and widely in the population. Replacement should also influence the timing. It should lead to a shorter interval, especially if the mother is nearing menopause. However, child mortality may also lead to a need for physical and emotional healing and/or to wait for a better economic situation and thus may have an opposite effect at times.

The third mechanism, hoarding (insurance), refers to the practice of bearing a number of children above the desired family size to insure against likely future losses. This mechanism is, however, difficult to measure. First, the insurance strategy depends on an individual couple’s adjustment to actual and perceived child mortality experiences in the surroundings and such perceptions are impossible to verify. Second, because hoarding is spread over the entire reproductive career of a couple the resulting association of mortality and fertility may require considerable time lags to be visible and detectable, both in terms of overall propensity and timing.

Data and study population

The Family Registry of the Estonian Republic is an individual-based historic population registry. It was kept from 1926 to 1949 by the lowest-level local governments (towns and boroughs) on all its residents who were Estonian citizens. It is national, although for those municipalities where there was fighting during the Second World War some or all of the registry books were destroyed. The Registry was computerised a few years ago as a cooperation between the Estonian Interior Ministry and Tallinn University. A sample of it has been studied

---

2 In the main towns from 1920. Before this vital event registration was conducted by religious authorities.
3 Out of a total of 378 municipalities, in 35 the Registry books have been destroyed completely and in 48 partially.
4 1939 local government reform created a problem of false linkages. Data insertion was not duplicated. So, if a person had an individual record then information from this was linked to the records of his or her parents. This was done using registry book and page index. With the merging of municipalities registry book number 1 in the old municipality, for example, was turned into book number 4 in the new municipality. And so false information was linked. After extra work the false linkages were limited to 6% of all mothers.
(Läll 2014), but the project itself was finalized in 2017 and the cleaning process was done by the author of the current article.

All local governments registered their population in a standardised way (Teder, 1939). They exchanged information if people changed their residence so there are only small problems of unregistered internal migration. Compiling the Registry was a big task for the municipalities and it is evident from comparing it with census data, that an individual record was not opened for all individuals. These problems affected mainly the elderly for whom the only vital event happening at the later stage of their lives was their own death, in which case a record was not opened.

The Registry contains retrospective information on vital events of the persons born preceding 1926. This is of high quality, since people could not solely rely on their own memory or oral family history, but documentary evidence from the church books was required.\(^5\) Still problems do exist. Due to the fact that the church books were event- and not individual-based, some events were not reregistered in the Family Registry. Looking at the proportion of illegitimate children to all children, it is evident that women were able to conceal their dead illegitimate children before 1926.

For an individual record to be opened the *de facto* requirement was that an individual had to live in Estonia between 1926 and 1949, which creates a significant left-truncation effect. This means that the coverage of each cohort decreases as we go back in time from 1926. For the cohorts born in 1880s we have approximately 1/3 of all births with a Family Registry individual record. The coverage ratio for mothers is of course a lot higher, since pre-marital mortality was still high. This observation scheme does create issues of selectivity and representativeness. This is first and foremost due to mortality selection – we can expect that mothers with better socioeconomic conditions who live in urban areas are overrepresented. Another issue is emigration, which was high from the 1860s to 1917 (Raun, 1986). These effects cannot be tested but they must be borne in mind.

For the study population women born between 1860 and 1899 with at least one child are chosen. The timeframe considers the coverage and selectivity issues that grow more serious for the cohorts born before 1860. The end-point is chosen so that all women under study would be over age 40 by 01.07.1941 after which losses occur for reasons due to war and foreign occupation, which are not wholly accounted for in the

---

\(^5\) Exceptions were made if events took place and were registered abroad and thus could not be accessed by the individuals (Teder 1939).
registry. Only mothers with legitimate, biological, singleton live births are analysed and those with at least one data quality problem are excluded. Parity progression to the second, third, fourth and fifth child are under analysis. To see the change of the effect during the transition women are separated into two cohorts – women born 1860–79 and 1880–99. Women are observed until the next birth or age 45. If a woman is not under observation until age 45 or her marriage ends, they are excluded.

Table 1

<table>
<thead>
<tr>
<th>Dataset sizes (number of mothers)</th>
<th>1&gt;2</th>
<th>2&gt;3</th>
<th>3&gt;4</th>
<th>4&gt;5</th>
</tr>
</thead>
<tbody>
<tr>
<td>Born 1860–1879</td>
<td>40015</td>
<td>33309</td>
<td>25759</td>
<td>18719</td>
</tr>
<tr>
<td>Born 1880–1899</td>
<td>67886</td>
<td>49795</td>
<td>31916</td>
<td>19291</td>
</tr>
</tbody>
</table>

**Demographic transition in Estonia and its broader context**

Until the 1880s the Baltic provinces of *Estland* (modern-day North-Estonia), *Livland* (modern-day South-Estonia and North-Latvia) and *Kurland* (modern-day West-Latvia) enjoyed wide autonomy from the imperial government. Local political, economic and cultural life in both rural and urban areas was dominated by the Baltic Germans (5% of the population) and their Medieval corporate institutions which were, of course, not at the forefront of building a modern statistical system. The first census in the Baltic provinces took place only in 1881 and only from that point onwards we have a fairly good understanding of national trends. A modern statistical organisation only began after the establishment of the Republic of Estonia in 1918. This is why we cannot pinpoint the start of the demographic transition exactly and cannot follow its early development in detail.

By the beginning of the 18th century when the territory of modern Estonia went from Swedish to Russian rule the population was quite devastated due to the interplay of war, famine and disease. Still, however, as the family reconstruction research of Heldur Palli prove (1973, 1984, 1988), Estonia was west of the so-called “Hajnal line” in terms of female age at (first) marriage already during the second half of the 18th century that saw quite substantial population growth (see
also Chojnacka 1976). By the time of the first census in 1881, infant and child mortality had probably decreased already for decades (Katus, 2000, Vallin et al, 2017). For fertility Katus (1994a) showed that using Princeton indices Estonia experienced the 10% decline in fertility in 1888, the fourth country in Europe by this measure. Marital fertility was by the time of the first census lower only in Sweden and Hungary. From the first Russian imperial census in 1897 it can be ascertained that Estonia had considerably lower fertility than other Russian provinces, including the neighbouring governates (Anderson et al., 1979). Using the same data of the Registry it is estimated that in cohort view the fertility transition in Estonia began from cohorts of women born in the 1830s or 1840s (Gortfelder, Puur, 2019).

The demographic transition in Estonia began early despite feeble economic development. Industrialisation and urbanisation only began at the end of the 19th century and by the time of the 1897 census less than 1/5 of the population lived in towns. This grew only a few percent by 1914 and reached 1/3 by the start of Soviet occupation in 1940 (Raun, 2002). This means that Estonia remained an agrarian land. Estonia, however, was at the forefront in Europe when it comes to mass literacy. By 1881 over 90% of the people were able to read and the figures for urban vs rural residents, men vs women, Baltic-Germans vs Estonians were the same by this measure. The development of literacy is based on the Lutheran Church and its teachings that emphasized the importance of reading scripture for all Christians (Kasekamp, 2010).

Spatial and social differentials in the timing of both mortality and fertility decline were considerable. Looking at the correlation of under-10 mortality and total fertility that the Family Registry and its municipalities offer, the coefficients are above 0.7. Pioneering areas were in South Estonia with its more advanced agricultural development, while the laggards were the West, Northeast and especially the Southeast, where ethnic Russians made up the majority (Katus, 1994b). Also, of course, the towns had an earlier fertility decrease. Nõges (1925) showed that in the economically and educationally most advanced county of Viljandi infant mortality declined from the 1850s-60s, while fertility started to decrease there from the 1850s and this grew to be more rapid in the 1870s.6

Infant and child mortality began to decline in rural areas as con-

---

6 He gives figures by parish starting from the 1830s. Due to the agricultural crises in the 1840s it is difficult to pinpoint the start of the decrease, since it caused an upswing in mortality and downswing in fertility.
temporary observations reveal (Grosset, 1883; Huebner, 1861) and this is evident also with the Registry (Läll, 2014). These appeared among cohorts of women born in the 1860s. Contemporary studies indicate the pioneering role of the Baltic Germans in the fertility transition (Rahamägi, 1923), but this may have been more due to lower nuptiality. Germans definitely had lower infant mortality than Estonians (Jaadla, Puur, Rahu, 2017).

Thus, we conclude that ideally, we would like to extend the Family Registry back in time by 2–3 decades to be able to see the demographic transition from the pre-transition age to its end both in the pioneering areas and among the leading social groups. But still the Family Registry is an appropriate tool for detecting the changes occurring in the course of the demographic transition and therefore can be useful to study the research questions proposed here.

Results

On figure 1 the course of the decline of (overall) fertility, under-10 mortality and net fertility are shown. Due to the fact that the we cannot extend the dataset back in time, the figure does not cover the whole of the demographic transition. But the basic shape of the transition is evident and already hints at the importance of child mortality improvements for the decline of fertility. Using the fact that the population is divided by the lowest-level local government units it is also easy to do a correlation test for the under-10 mortality and fertility relationship. For women born 1860–79 the correlation coefficient is 0.73, for the 1880–99 cohort the measure is 0.76. Using the change of mortality and fertility the coefficient is 0.62. This provides another proof that the relationship for both the level and change is strong.

To have a clear understanding of the effect of child mortality on subsequent fertility parity progression ratios (PPR) and median birth intervals are compared based on child survival status. PPR shows the probability of having a subsequent birth, while the median birth intervals, of course, show the timing of this next birth. In table 2 we present the main results for all women under study.

---

7 317 municipalities were used for this exercise.
Figure 1. Mean total fertility, net fertility and children died for mothers born 1860–99

Table 2

Probability of having another birth and time to it (in months) by parity and child survival status of previous children

<table>
<thead>
<tr>
<th>Child deaths</th>
<th>Parity progression ratios</th>
<th>Median birth intervals</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1&gt;2</td>
<td>2&gt;3</td>
</tr>
<tr>
<td>0</td>
<td>0.85</td>
<td>0.75</td>
</tr>
<tr>
<td>1</td>
<td>0.88</td>
<td>0.80</td>
</tr>
<tr>
<td>2</td>
<td>0.84</td>
<td>0.77</td>
</tr>
<tr>
<td>3</td>
<td></td>
<td>0.79</td>
</tr>
<tr>
<td>4</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>0.85</td>
<td>0.77</td>
</tr>
</tbody>
</table>

Notes: Risk set is defined as women that have had at least one birth for the 1>2 progression, two births for the 2>3 progression and so on. Number of deceased children are counted for the time of the conception of the next child.
First, we will discuss the results for the PPR that are depicted on the left. Comparing the ratios by rows an obvious pattern emerges. At each parity progression women who have experienced child mortality by the time of the conception of the next child have higher likelihoods for progressing to a higher parity. For example, the progression to third birth shows that women with no child deaths had likelihood 0.75, those with one deceased child 0.8 and those unlucky ones whose both previous children had died it was even higher 0.84.

Reading table 2 by columns we can compare the results for the four parity progressions. It is clear that at higher parities the likelihood differences by child survival status are smaller. A selection effects probably explains this. Mothers that had already reached parity 4 did not limit their fertility as consistently as those that stopped at younger parities. Also, it has to be borne in mind that the median age for those at parity 4 is higher than at lower parities and therefore they were not as fecund. In other words, women at parity 4 had more difficulties to conceive regardless of child survival status.

Turning to median birth intervals that are depicted on the right in table 2 we can see that birth intervals were shortened by experiencing child mortality. This means that couples aimed to replace the deceased offspring quickly. However, a physiological effect also plays a part for the timing if the deceased child was an infant whose death also meant an end to breast-feeding and return to ovulation. As with PPR-s the effect of child survival decreases to non-existence for the last parity progression under study.

Table 3

<table>
<thead>
<tr>
<th>Child deaths</th>
<th>1&gt;2</th>
<th>2&gt;3</th>
<th>3&gt;4</th>
<th>4&gt;5</th>
</tr>
</thead>
<tbody>
<tr>
<td>Parity progression ratios</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0</td>
<td>0.90</td>
<td>0.82</td>
<td>0.83</td>
<td>0.70</td>
</tr>
<tr>
<td>1</td>
<td>0.90</td>
<td>0.87</td>
<td>0.85</td>
<td>0.77</td>
</tr>
<tr>
<td>2</td>
<td></td>
<td>0.86</td>
<td>0.83</td>
<td>0.80</td>
</tr>
<tr>
<td>3</td>
<td></td>
<td></td>
<td>0.77</td>
<td>0.81</td>
</tr>
<tr>
<td>4</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>0.90</td>
<td>0.82</td>
<td>0.84</td>
<td>0.72</td>
</tr>
</tbody>
</table>
We now turn to the question: how did these effects change during the transition? Table 3 presents PPRs and median birth intervals ordered also by the birth cohort of the mother. Those born 1860–79 are the cohort for which the demographic transition was in effect only in the urban areas and some agriculturally more developed rural areas. For the 1880–99 cohort the fertility transition had begun nearly everywhere and fertility levels had already decreased substantially for the pioneering areas.

We can see that in the later cohort the probabilities of next birth are lower across all parity progressions. However, a substantial change does occur during the transition. For the earlier cohort there appears to be little if any difference in the likelihood of having another birth depending on child survival status. This indicates that although at least some replacement behaviour existed, it was limited and perhaps most evident in the pioneering areas. The comparison with the later cohort shows much bigger discrepancies based on child survival status. This holds also for the last parity progression under study. Thus, we can infer that the role of human agency strengthened in fertility decision-making process. More and more couples were trying to reach parity-specific aims, which, if ruined by child mortality, they were highly motivated to mend.

With regard to birth intervals, an equivalent change is not seen. If anything for the later cohort the differences of median intervals are smaller. This is unexpected even if total change (last row) was also miniscule. Looking at median birth intervals for 5-year cohorts it is evident that unlike quantum the timing of fertility did not experience an uninterrupted linear change, but the length of birth intervals first de-
creased for women born in the 1870s and then began to increase again for those born in the 1890s. This pattern holds also if we divide the study population by child survival status. This would indicate that the fertility transition in Estonia was mostly based on stopping rather than spacing behaviour.

Table 4

Probability of having another birth and time to it (in months) by mothers’ residence, parity and child survival status of previous children

<table>
<thead>
<tr>
<th>Child deaths</th>
<th>1&gt;2 urban</th>
<th>1&gt;2 rural</th>
<th>2&gt;3 urban</th>
<th>2&gt;3 rural</th>
<th>3&gt;4 urban</th>
<th>3&gt;4 rural</th>
<th>4&gt;5 urban</th>
<th>4&gt;5 rural</th>
</tr>
</thead>
<tbody>
<tr>
<td>Parity progression ratios</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0</td>
<td>0.78</td>
<td>0.88</td>
<td>0.66</td>
<td>0.79</td>
<td>0.64</td>
<td>0.74</td>
<td>0.64</td>
<td>0.72</td>
</tr>
<tr>
<td>1</td>
<td>0.84</td>
<td>0.90</td>
<td>0.75</td>
<td>0.83</td>
<td>0.66</td>
<td>0.75</td>
<td>0.62</td>
<td>0.69</td>
</tr>
<tr>
<td>2</td>
<td></td>
<td></td>
<td>0.80</td>
<td>0.87</td>
<td>0.72</td>
<td>0.80</td>
<td>0.64</td>
<td>0.71</td>
</tr>
<tr>
<td>3</td>
<td></td>
<td></td>
<td></td>
<td>0.71</td>
<td>0.83</td>
<td>0.66</td>
<td>0.75</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.67</td>
<td>0.78</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>0.79</td>
<td>0.89</td>
<td>0.69</td>
<td>0.80</td>
<td>0.66</td>
<td>0.75</td>
<td>0.64</td>
<td>0.71</td>
</tr>
</tbody>
</table>

Median birth intervals

<table>
<thead>
<tr>
<th>Parity progression ratios</th>
<th>0</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>20</td>
<td>16</td>
<td>20</td>
<td>22</td>
<td>22</td>
<td>20</td>
</tr>
<tr>
<td>1</td>
<td>21</td>
<td>15</td>
<td>17</td>
<td>18</td>
<td>22</td>
<td>21</td>
</tr>
<tr>
<td>2</td>
<td>23</td>
<td>22</td>
<td>22</td>
<td>24</td>
<td>24</td>
<td>23</td>
</tr>
<tr>
<td>3</td>
<td>25</td>
<td>24</td>
<td>22</td>
<td>24</td>
<td>24</td>
<td>24</td>
</tr>
<tr>
<td>4</td>
<td>24</td>
<td>24</td>
<td>24</td>
<td>24</td>
<td>23</td>
<td>24</td>
</tr>
<tr>
<td>Total</td>
<td>20</td>
<td>21</td>
<td>23</td>
<td>24</td>
<td>24</td>
<td>24</td>
</tr>
</tbody>
</table>

Notes: Risk set is defined as women that have had at least one birth for the 1>2 progression, two births for the 2>3 progression and so on. Number of deceased children are counted for the time of the conception of the next child.

Lastly, we turn to the urban-rural comparison of the mortality-fertility synergy. Table 4 has instead of cohort the residence type of the mother as a distinction. The likelihoods of experiencing another birth are higher for rural women regardless of child survival status. However, we can see clearly that among the urban women the discrepancies based on the number of dead children are more pronounced, meaning that urban women had more modern reproductive lives. This we would expect since the urban areas were the pioneers in new fertility behaviour.
Additionally, an extra analysis was done (not reported in here) which divided urban women into those living in the bigger towns\(^8\) and those living in the small towns. The pattern that emerged was that women living in small towns were in between the rural and big town women in their likelihoods for experiencing another birth, as expected.\(^9\)

Birth intervals were only a tiny bit shorter in urban areas, indicating that what set urban women apart was their implementation of stopping strategy to control their childbearing. Looking at the differences based on number of child deaths, the impact of child fatalities is more pronounced for the rural women. This again underlines that the increasing agency was more a factor of parity-specific rather than birth timing control.

**Conclusion**

The results reported prove that the experience of child deaths was an important factor for parents considering whether or not to have an additional birth. For the first research question proposed, we confirm that having deceased children raised the likelihood of having another birth across the four parity progression that were under study, although the effect decreased in size for the higher progressions. The effect was significantly more pronounced for women born later and women living in urban areas. This answers in affirmative the second research question proposed. The increase of family planning during the course of the transition and among the pioneering groups made the differences in probabilities of next birth more pronounced. This confirms the classical understanding of the demographic transition, that in the traditional era behavioural replacement of dead children was not common, but as early-life mortality began to decrease, which enabled parents to start having explicit fertility goals in terms of number of children reaching adulthood, parents began to deliberately replace deceased children to achieve their goal. Experiencing child mortality also shortened the birth intervals, which was the third question proposed, although for the last parity progression under study this effect largely disappears. Finally, for the fourth research question we have to answer in the negative. The effect of child mortality on birth intervals decreased for the last cohort and was not more pronounced for the women residing in urban areas.

---

\(^8\) Tallinn, Tartu, Narva, Pärnu.

\(^9\) The picture got more complicated for the last progression, since the numbers of women get small at some child survival levels.
From the results we can infer that even in a traditional society (in our case the earlier cohort and rural women) at least some behavioural replacement response can be detected as also others have shown (Knodel 1982; Reher et al. 2017). But the power of this effect grows a lot when couples practice modern family limitation.

References


Chapter 6
Gortfelder


Chapter 6


Chapter 7.
RELIGION AND MARRIAGE IN EARLY TWENTIETH CENTURY EKATERINBURG, RUSSIA:
A MICRODATA ANALYSIS

Elena Glavatskaya,
Alexander Bobitsky,
Elizaveta Zabolotnykh,
Anastasia Vishnevskaya

Introduction

Rapid industrialization and urbanization characterized late nineteenth century Russia. Ekaterinburg was one of the cities that received migrants from both the countryside and the western provinces of the Russian Empire. Especially in the first quarter of the twentieth century the city became more ethnically and religiously diverse, even if the majority, over 90 per cent, were ethnic Russians and members of the Russian Orthodox Church. While there are many studies conducted on European urban nuptiality, less has been done on the Russian sources. However, the Princeton Fertility Project analyzed differential marriage ages in the Russian gubernias (provinces) based on the 1897 census aggregates in their study of fertility in Russia (Coale et al.). Their study included Perm’ province (even including Ekaterinburg in the Asian part) with a population of about 3 million, but did not consider internal regional or religious differences. The same applies to a comprehensive study of Russia’s social history that has a section on demography (Mironov, 2014). Urban nuptiality in the parts of Russia to the east of the Urals got even less attention and most studies were based on aggregate data. The exceptions are research focusing on Ekaterinburg’s Catholic and Jewish communities by the authors of this paper (Glavatskaya, Borovik, 2016; Glavatskaya, Borovik, Bobitsky; Glavatskaya, Zabolotnykh) and one on the Siberian city Barnaul (Vladimirov, Sarafanov 2013). The latter, however, studied only the Russian Orthodox population and provided only information on the
intervals of marriage age among different social groups, and not the **mean age at first** marriage.

The ‘European marriage pattern’ concept, introduced by John Hajnal in 1965, refers to two basic characteristics: high age at first marriage: 27–29 for men and 24–26 for women, and a large proportion of people remaining unmarried: 10–15 percent (Hajnal). The concept formed a turning point in marriage studies both in Europe and beyond. Lynch found that the ‘European marriage’ characteristics were especially manifest in cities (Lynch). Hajnal’s geographic demarcation between the East and the West – the Trieste – Saint-Petersburg line was modified by researchers who took social status, cultural values, the place of residence and migration status into consideration (see for example Moreels & Matthijs, Holom et al). Our goal is to analyze differential marriage patterns, including age at first marriage and proportion of final celibacy in a city situated near the border between Europe and Asia. Ekaterinburg had been Europe-oriented with its industries and western Christian communities since the eighteenth century. It was a major Russian city throughout the nineteenth and twentieth centuries, and its situation can be a key to understanding demographic dynamics in the core of Russia.

Even if limited by source availability, we analyzed several marriage characteristics:

1. Singulate mean age at marriage (SMAM) as well as number and percent of never married in Ekaterinburg compared with European Russian data and its development between 1897 and 1926.

2. Age at first marriage by denomination, parish, social and migration status. This calculation we made using the nominative data transcribed from the parish registers of Ekaterinburg.

**Sources and methods**

There are two censuses, with aggregates allowing us to compute singulate mean age at marriage, as well as the number and percent of the never married: the first All-Russian Census conducted in 1897 and the first All-Soviet Union census conducted almost three decades later. Both were carefully prepared, run and processed according to international standards of census taking (Thorvaldsen 2018). However, the microdata were destroyed after the information was processed and aggregated into tables in accordance with the Russian practice. After a successful start in population registration during the all-Russian
census in 1897 there was a break due to the first Russian revolution in 1905, followed by yearlong insurgences that made the next census impossible. The economic hardships and food crises after Russia entered World War I, required urgent information about population and supplies, and desperate attempts were made by several Russian municipalities in 1916 and 1917 to register its population in order to arrange efficient food supply. These and the next Bolshevik attempt to take a census in 1920 failed, due to lack of resources and turbulent time of foreign interventions and civil war (Thorvaldsen, Glavatskaya, 2017). However, we have aggregate census data on 1920 Ekaterinburg. Since 1913 and 1920 aggregates did not have data on ages we decided to use 1897 and 1926 censuses to compute singulate mean age at marriage (SMAM) to see overall developments. Our microdata research is based on vital events registered on the individual level in the parish registers (metricheskie knigi, hereafter referred to as parish registers or church books) from Ekaterinburg’s religious communities. Tsar Peter the Great ordered the registering of vital events in church books from 1722. During the next two centuries, the forms evolved and the religious denominations’ parish registers became more standardized. They all had three common parts – about baptisms/births, about weddings and about burials/deaths, Muslim and Jewish books in addition had a section on divorces, which these religions justified. This registration of vital events took place until the October Revolution in 1917 all over the Russian Empire. After that, a civil office established by the Bolsheviks took over the registration in each municipality. Russian legislation regulated the parish registers’ accuracy. The religious community board had to check and verify the books frequently and religious leaders and communities were to be fined was there any disorder found in the records.

We found parish registers in the State Archive of Sverdlovsk oblast’ (GASO) with births, weddings and funerals. The information from the parish registers was transcribed parish after parish into separate databases and then exported into the database Ural Population Project (hereafter referred to as UraPP). The registration of marriages in the church books provides names (first, family and patronymics), marital status, social status and/or occupation of grooms and brides, their place of origin or registration, age, religion (when appropriate) and date of the wedding; person who performed the ritual service; and there is also information on the parents and witnesses. We have transcribed marriage data for four Russian Orthodox Church parishes
with altogether almost 10000 weddings from 1880 until 1923, among them almost 6000 for the period from 1900 until 1919. In addition, we have transcribed data from the Old Believers\textsuperscript{10} parishes; St. Anna Catholic Church books, the St. Paul Lutheran Church books, as well as similar sources from the city’s Synagogue and Muslim community and analyzed them. Altogether, we analyzed about 7000 marriages registered in the 10 city parishes (See table 1).

Table 1

Number of marriages by year and denomination in Ekaterinburg

<table>
<thead>
<tr>
<th>Year</th>
<th>Russian Orthodox Church</th>
<th>Old Believers</th>
<th>Catholics</th>
<th>Lutherans</th>
<th>Jews</th>
<th>Muslim</th>
<th>Sum</th>
</tr>
</thead>
<tbody>
<tr>
<td>1900</td>
<td>308</td>
<td>4</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>312</td>
</tr>
<tr>
<td>1901</td>
<td>314</td>
<td>8</td>
<td>3</td>
<td></td>
<td></td>
<td></td>
<td>325</td>
</tr>
<tr>
<td>1902</td>
<td>277</td>
<td>5</td>
<td>5</td>
<td></td>
<td></td>
<td></td>
<td>287</td>
</tr>
<tr>
<td>1903</td>
<td>284</td>
<td>4</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td>289</td>
</tr>
<tr>
<td>1904</td>
<td>136</td>
<td>4</td>
<td>4</td>
<td></td>
<td></td>
<td></td>
<td>144</td>
</tr>
<tr>
<td>1905</td>
<td>279</td>
<td>6</td>
<td>3</td>
<td></td>
<td></td>
<td></td>
<td>288</td>
</tr>
<tr>
<td>1906</td>
<td>385</td>
<td>7</td>
<td>6</td>
<td>10</td>
<td></td>
<td></td>
<td>408</td>
</tr>
<tr>
<td>1907</td>
<td>399</td>
<td>13</td>
<td>7</td>
<td>11</td>
<td>9</td>
<td></td>
<td>439</td>
</tr>
<tr>
<td>1908</td>
<td>314</td>
<td>19</td>
<td>9</td>
<td>10</td>
<td>11</td>
<td></td>
<td>363</td>
</tr>
<tr>
<td>1909</td>
<td>329</td>
<td>21</td>
<td>6</td>
<td>11</td>
<td>7</td>
<td></td>
<td>374</td>
</tr>
<tr>
<td>1910</td>
<td>301</td>
<td>19</td>
<td>8</td>
<td>25</td>
<td>7</td>
<td></td>
<td>360</td>
</tr>
<tr>
<td>1911</td>
<td>283</td>
<td>17</td>
<td>7</td>
<td>23</td>
<td>12</td>
<td></td>
<td>342</td>
</tr>
<tr>
<td>1912</td>
<td>350</td>
<td>7</td>
<td>3</td>
<td>38</td>
<td>8</td>
<td></td>
<td>406</td>
</tr>
<tr>
<td>1913</td>
<td>377</td>
<td>6</td>
<td>4</td>
<td>23</td>
<td>5</td>
<td></td>
<td>415</td>
</tr>
<tr>
<td>1914</td>
<td>370</td>
<td>9</td>
<td>6</td>
<td>12</td>
<td>8</td>
<td>18</td>
<td>423</td>
</tr>
<tr>
<td>1915</td>
<td>239</td>
<td>10</td>
<td>6</td>
<td>21</td>
<td>14</td>
<td>9</td>
<td>299</td>
</tr>
<tr>
<td>1916</td>
<td>187</td>
<td>6</td>
<td>25</td>
<td>63</td>
<td>12</td>
<td>11</td>
<td>304</td>
</tr>
<tr>
<td>1917</td>
<td>267</td>
<td>12</td>
<td>32</td>
<td>70</td>
<td>33</td>
<td>15</td>
<td>429</td>
</tr>
</tbody>
</table>

\textsuperscript{10} Ethnic Russians whose ancestors dissented from the Russian Orthodox Church over religious reforms in the 17th century. For more details on Ekaterinburg Old Believers see (Borovik, 2018).
These microdata allowed the computing of mean age at first marriage, share of remarriages by denomination, and how it was influenced by WWI, Revolution and Civil War.

**Ekaterinburg: ethno-religious and demographic portrait**

Pre-Revolutionary Ekaterinburg was an industrial city in the Middle Urals with growing ethnic and religious diversity, especially in the first quarter of the twentieth century (Glavatskaya 2018). The vast majority belonged to the Russian Orthodox Church which we shall refer to here as the State Church, and there were additional communities of Old Believers, Muslims, Catholics, Lutherans and Jews (see table 2).

### Table 2

Religious denominations in Ekaterinburg 1897 (N, %)

<table>
<thead>
<tr>
<th>Denomination</th>
<th>Men</th>
<th>Women</th>
<th>Sum</th>
<th>%</th>
</tr>
</thead>
<tbody>
<tr>
<td>State Church</td>
<td>18534</td>
<td>21211</td>
<td>39745</td>
<td>91.9</td>
</tr>
<tr>
<td>Old Believers</td>
<td>766</td>
<td>1024</td>
<td>1790</td>
<td>4.1</td>
</tr>
<tr>
<td>Muslims</td>
<td>386</td>
<td>292</td>
<td>678</td>
<td>1.6</td>
</tr>
<tr>
<td>Lutherans</td>
<td>167</td>
<td>176</td>
<td>343</td>
<td>0.8</td>
</tr>
<tr>
<td>Catholics</td>
<td>167</td>
<td>156</td>
<td>323</td>
<td>0.7</td>
</tr>
<tr>
<td>Jews</td>
<td>150</td>
<td>153</td>
<td>303</td>
<td>0.7</td>
</tr>
<tr>
<td>Other</td>
<td>23</td>
<td>34</td>
<td>57</td>
<td>0.1</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td>20193</td>
<td>23046</td>
<td>43239</td>
<td>100</td>
</tr>
</tbody>
</table>

*Source: 1897 Census aggregates (Troinitskii 1904: 92–93).*
Chapter 7

The more than 91 percent of the State church members in the city according to the 1897 census were overwhelmingly ethnic Russians, which was also the case for the additional four percent Old Believers. After them, the city’s Muslim community was the second biggest religious minority, composed of Tatars and Bashkirs – in-migrants to the city. The overwhelming majority of the Lutherans were Germans and most Catholics were of Polish origin. The same proportion of Jews came from diverse places, mostly in Western Russia.

The survey conducted in Ekaterinburg in 1913 by the city’s address office showed increasing religious diversity. All the non-Orthodox denominations expanded their share in the religious composition of the city, mainly due to in–migration, but also natural population growth (see table 3).

Table 3

<table>
<thead>
<tr>
<th>Denomination</th>
<th>Men</th>
<th>Women</th>
<th>Sum</th>
<th>%</th>
<th>Constructed, N¹</th>
<th>Constructed %</th>
</tr>
</thead>
<tbody>
<tr>
<td>Russian Orthodox²</td>
<td>35024</td>
<td>34177</td>
<td>69201</td>
<td>90.6</td>
<td>81833.3</td>
<td>90.5</td>
</tr>
<tr>
<td>Jews</td>
<td>679</td>
<td>456</td>
<td>1135</td>
<td>1.5</td>
<td>1453.4</td>
<td>1.6</td>
</tr>
<tr>
<td>Catholics</td>
<td>587</td>
<td>364</td>
<td>951</td>
<td>1.2</td>
<td>1104.1</td>
<td>1.2</td>
</tr>
<tr>
<td>Lutherans</td>
<td>512</td>
<td>377</td>
<td>889</td>
<td>1.2</td>
<td>1031.6</td>
<td>1.1</td>
</tr>
<tr>
<td>Muslims</td>
<td>2854</td>
<td>1139</td>
<td>3993</td>
<td>5.2</td>
<td>5112.0</td>
<td>5.7</td>
</tr>
<tr>
<td>Other</td>
<td>136</td>
<td>43</td>
<td>179</td>
<td>0.2</td>
<td>201.0</td>
<td>0.2</td>
</tr>
<tr>
<td>Sum</td>
<td>39792</td>
<td>36556</td>
<td>76348</td>
<td>100</td>
<td>90422.6</td>
<td>100</td>
</tr>
</tbody>
</table>


Even if the municipal office in 1913 likely used instructions distinct from the 1897 census, we think it mirrors the general dynamics of the religious groups’ sizes. The 1920 census reflected the general

¹ Children under 14 were not registered in 1913. We solved this problem by constructing cohorts of children using data on their share in each denomination according to the 1897 census. However, children aged 10–14 are still missing even in this calculation due to the age grouping in the census aggregates. Therefore, the constructed relative numbers for 1913 are closer to reality than the absolute.
² Including Old Believers.
population decline and the changes in the ethnic composition of the city caused by the WWI, Revolution and Civil War. There was no question about religion in the 1920 census forms, but we used ethnic markers to distinguish religious identity, for they were closely related. We distinguish Poles as ‘ethnic Catholics’, Germans as ‘ethnic Lutherans’, Tartars and Bashkirs as ‘ethnic Muslims’, Russians, Ukrainians, Belarussians as ‘ethnic Orthodox’. The Old Believers most of whom were ethnic Russians we naturally had to include in this group as well. According to the 1920 census aggregates, the Jewish population increased significantly, the Catholics’ did not change, while all other ethnic groups and denominations suffered drastic decreases, Muslims in particular (See table 4).

Table 4

<table>
<thead>
<tr>
<th>Denomination</th>
<th>Men</th>
<th>Women</th>
<th>Sum</th>
<th>%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ethnic Orthodox</td>
<td>37121</td>
<td>43142</td>
<td>80263</td>
<td>90.8</td>
</tr>
<tr>
<td>Jews</td>
<td>1682</td>
<td>1923</td>
<td>3605</td>
<td>4.1</td>
</tr>
<tr>
<td>Ethnic Catholics</td>
<td>575</td>
<td>507</td>
<td>1082</td>
<td>1.2</td>
</tr>
<tr>
<td>Ethnic Lutherans</td>
<td>199</td>
<td>161</td>
<td>360</td>
<td>0.4</td>
</tr>
<tr>
<td>Ethnic Muslims</td>
<td>774</td>
<td>648</td>
<td>1422</td>
<td>1.6</td>
</tr>
<tr>
<td>Other</td>
<td>1050</td>
<td>618</td>
<td>1668</td>
<td>1.9</td>
</tr>
<tr>
<td>Sum</td>
<td>41401</td>
<td>46999</td>
<td>88400</td>
<td>100</td>
</tr>
</tbody>
</table>


Marriage patterns in Ekaterinburg and Ekaterinburg uezd\textsuperscript{11}

*The singulate mean age at marriage*

Since the 1897 census microdata are not available we extracted information from the aggregates and computed the singulate mean age at marriage (SMAM) for men and women both in the city with a popula-

\textsuperscript{11} Rural administrative area
tion of 43239 and the uezd – the surrounding, administrative rural area with a population of 369057. As expected urban SMAM was higher than rural: 3.6 years difference among men and 4 years among women (See table 5)

Table 5

SMAM in Ekaterinburg city and Ekaterinburg’ uezd in 1897

<table>
<thead>
<tr>
<th></th>
<th>Ekaterinburg</th>
<th>Ekaterinburg’ uezd</th>
</tr>
</thead>
<tbody>
<tr>
<td>Men</td>
<td>27.1</td>
<td>23.5</td>
</tr>
<tr>
<td>Woman</td>
<td>24.6</td>
<td>20.6</td>
</tr>
</tbody>
</table>

Source: 1897 Census aggregates (Troinitskii 1904: 36–37).

Ekaterinburg’s SMAM corresponds well with the mean age at marriage among the Russian urban population in 1910: 27.7 years for men and 24.8 for women in big cities\(^\text{12}\) and 26.8 and 22.8 respectively in smaller ones (Mironov, 2003, 169); with women’s age at first marriage pattern closer to the big cities. To see the dynamics we also computed SMAM using 1926 census data for both Sverdlovsk (Ekaterinburg)\(^\text{13}\), and its’ subordinated rural area – Sverdlovskii okrug’, with a population of 272179. To make it comparable with the 1897 data we constructed the same age groups as in 1897 and calculated SMAM based on the same cohorts (See table 6).

Table 6

SMAM in Sverdlovsk and Sverdlovskii okrug in 1926

<table>
<thead>
<tr>
<th></th>
<th>Sverdlovsk</th>
<th>Sverdlovskii okrug/rural</th>
</tr>
</thead>
<tbody>
<tr>
<td>Men</td>
<td>24.9</td>
<td>22.0</td>
</tr>
<tr>
<td>Woman</td>
<td>21.8</td>
<td>20.0</td>
</tr>
</tbody>
</table>

Source: 1926 census aggregates (1926 Census’ tables, 1930:82–83).

\(^{12}\) Those with population over 100 000 – Saint-Petersburg, Moscow, Warsaw, Lodz’, Odessa, Riga, altogether 20 cities in the 1897 Russian Empire.

\(^{13}\) Ekaterinburg was renamed Sverdlovsk in 1924 after the regional revolutionary leader Iakov Sverdlov. After 1992, it again became Ekaterinburg.
As expected urban SMAM was higher than rural for both men and women, however it had dropped several points since 1897. The differences between urban and rural areas were lower as well, especially in the case of women: 1.8 years compared to the 4 years in 1897. From the SMAMs computed for 1897 and 1926, we might assume that the mean ages at first marriage in the early 20th century were not higher than 26 for men and 23.2 for women.

**The never married**

Since there is no alternative data available, we computed the marriage distributions in Ekaterinburg according to the 1897 and 1926 census data and the share of the never married for the cohorts 40–49 and 50–59 (see Table 7 a,b).

**Table 7a**

<table>
<thead>
<tr>
<th>Age</th>
<th>Men</th>
<th>Woman</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>N</td>
<td>% never-married</td>
</tr>
<tr>
<td>15–16</td>
<td>954</td>
<td>100.00</td>
</tr>
<tr>
<td>17–19</td>
<td>1478</td>
<td>96.68</td>
</tr>
<tr>
<td>20–29</td>
<td>3776</td>
<td>56.99</td>
</tr>
<tr>
<td>30–39</td>
<td>3266</td>
<td>17.30</td>
</tr>
<tr>
<td>40–49</td>
<td>2346</td>
<td>9.80</td>
</tr>
<tr>
<td>50–59</td>
<td>1277</td>
<td>6.97</td>
</tr>
<tr>
<td>Average for 40–49 and 50–59</td>
<td>8.39</td>
<td></td>
</tr>
</tbody>
</table>

*Source: 1897 census aggregates (Troinitskii 1904: 36).*

**Table 7b**

<table>
<thead>
<tr>
<th>Age</th>
<th>Men</th>
<th>Woman</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>N</td>
<td>% never-married</td>
</tr>
<tr>
<td>15–16</td>
<td>2382</td>
<td>99.62</td>
</tr>
</tbody>
</table>

Numbers and % of never married in Ekaterinburg by gender and age group

146
As seen from the tables above the tradition of compulsory marriage in Ekaterinburg was still strong after WWI, Revolution, Civil War and nine years of secular history. What is more, it became even stronger, for the proportion of never married dropped from 8.39 % to 3.72 % for men and from 13.52 % to 6.14 % for women.

*Age at first marriage group wise: microdata analyses*

The approximate mean age at first marriage we computed above for early 20th century Ekaterinburg – not higher than 26 years in case of men and 23.2 in case of women, was almost the same as in most Russia’s cities – 26.8 and 22.8 correspondingly (Mironov, 2003:169). However, the UraPP data analyses based on nominative data for the whole city while giving the same result for the mean age at first marriage for men: 26.2 years, showed lower age at first marriage for women – 21.4 (see table 8).
According to the Russian legislation at the time, the conventional earliest marriage age was 16 for the brides and 18 for the grooms. The Archbishop could allow a wedding even if the bride or groom were more than half a year younger than they had to be, however [Grigorovskii]. According to the UraPP, there were few weddings of underage newlyweds in early 20th century Ekaterinburg: 89 brides aged 15 years old (2%) and 39 grooms aged 17 years old (0.9%). Such early marriages were not a widespread phenomenon in late Imperial Russia, only 2.5 percent of the brides married before reaching 17 years (Tol’ts, 141). Most Orthodox Church brides in Ekaterinburg married between 20 and 22 years old just like their peers in Barnaul (Vladimirov, Sarafanov, p. 99–101). As to the grooms, most married between 19 and 27. Some entered the marriage market before they reached 20, the age of conscription in Russia. Depending on the regiment, young men served three to four years in the army (five in the navy). The returned conscripts entered the marriage market at the age of 25–26.

Table 8 shows that the Orthodox Church grooms’ age at first marriage differed by one year depending on the parish they belonged to or registered their wedding in. On average they married at the age of 25.4 and the average age of their brides was 21.1. Ekaterinburg’s Jews married almost at the same ages. Catholic and Lutheran grooms on average married at the age of 28.5 and 27.7 respectively. Muslim men usually...
married once they reached the age of 28. However, we consciously put their data separately because we have their records transcribed for the very short period 1914–1918, which were turbulent years. There were fewer differences in the Orthodox Church brides’ mean age at first marriage in different parishes, just half a year. Lutheran and Catholic brides married on average a few years older than the Russian Orthodox Church brides did, while the Muslim brides contrary to the Russians married even younger.

The microdata allowed us to compute mean age at first marriage according to social status. On average Ekaterinburg Russian Orthodox Church petty bourgeoisie (meschane) and peasant grooms married at the age of 24.7 and 24.8; bourgeoisie group involved in trade (kuptsy) – 27.4 and nobility (dvoriane) at the age of 28.2. However, there were noted differences between parishes. For example, among St. Epiphany Church grooms the petty bourgeoisie on average married one year later than peasants belonging to the same parish.

It is interesting that their co-ethnics, the Old Believers, displayed the opposite pattern with respect to social status: peasant grooms married one year later than their coreligionist petty bourgeoisie did. For the Old Believers we also computed age at first marriage according to migration status, while other variables are still in the process of coding. More than 55 percent of the Old Believers who registered their weddings in Ekaterinburg were peasants who in-migrated to the city. Finding themselves in the new social and cultural environment, they could have experienced problems starting a family and therefore had to postpone marriage. That corresponds well to the findings based on Historical Population Database of Transylvania. Those who migrated to more developed regions married on average one year later than their non-migrating counterparts (Holom et al.). Brides’ mean age at first marriage correlated directly with their social status as well: higher social status meant higher mean age at first marriage. In addition, girls from the Orthodox St. Epiphany Church grooms the petty bourgeoisie on average married one year later than their peers from other parishes.

**Age differences between the spouses**

On average, the Ekaterinburg grooms were 4.8 years older than their brides, while in case of the Muslims it was 8.7 years (see table 8). Among couples with a noted age difference of 12–15 years, we note the wedding between Old Believer Anisya Narbutovskaya, a 19-years-old peasant girl, who married Ivan Kulik, a 31-years-old Pol-
ish peasant from Galicia then part of Austria-Hungary on 23 October 1917. Ivan probably was one of the war prisoners, exiled to the Urals in thousands after their capture in 1914–1915. We do not know what was his denomination in Galicia, but two days before the wedding, Ivan was baptized by the Old Believer minister Porfirii Mokrushin in the Virgin’s Assumption Chapel and so became an Old Believer himself. As the parish record said, it was his first marriage (GASO. F. 6. Op. 13. D. 334. L. 12 vv.–13; D. 336. L. 2 vv.–4). However, taking into consideration the circumstances, the general Christian belief that the baptism rite ‘cleans’ a person from all the former sins, and the fact that the Old Believers did not recognize other religions’ sacraments, we can be certain that Porfirii – the Old Believer minister who performed both rites, would not consider Ivan’s previous marriage as valid. There are examples of POWs marrying Russian girls while in captivity, even being already married back at home before captivity (Glavatskaya, Borovik, 2016).

WWI and age at first marriage

The UraPP microdata allowed not only computing how the WWI, Revolution and Civil War affected age at first marriage in Ekaterinburg, but also looking beyond the averages (see Table 9).

Table 9

Mean age at marriage by denomination in Ekaterinburg, 1901–1914 and 1915–1919

<table>
<thead>
<tr>
<th>Denominations</th>
<th>Grooms, MAM</th>
<th>Brides, MAM</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>cases</td>
<td>Before 1914</td>
</tr>
<tr>
<td>St. Epiphany Cathedral</td>
<td>301/96</td>
<td>25.7</td>
</tr>
<tr>
<td>Ascension Church</td>
<td>1051/226</td>
<td>25.0</td>
</tr>
<tr>
<td>St. Catherine Cathedral</td>
<td>827/172</td>
<td>25.1</td>
</tr>
<tr>
<td>Church of the Holy Spirit</td>
<td>809/40</td>
<td>24.8</td>
</tr>
<tr>
<td>State Church total/average</td>
<td>2988/534</td>
<td>25.2</td>
</tr>
</tbody>
</table>
Chapter 7

<table>
<thead>
<tr>
<th>Denominations</th>
<th>Grooms, MAM</th>
<th>Brides, MAM</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>cases</td>
<td>Before 1914</td>
</tr>
<tr>
<td>Catholics</td>
<td>50/155</td>
<td>31.4</td>
</tr>
<tr>
<td>Lutherans</td>
<td>142/158</td>
<td>26.0</td>
</tr>
<tr>
<td>Jews</td>
<td>70/53</td>
<td>27.4</td>
</tr>
<tr>
<td>Non-Orthodox total/average</td>
<td>262/366</td>
<td>28.3</td>
</tr>
</tbody>
</table>

Source: UraPP

As expected, age at first marriage increased among both grooms and brides in all Russian Orthodox Church parishes by 1.8 and 1.3 years respectively, even if there was some differences between parishes. However, age at first marriage increased even more among the Lutherans: by 2.3 and 2 years respectively. This group suffered population decrease between 1913 and 1920. On the contrary, in-migrated Catholic and Jewish men started to marry much younger than their co-religionists did in Ekaterinburg before the WWI, as did the Catholic women.

Conclusion

Our analyses of marriage behavior in the church records allow us to get insights into different religious groups’ life in early 20th century Ekaterinburg. In addition, we computed the singulate mean age at marriage (SMAM) on the basis of age group tables by marital status from the 1897 and 1926 census aggregates. That allowed us to put our data into some perspective.

Our calculation of SMAM for Ekaterinburg city separately from the same census are somewhat higher, 24.6 years for brides vs 22.5–24 calculated by the PFP (Princeton). This can be due to the whole Perm’ province urban population data being affected by the lower SMAM rates in other cities with a higher proportion of peasants in their population. Ekaterinburg’s industrial character attracted urban population elements from the West, especially Catholics, Lutherans and Jews. In addition, in-migrants usually postponed their marriage one year.

The 1897 aggregates do not subdivide the age group/marital status tables further, but our analysis of the parish registers on the individual level provided clues to understand the religious affiliation’s influence on age at first marriage. The Catholic, Lutheran and Muslim men mar-
ried a few years older than the Russian Orthodox and Jewish men did. The difference in age at first marriage among their brides was less significant: 1–2 years higher among Catholic and Lutheran women compared to the Russian Orthodox and one year less among the Muslim brides. We also found out that belonging to a certain parish within the same confession might influence mean age at first marriage in case of both brides and grooms.

Our research confirmed the general opinion that social status matters. In addition, as our data showed, the in-migrants postponed their marriage one year compared to their peer co-ethnics and co-religionists if not pressed by crises in times of war or political instability. For example, the male peasant Old Believers from the countryside who moved into the city, married almost one year later than their co-religionist townsmen. Both findings correspond well to similar results from Transylvanian data (Holom et al.). The likely reason is that the city created opportunities unavailable in the countryside and thus more freedom to start or postpone family life. It also required some time to settle before joining the marriage market.

While invisible as part of Perm’ province in the Princeton Project, prerevolutionary Ekaterinburg seems to be closer to the ‘European’ marriage pattern, at least in terms of both SMAM and men’s age at first marriage. Individual level data show that mean age at first marriage was slowly increasing in the State Church parishes until WWI, when the age at first marriage increased by 1.8 years for grooms and 1.3 for brides. The Lutherans, mostly ethnic Germans, postponed their weddings even more: 2.3 and 2 years respectively: likely due to the shortage of eligible partners on the marriage market, since the Lutheran-German population decreased dramatically between 1913 and 1920. In addition, their life in a country being on war with Germany was rather difficult. Unfortunately, we do not have individual data after 1919 to explain the fact that ten years after the Revolution for one reason or another Ekaterinburg/Sverdlovsk moved away from the European tradition to marry later. According to the 1926 Census data analyses, both SMAM and the percentage of never married became lower. We may assume as one cause that after the Revolution, the new regime shut down in-migration from the West. Also, in-migration increased from the rural suburbs, and most important, changed the very basics of the marriage institution in the country. The in-migrants from rural areas paradoxically at the time brought with them the Orthodox Church or Muslim marital eth-
ics, which required obligatory marriage. Being under stress during the Civil War these in-migrants tried to marry earlier to get social support in turbulent times, like prisoners of war and refugees did during the WWI. In addition, the Bolsheviks’ legislation separated the church from the state, deprived it the right to register weddings, and introduced freedom of divorce already in 1917. Both marriages and divorces were to be registered by a civil officer in a state office at once, with no sacraments required. To prove that hypothesis we need to get individual level data for the period between 1919 and 1926, which is our next project.

References


Chapter 7

Vsesoyuznaya perepis’ naseleniya 1926 goda. Ural’skaya oblast’ : [tabli-
cy]. Otd. 3. Semejnoe sostoyanie. Mesto rozhdeniya i prodolzhitel’nost’ prozhi-
vaniya. Uvechnost’. Otd. ott. iz izd.: Vsesoyuznaya perepis’ naseleniya 1926
Section 3. Marital status. Place of birth and duration of residence. Injury. Sep-
arate print from volume 38 of the results of the all-Union census of 1926].
handle/123456789/12871
Chapter 8.
URBAN AND RURAL FAMILIES IN THE LATE 18TH CENTURY SAINT PETERSBURG PROVINCE ACCORDING TO THE 5TH TAX REVISION (REVIZSKYE SKAZKI)

Maria Markova

Introduction

Family and other population history studies based on microdata is a rapidly developing strain in modern Russian historiography. Thus, the second half of the 20th century to the beginning of the 21st centuries are characterized by active use of primary, individual level demographic data on the Russian population (Ulyanova, Troitskaya). In most cases Russian scholars have analyzed rural populations, peasants in particular, and only a few studies consider urban populations or the clergy (Avdeev, Troitskaya, Ulyanova; Postnikov). Our research is based on the revizskie skazki and focuses on a comparative analysis of demographic trends in different social groups registered in the late 18th century Saint Petersburg province: merchants and low-middle class city dwellers and peasants: privately owned serfs and those who belonged to the Tsars’ family.

Revizskie skazki (hereafter referred as revisions) were originally fiscal registers designed to list all persons who were subjects to taxation. Peter the Great introduced the system in 1718 and there were ten such revisions run in 1719, 1745, 1763, 1782, 1795, 1811, 1815, 1843, 1850 and 1858 before the abolition of serfdom in 1861\(^1\). Due their cross-sectional nature, we can classify them as census-like, but they were not only kept for statistical purposes. In Saint Petersburg province, the revisions are well preserved and researchers have at their disposal significant amounts of information dating from the 18th and 19th centuries covering large areas. Of course, the 19th century manuscripts are preserved better than the 18th century primary data. For example, only a small number of documents remain from the fourth revision run in the 1780s. However, we found representative amount of the fifth revision’s primary data from 1795. Most of the revision lists used in this study are

\(^1\) For more information on Revizskie skazki see Troitskaya, 1995.
kept in the two central Russian archives: Tsentral’nyi gosudarstvennyi istoricheskii arkhive v Sankt Peterburge (Central State Historical Archive in Saint Petersburg) (hereafter TsGIA) and Rossiiskii gosudarstvennyi istoricheskii arkhiv (Russian State Historical Archive) (hereafter RGIA). The revision lists included data on various social groups of the Russian capital’s population and its neighboring city Sofia with merchants, artisans, house serfs, foreigners, etc, as well as various categories of peasants, priests, clerks and German colonists.

**The source basis**

For this study, we used revision lists with data on Saint Petersburg and Sofia merchants and lower-middle class residents, as well as lists with data on serf-peasants and peasants belonging to the Tsars’ family.

**Merchants and lower middle class residents**

Revision lists about Saint Petersburg merchants and lower-middle class families are kept in TsGIA’s collection 221 (Petrograd Merchant Board) in Saint Petersburg) (TsGIA. F. 221. Op.1. D. 169). The collection contains revision lists about the tax-paying population of the capital between 1740s and the 1830s.

Revision lists about Sofia merchants and lower-middle class families are kept in RGIA’s collection 488 (Tsarskoye Selo Town Hall) (RGIA. F. 488. Op.1. D. 451, 469a, 469b). It should be noted, that this collection also contains lists from Gatchina and Pavlovsk estates.

**Peasants**

1. Revision lists about the Saint Petersburg province peasant population are stored in a number of collections in central and regional archives. Revision lists on the Tsars family peasants we found in RGIA’s collection 487 (Tsarskoye Selo Palace Administration) (RGIA. F. 487. Op.9. D.901).

2. Revision lists with data on the serf-peasants are preserved partly in TsGIA’s collection 479 (Petrograd Treasury Chamber) (TsGIA. F. 479. Op.23. D.132).

The revision lists are organized by “domokhozyaystvo”, a term which can be translated as “household”. They contain information on the head of the household or family and then about all its members and how they relate to the head (wife, son, daughter etc).
The revision also provides data on age, occupation and marital status of the persons listed, including women (Kashchenko, Markova, 103; Matison, 134). They may occasionally contain information about a couple’s previous marital status when applicable, place of residence and premarital social status, as well as the time of the marriage. For example the Runovskaya farm’s revision list had such information registered in a separate column entitled ‘A imenno’ (namely). One household head had the wife Kristina, daughter of peasant Davyd Davydov. This tax payer had married her in 1788 in the village where she lived. During the fourth revision in 1768 she was registered as eloped; but in 1795 she was back (TsGIA. F. 479. Op.23. D.132. L.200vv.)

Information on the female relatives (daughters, nieces) who had married and left the family during the inter census period, can be found in the separate column “These have left after the revision and are still absent for various reasons.” This section often provides data on their marriage, residence and husband’s social status.

The revision lists also contain information about the age of all family members, stated during a previous or the contemporary revision, or in both cases. This allows computing the average age at marriage as well as the average age difference of the spouses among the social groups in Saint Petersburg province for the entire 18th century. Even if affected by noted age heaping this information is of great value, for the vital event registers of the time did not provide information on the newlyweds’ ages. It was only since 1838, due to the legal requirement, that priests started to register the brides’ and grooms’ ages.

However, the marriage date was not always stated in the documents; for example, there is no such information in the revision lists of the Sofia merchants and lower-middle class citizens. As a compensation, these documents provide information on social mobility, the ratio of first marriages and re-marriages, marital migration, as well as information on family’s composition among various categories of the population.

**Settlements chosen for analysis**

This paper presents analyses of population in a few settlements with both rural and urban population elements. As examples of urban settlements we use Sankt-Petersburg and Sofia – a town located
near the Emperor’s summer residence – Tsarskoye Selo some 23 km away from the capital at the time – Saint Petersburg. Empress Catherine the Great founded Sofia in 1780 with a plan to construct a modern European city with developed infrastructure, straight streets, water supply system, etc. Its first settlers were those employed at the Emperor’s summer residence. In administrative terms Sofia used to be the center of a uezd – a sub province administrative unit, but eventually it also included the Emperors’ residence Tsarskoye Selo.

Two settlements, Pulkovo and Podgornoye Pulkovo, were chosen to represent the rural population in this study. Both were located near the capital as well and populated by peasants belonging to the Tsars’ family (Gleserov, 352). In addition, we analyzed the Taytskaya estate populated by serf-peasants privately owned by the noblemen Alexander Demidov (Aleksandrova, 266–286).

In sum, we have processed information on both urban population groups in Saint Petersburg province: 593 households with 1244 persons from Saint Petersburg itself and 474 households with 1187 persons from Sofia; and rural population groups: 159 peasant households with 1243 people owned by the Tsars’ family and 173 serf-peasant households with 1251 people.

Family structure

The analysis of family structure among Saint Petersburg and Sofia artisans and merchants showed significant differences between the two cities. In Saint Petersburg, the percentage of single persons in the revision was very high: 51% of the total number of “families”; 37% were nuclear families; extended and multiple families were equal and made up 3% each. Lidia Semenova in her analyses of qualified workers (masterovye) employed in Sankt Petersburg military department also noted a high percentage of families (more than 30%) consisting of single men. It turned out that they were married but lived in Sankt Petersburg on their own while their family members lived elsewhere. For example in 1797, 36 of 51 qualified workers employed at the Okhtinskii factory in Sankt Petersburg had their wives and children staying in the countryside. Being peasants daughters, they had to live on the farm. Only 15 workers had their wives living together with them at the factory. A similar situation was observed while studying the revision lists of the workers employed at Saint Petersburg gun-
powder producing plant in 1798: half (33 of 65) of their wives lived in the countryside (Semenova, 177–179).

Most of Sofia’s merchants and low-middle class citizens had nuclear families (56%); 37% consisted of single men; extended and multiple families made up 3% each like in Sankt Petersburg.

The rural families as defined in the relationship column turned out to be much bigger than the urban ones; however, the differences between the various strata of the peasant population were small. Peasant families belonging to the Tsars on average had 7.8 people; while serf-peasants families on average consisted of 7.2. That corresponds well with the Russian family size with 5–7 members on average throughout the 15th and early 19th centuries found by others (Aleksandrov, 57).

Multiple family households (see above) were the most common in rural areas; the proportion of such families in the Tsars’ settlements was 51%, and in the serf-peasants villages – 50%. The proportion of nuclear families among the Tsars’ peasants was 27%; and among privately owned serf-peasants – 30%. Only 13% of the Tsars’ peasants and 10% of the serf-peasants had extended families; and families consisting of single person were rare in both cases – 8% (see figure 1).

Figure 1. Distribution of peasant families by type

![Figure 1. Distribution of peasant families by type](image-url)
Migratory status of the urban population

The Empress wanted to build the city of Sophia as fast as possible, so she issued special decrees to employ people regardless of their social status. As a result, runaway serfs inhabited Sofia’s outskirts (posad) and legalized their new status. Analysis of the revision data shows that 20% of Sofia’s merchants and lower-middle class family heads (in total 99) registered in 1782 were former fugitives. Their former place of residence proves that the urban population growth between the Fourth and Fifth Revisions was due to in-migration. For example, only Sophia’s 142 household heads were listed in both the 1782 and 1795 revisions, which composed only 38% of its population. The vast majority had moved from a different place of origin to their present place of residence. Migrants from western provinces accounted for 3%; northwestern (Novgorod, Pskov and Saint Petersburg) – 17%, central – 20% (mostly from Tver’, Yaroslavl’ and Moscow provinces); 22% – from northern (Arkhangelsk, Vologda, Olonets provinces) (See figure 2).

Figure 2. Sofia household’s heads’ previous places of residence
Basic parameters of marital behavior

Urban population

The intense migration processes that took place in Sophia between the fourth and fifth revisions affected the marriage behavior of its citizens. We have calculated the age difference of the spouses married in the 1780s and 1790s in Saint Petersburg and Sofia (152 and 252 couples respectively).

The analyses showed that the spouses’ age at marriage was either similar or did not exceed 5 years (with older husbands) in 52% of all cases in Saint Petersburg and 41% in Sofia. Couples where husbands were ten or more years older than their wives, composed up to 15% of the total number of marriages in Saint Petersburg and more than 30% in Sofia. Merchants and lower-middle class men in the town likely did not want to start families too early, trying to settle into the new place. Active migration processes could have caused postponement of marriages by Sofia’s grooms.

Sofia single men’s ages indirectly support that assumption: about one third of them were 31 – 40 years old, and 28% belonged to the age group 41 to 50 years old. Among the single there were many who stated that they ‘did not remember their parental origins’ (nepomnyashchiye rodstva). One such person, a 44 years old Sofia’s lower-middle class fellow, Ivan Fedorov, used to be a serf belonging to the landowner Ivan Andreev Yemelyanov who lived in the village of Vasil’kovo, Vologda province, wherefrom he run away and lived as a worker in Sofia. According to the revision in 1784, he was already registered as a settler on the Sofia outskirts (Posad). People like Ivan could not easily join the marriage market, having neither a clear status nor confidence. Most of Saint Petersburg’s single men, on the contrary, were under 30 years old. Those between 11 and 20 made up 24% of all single men; and those 21 to 30 years old – 45%.

We have analyzed the age-difference between spouses in 135 couples married in the 1780s and the 1790s. In most cases (up to 70%) the spouses had almost the same age at marriage or one of the spouses was maximum five years older. Cases when the husband was 10 or more years older composed up to 7%; and those with a wife more than six years older than her husband composed only 2%. 

162
Chapter 8

**Rural population**

In most serf-peasants families both husband and wife belonged to the same landowner: such unions composed up to 88% in the settlements under consideration; every third union was concluded between fellow villagers. However, there were exceptions: we found sixteen cases when the Tsars’ peasant brides married privately owned serf-peasants; and a soldier’s daughter married a serf.

Couples who lived in Pulkovo and Podgornoye Pulkovo had similar marital behavior. Spouses in 76 of the 91 couples (84%) had married at the same age or one of them was up to five years older. There were only thirteen cases registered when the husband was more than six years older and only two cases when the wife was six or more years older.

Also, grooms from the Tsar settlements preferred marrying the same status brides (89%). Only four grooms married brides from the state-run village; another four married daughters of the privately owned serf-peasants; two married a soldier’s daughter and a sailor’s daughter.

**Conclusion**

Our study proved that most of the fifth revision primary lists are well preserved, which allowed analyses of family structure and main characteristics of marital behavior among selected urban and rural population groups in Saint Petersburg province in the 1780s and 1790s.

We found significant differences in family size and structure among the urban and rural population. Saint Petersburg merchants and lower-middle class had extremely high percentages of families composed of a single man, whose family members likely stayed in the villages. A majority of the neighboring Sofia town’s merchants and lower-middle class had nuclear families.

Peasants in the rural areas of Sankt Petersburg province had big extended families: there were three times more family members than in the merchant and low-middle class families living in the cities. When comparing the two categories of peasants, serfs and those who belonged to the Tsars’ family, we have not found much differences in family structure, size, or preferences in choosing a spouse.

Thus, the revisions have good informational potential for studying family and marriage behavior among different social groups and its dynamics in late 18th mid 19th centuries Russia.
Sources


Sobranie raznyh znanij o zakonah rozhdeniya i smerti v rode chelovecheskom [Collection of different knowledge of the laws of birth and death in the human race]. Sobranie sochinenij vybrannyh iz mesyaceslovov na raznye gody. St. Petersburg. 1787. (In Russian)

References


Chapter 8


At the national level, we can distinguish between three main flows of migrants: emigration, immigration and domestic migration. It is strange to what degree the first two flows have been the subject of summary Norwegian studies relative to the latter. Already shortly after the War, Ingrid Semmingsen published her comprehensive summary of US emigration in two volumes, a pioneering work also in an international context (Semmingsen 1941, Semmingsen 1950). And just after the turn of the millennium, immigrants received their three volume history, with the addition of a summary book in English - the result of a comprehensive team work (Kjeldstadli 2003, Brochmann and Kjeldstadli 2008). Although there are a number of local and regional studies of internal migration in Norway, we lack a broad summary overview. A brief and concise overview can be found on the website Norgeshistorie.no under the heading “On migration to and within Norway” authored by Jan Myhre, a prolific author of migration history. The website contains as much about the quantitatively modest immigration as about the extensive internal migration.¹ The only reference to the latter theme is to an edited book about the 19th century, mainly a collection of regional and local history articles (Gjerdåker 1981). These include local moving on the west coast, long distance migration from southern to northern Norway as well as labour migration in southern Norway. Three articles are based on the social history project about Ullensaker parish (with

Gardermoen airport) and Kristiania / Oslo. It is symptomatic that while there are master theses on emigration to America from both these places, the study planned to link the two subprojects – treating migration from Ullensaker to the capital, remains unwritten (Koren 1979, Østrem and Rinnan 1979). Internal migration is also absent as a topic in the recent history of vital statistics aggregates (Søbye 2014). A systematic search in the local history journal Heimen and the national Historisk tidsskrift until 2017 shows ten and two articles, and four and one, respectively, in which words with the texts “migrasjon” or “migr” appear in the title and refer to internal migration. This distribution shows the extent to which internal migration has been a local history theme. The text string “migr” otherwise occurs a lot in Historisk tidsskrift because of the many references to emigration on the national level.

The historiography looks somewhat better in our neighbouring countries. In particular, Russia is well covered due to the collaboration between a Russian historian and a migration historian, both teaching in the United States, resulting in a broad presentation of internal Russian migration in the 20th century (Siegelbaum and Moch 2014). Hans Christian Johansen overviewed internal migration in Denmark in two major works (1975, 2002). For Sweden, it is peculiar how the overview article about migration history in Wikipedia starts with emigration before treating immigration, so that internal migration implicitly is defined out of the migration concept. However, internal migration was treated thoroughly in geographer Hägerstrand’s work, which is fundamental also for international migration theory (Sundbärg 1910, Hägerstrand 1947, Hägerstrand 1953). Erik de Geer summarized internal migration in Sweden and Finland through most of the 1800s and 1900s in connection with his emigration studies (1977). For Finland, we have a summary of internal migration for the six decades before World War II (Jutikkala 1953) and a series of recent studies of the connection between migration and mortality (Saarela and Finnäs 2008). We can conclude that outside Denmark, summaries of internal migration history are in short supply in the Nordic region. Nevertheless, there is no shortage of sources to base this article on, which will mainly describe the evolution of internal migration from the census of 1865 and provide a history of important internal relocations. We ask the following research questions: Where and when did the main internal migration streams go? Did women or men constitute the majority at different times? What are the most important series of source materials?

---

2 https://sv.wikipedia.org/wiki/Sveriges_migrationshistoria
Important internal migration currents

Sølvi Sogner’s PhD work on population growth and migration based in Rendalen parish northeast of Oslo on the border with Sweden, was the earliest study of internal geographic mobility, and determined that the main direction has been from the inland towards the coastal areas for centuries (1979). To a greater extent than Michael Drake (1969), she accepted migration from Akershus province surrounding the capital to Northern Norway rather than underenumeration as the reason for the relatively larger population growth in the northernmost provinces according to the population censuses of 1779 and 1801. However, the lack of nominative data and a birthplace column in these censuses respectively, make it hard to quantify this issue thoroughly.

The Colonization of Northern Norway

The most comprehensive and most widely mentioned wave of internal migrants is that from South-Eastern Norway and Southern Trøndelag to the Målselv and Bardu valleys southeast of Tromsø. The land clearing in these valleys in the southern part of Troms province started in the era of mercantilism, and is linked to Bailiff Jens Holmboe who was born in Lesja, thus in the origin of this migration wave. In 1781 he was appointed as top administrator of the taxation district outside Tromsø - thus getting first hand knowledge of both the starting and ending points for the migrants. These territories have been described wrongly as virgin and unpopulated, but for a long time formed the boundary between Sami and Norwegian settlements. The Sami had used the valleys in annual trekking with their reindeer from the winter pastures in Inland Sweden to their summer camps by the Norwegian coast. These valleys bordered on Malangen, where fish-farmers have cultivated the land and fished at least since the mid-1600s. Thus, several of the first who travelled north settled in the extended areas surrounding the valleys, including the big island of Senja. Erik Dørrum’s master thesis provides a detailed chronological overview of the settlements in Målselv valley, explaining how they were related by family ties, how some moved on, married and had children until 1835 based on church records, probate and land registers. For Bardu valley, there is a somewhat simpler list of settlers (Granlund 1975). The 1801 census lists the people living in the new settlements, in Bardu 74 people and in Målselv 263 with a significant surplus of men and a low average age, especially in Bardu – the most remote part. The farm cadastres of 1797 and 1807 show increas-
ing herds of horses, cattle, and that the settlers were pioneers when it came to potatoes and peas, while still holding on to grain cultivation, although it did not yield returns every year.

This pioneer immigration northwards continued until 1805, when 40 settlers had settled in the valleys, especially below the waterfalls. During the next fifteen years, only nine new colonists came, but in the 1820s this migration wave picked up new speed and a total of 159 colonists were counted in the numeric census of 1835. The migration was interrupted from the late 1790s since all state aid disappeared from 1797, along with a ban on the sale of timber from state forests and several years of failed harvests. It is apparent from the numeric population censuses 1815 to 1855 that the population in inland Troms grew slowly until the 1830s when it reached 1,000 inhabitants and the flow of migrants dried up. The easiest available market with the most heterogeneous business community close to the coast expanded faster than the settlements further inland, where we find a surplus of men through 1855. The lay preacher Hans Nielsen Hauge visited his most faithful congregation with advice on economic activities and suggesting well-sounding farm names until he was imprisoned for illegal sermons. It has probably happened that he sent bride candidates northwards, but the birthplaces of married women in the 1865 census suggests this to be mostly a myth— the surrounding province of Troms was a far more common birthplace.

The population grew until 1865, in the previous decade the increase was 914 extra persons when we treat the Målselv and Bardu valleys together. This was record growth, but not much greater than the population growth of 1846–1855 with 882 persons. Only from 1855, after the municipalities were split into separate parishes, is it easy to calculate the birth surplus from the church records (number of births minus the number of deaths in the same period). The surplus was 707 inhabitants for the next decade and only slightly lower in the previous decade of 1846-1855. Thus, we see that fertility in the period 1846–1865 meant more than migration for the population development. If we assume that

---

1 While Bardu and Målselv parish are marked in the baptismal lists for Ibestad and Lenvik before 1851/1853, this is not done in the respective funeral lists. By estimating a reasonable number of deaths, I calculate a birth surplus of 644 people. The inhabitants of the most peripheral parts of inner Troms used churches in other priestships for baptism and burial (Volden 1979). This will lead to under-registration of the birth surplus, meaning that the culmination of the in-migration should be pushed even a few years back in time.
Figure 1. Birthplaces in the census 1865 for the population resident in Målselv and Bardu parishes
the birth surplus had increased as much as the number of people since
the last decade, there is reason to believe that immigration was consid-
erably higher before 1845 than afterwards, i.e. immigration to inland
Troms culminated earlier, as the number of new homestead permissions
show. Therefore, the average annual surplus of immigrants was less
than 25 people a couple of decades after immigration into inland Troms
culminated in the 1820s. From about 1865 the migration flow reversed.
The valleys of Målselv and Bardu turned into an area with more out-
migrants than in-migrants since towns and overseas areas became the
main destinations.

The 1865 census was the first to give a detailed overview of resi-
dents’ birthplaces. It shows that Inland Troms formed an exception to
the general pattern that while the towns had a large proportion of in-
migrants, most of the people were born where they resided in the rural
parishes. Admittedly, non-migrants represented the majority also here
according to the 1865 census, but like in many towns, there was al-
most equilibrium: About 43 percent were born elsewhere. 11% of the
population were born elsewhere in Troms province, 14 % in Hedmark
province, 5% in Oppland province and 4% in Trøndelag provinces; the
three latter in southern Norway. Overall, birthplaces in northern Nor-
way dominated in the valleys, with over 2/3 of the population in 1865.
The 223 fathers from southern Norway had from one to nine children
each, who in total constituted 624 second generation persons from the
south, and with the second generation included, there was equilibrium
between people originating in the southern and northern parts of Nor-
way in 1865. The maps in figure 1 show where the inhabitants of Mål-
selv and Bardu parishes were born in greater detail. Compilation of the
birthplaces of the fathers and children can also be used to estimate the
frequency of step migration. However, few stays along the way were so
permanent that children resulted. Of 93 children born to fathers origi-
nating outside Målselv / Bardu, both generations were born in the same
parish in 72 cases. Only one child born in Nordland and six children
born in Tromsø can indicate step migration. In other words, it was far
more common to establish a family before leaving from the south than
to start family life along the way northwards.

Among the 1865 census aggregates there is a table indicating the
type of birthplace (births in own parish, foreigners, born in a town
and born in a rural area) for each parish. According to this, 17% of
the inhabitants in the northernmost Nordland, Troms and Finmarken
provinces were born in rural places without specifying in which part
of the country. Two smaller migratory streams to Northern Norway have been subjected to detailed studies, one about migrants from Voss east of Bergen and one about relocation from Gudbrandsdalen north of Lillehammer to Salten (Hanssen 1979, Gjerdåker 1981). At the time, these students did not have nationally transcribed censuses at their disposal. The 1865 census lists 216 immigrants from Voss and Vossesstrand to Northern Norway, 156 in Nordland, 56 in Troms and only 4 in Finnmark province – thus concentrated further south between Bodø and Harstad in Nordland province. Another stream ended in the 1800s in the Tverrelv valley in Alta in Finnmark province; there were 195 persons from southern Norway in Alta according to the 1865 census. Smaller groups were easily identified in other parishes, especially in neighbouring Troms province, e.g. 22 people born in Gudbrandsdalen residing in Lyngen parish east of Tromsø in 1865.

Yngve Nedrebø’s theory of the migration from Sogn treats a theme at the intersection of internal relocation and emigration, showing that from the outer parts of Sognefjord, people moved primarily to other places in western Norway and preferably to Bergen, while the migrants from the inner parts of the Sognefjord more often ended up in America. This has been called the “uncle effect”; the rationale being that youngsters migrated to where older family members had moved earlier. The farmers from the inner parts of the fjord primarily wanted land and so emigrated, while the fishermen further west traditionally traded with Bergen and had a network of contacts there.

**Forced Migration during World War II**

Apart from the individual relocation of prisoners and persons with mental illness, migration under compulsion in modern times is linked to events during World War II. Most dramatic was the evacuation of nearly 50,000 people from northern Troms and major parts of Finnmark in the fall of 1944 after the Red Army liberated parts of Norway, but also the arrests of opponents against the Nazi regime with imprisonment in different camps reached a significant amount. Step migration with temporary residence inside Norway for the approximately 50,000 who fled to Sweden or the United Kingdom was also usual. Involuntary detention of the population of Telavåg south-west of Bergen in the spring of 1942

---

4 The maps and the 1865 statistics are accessible via the internet in my article in the book *Folketellinger gjennom 200 år*, see url http://www.ssb.no/befolkning/artikler-og-publikasjoner/folketellinger-gjennom-200-aar?fane=om
included fewer persons (estimated at 341), but was all the more dramatic because of the destruction of the entire settlement as a vengeance for their relations with the United Kingdom. The shipment of almost 700 teachers to Kirkenes by the Russian border lasted more briefly, from mid-April to the autumn of 1942. The latter two deportations happened on the background of the Norwegian resistance struggle being stepped up, shortly after the expansion of WWII into the Soviet Union and the Pacific caused more desperate warfare from the Germans.

The destruction and migration of Telavåg

The two cases of forced geographic mobility in 1942 led to significant suffering for those involved. However, they were in principle fundamentally different because the people from Telavåg were migrants who had no residency to return to and were relocated for two years, while the teachers could return to their homes after half a year’s temporary stay in Kirkenes. Like so many demographic factors in the period from 1920 to 1960, mobility during the war is an understudied area; the quantitative data in the historical literature are approximate and often inconsistent. This is due to the lack of record keeping during the War, and that the nominative manuscripts for the vital statistics and the censuses of 1920, 1930, 1946 and 1950 are difficult to utilize in the archives. The situation is now changing since much of the material has been scanned and posted on the web for certified researchers. Still, much of the evidence from the War rests on oral sources and accounts written by the migrants.

As to the male population from Telavåg, several texts indicate that 72 or 76 people aged 16 to 60 were sent to the concentration camp Sachsenhausen in Germany, where 31 of them died. An early presentation states that 260 women and children were detained in Norway, but it is unclear whether that figure includes men over 60 years (Christensen 1964, 91). The women and children first staid at Storetveit school outside Bergen and were then sent further east to Framnes Folk High School in Hardanger where they stayed for two years. The plans to send the women to forced labour in Eastern Norway were thwarted by medical doctors, exploiting the Germans’ well-known fear of epidemics. After two years, most moved back to the Sotra island they came from, but everyone were banned from visiting their burnt-down Telavåg settlement for the remainder of the war. The surviving men were rescued from Germany with the white buses of the Swedish Red Cross. Telavåg was undoubtedly the southern Norwegian community most affected by
deportation, extinction and loss of half of the male population during the war. It was a slight consolation that, in comparison with the German revenge elsewhere, for example after the liquidation of the SS chief executive officer and Reichsprotektor Reinhardt Heydrich in Czechoslovakia, their reactions in Norway were relatively moderate.

**Deportation of teachers**

By the end of March 1942, approximately 1300 teachers were arrested in places scattered throughout the country, suspected of being at the forefront of the struggle against the Norwegian Nazi party and the Quisling government’s attempt to enrol all teachers in the newly created corporative Norway Teachers’ Union. At the same time, the German backed authorities aimed to impose service obligations for all persons aged 10 to 18 in their Youth League. Most of the approximately 14,000 teachers protested together with 200,000 parents in a letter to the Ministry of Church and Education, who at the same time had a conflict with the clergy. The attempt to break the opposition of the teacher prisoners with punishment drill gave small results, although many became seriously ill. On April 11, 500 teachers from Eastern Norway were marched from the camp at Jørstadmoen by Lillehammer to Fåberg railway station. Here a 17-hour train trip was waiting in congested wagons to Trondheim, while large crowds greeted them at the stations. On April 15, 50 German guards marched them aboard the coastal steamer Skjærstad, except one seriously ill teacher. The small ship was only intended for 150 passengers, but was sent northwards with passengers lacking everything: space, air, provision, nursing, etc. After stopping in Bodø and Tromsø they reached Kirkenes on April 28th; on the last leg the Germans used Skjærstad as a hostage to guard against possible allied attacks against a German convoy. From the harbour, they marched to barracks intended for Russian prisoners of war, getting food and water that caused a lot of disease (Christensen 1964, 86ff). On May 11, another 147 teachers came to Kirkenes with the coastal express ship Finmarken. The forced labour consisted mostly of road construction (also in Finland) and loading / unloading of ships for the front in Russia. In this connection, teacher Olav Hole died in an accident. Support from locals was important in the form of letters, gifts and meals by the road. When the schools reopened after the extra «fuel holiday» on 9 April, it became apparent that the Nazi plans to introduce corporatism built on professional organizations that the Quislings were unable to create due to lack of German support for a government that created con-
The teachers in Kirkenes would suffer more than one loss of life if they met the winter under the prevailing conditions. During the autumn, the teachers were given a chance to sign a pro forma declaration of affiliation with the Teachers’ Union and allowed to return to their homes. In retrospect, it should be stressed that the approximately 7000 Russian prisoners of war in the area were treated far worse, and about 2000 of them died.

The evacuation from Finnmark and North Troms provinces in 1944

The progress of the Red Army, the Finnish capitulation and the withdrawal of the Germans in the Nordic Arctic in the fall of 1944 became the background for the Germans causing the worst man-made disaster in Norwegian history. According to the census of 1930, Finnmark had a resident population of 53308, while the four northern Troms municipalities Kåfjord, Skjervøy, Nordreisa and Kvænangen had 10288. The 1940 population census was postponed to 1946, and then the affected areas had 58790 and 11939 inhabitants respectively. According to vital statistics, Finnmark had a birth surplus of 8898 in the period 1931-44, while the four municipalities in northern Troms had a birth surplus of about 1800. Even if the calculation of the population in 1944 does not take account of immigration and migration due to missing registers, it is clear that the northernmost areas of Norway experienced population decline.

The Norwegian and German Nazi authorities in October 1944 urged for voluntary evacuation of Finnmark’s population, but mainly people who feared reprisals because of their Nazi contacts more than the dangerous journey southwards left. Only after Hitler ordered the use of force on October 28, the evacuation progressed, and the Germans used the tactics of scorched earth almost to perfection. They only saved some churches, and some buildings in East Finnmark that they could not get at due to the advancement of Russian troops. The eight volume Norway at War indicates that between 40,000 and 45,000 persons were evacuated and that between 20,000 and 25,000 resisted German orders and remained as “cavemen” (Eriksen and Halvorsen 1987). Other reports state that between 19000 and 23000 inhabitants managed to escape and stayed during the winter mainly near their homes in Eastern Finnmark liberated by the Red Army. This was partly due to the urge for civil disobedience sent by the Norwegian government in London and partly in fear of what Germans could do while evacuated at sea. An estimated 36817 people were evacuated from Finnmark, and in addition...
to the 22730 remaining, the total amount was 2659 persons lower than the sum of the 1930s and the birth surplus until 1944. With a birth surplus in northern Troms of a few thousand in addition to the 1930 census population we have a population there of about 12300, which is close to the sum of over 12090 evacuated and 391 remaining. Especially in Finnmark, the numbers remain uncertain due to (illegal) refugees to Sweden and the Russian-controlled area in Eastern Finnmark, as well as German raids and arrests. The number of 25,000 persons remaining, cited from _Norway at war_ is probably too high. However, the number of evacuated must have been higher than 45,000, a likely approximation is 49,000 persons. An additional uncertainty in North Troms is that some residents were also evacuated contrary to German plans from areas west of the Lyngen Fjord (Bratrein 1994). Persons evacuated should fill out individual registration cards, which make their numbers more certain than the sum of “cavemen” staying behind.

Out of the nearly 50,000 evacuated, about half remained in Northern Norway while the rest came to Eastern Norway or to Trøndelag in mid-Norway. Tromsø served as an over-populated intermediate harbour and many came there on their own boats. The weather this autumn was unusually benign without any autumn storm. This explains why few died during the evacuation. Arvid Petterson has investigated how many perished at the individual level. There are three main categories: those who died while they were forced to evacuate, those who died while they were “cavemen” due to illnesses and accidents, and those who died while they participated in the War in Finnmark 1944-1945. He lists a total of 339 persons before removing elderly people and is left with 280 who died directly as a result of the deportation and scorched earth tactics. Petterson believes that the number would be significantly higher if this type of survey had been made right after WWII. It has not yet been realistic to review all relevant sources such as the funeral lists in the church books, which are now being transcribed. The overview in the four volume _Our War Victims_ is particularly inadequate when it comes to those who perished in or on their way from northern Norway in connection with forced evacuation. This is because the volumes were based on reports from the survivors, and these have been particularly underenumerated for northern Norway. Nevertheless, the 1946 population census is well matched by the sum of 1930 census figures and the birth surplus in the period between, so the real number of lives lost due to forced migration in the fall of 1944 needs hardly be significantly increased for source-critical reasons.
Historians disagree about the extent to which the government in London encouraged the forced evacuators to oppose the orders of the Germans. *Norway at War* has this comment: “However, neither at home nor abroad were the authorities sufficiently informed about the conditions in the north when they asked people to perform the impossible. An unarmed civilian population could in no way fight the 20th Gebirgsarmé and its 200,000 well-trained soldiers who had been assigned to carry out Hitler’s order.” Arvid Petterson documents that the Finnmark population had access to radio equipment to a significant extent. Thus, those evacuated by roads and at sea could convey information from London. There is, therefore, reason to believe that information from London contributed to the fact that many remained in caves etc, trying to oppose German orders at an early stage of the war. Initially, the parole to oppose German orders about evacuation, and to sabotage aid to those evacuated also created significant difficulties for the nearly 50,000 evacuated. Later, the Home Front and the London authorities understood that the paroles increase the stressful situation for those evacuated.

**Net versus gross migration**

Migration statistics often distinguish between gross migration, which in principle includes all changes of permanent residence, and net migration that includes change of residence since birth or another predetermined time such as marriage. Censuses typically inform us about whether a person lived in his parish of birth, but not about how many times he or she had changed address. Unless we can infer migration from other sources or the birthplaces of the children, the census usually provides net migration statistics only. From 1865 to 1950, the censuses were the main sources of knowledge about the volume of migration in Norway. The nominative census of 1801 did not include place of birth, and the enumerations from 1815 to 1855 were mainly only numeric, thus taken on the aggregated level without birthplaces or other information about migration. The lists of in- and out-migrants that the priests were instructed to include in the church books from 1812, are inadequate and contain only a fraction of the migrants, even those who moved long distances and could not easily obtain a migrant’s pass at a later date. The priest would charge a fee for every pass, and contrary to baptism, marriage and burial, much migration in addition went unrecorded because it cannot be linked to ecclesiastical events.
To compensate to some extent for the undercount in the parish migration lists, Ståle Dyrvik proposed rules for the indirect dating of migration using other information at the individual level from the church books and census records (1983 159ff). The Historical Population Register, when it becomes more complete, will provide for the creation of statistics on gross migration before 1950. As long as we base the statistics on cross-sectional data from the censuses about birthplace and place of residence, the aggregates will be limited to migrants living in a parish or municipality on census day different from where they were born, i.e. net migration. Admittedly, card based population registers were introduced gradually in the towns from 1905 and in all rural municipalities from 1946, but these have not been used to develop migration statistics and are not data processed for statistical purposes (Thorvaldsen 2008). The Danish sources are similar to the Norwegian, but here the birthplace column was introduced in the nominative census already in the 1840s, as the first in the world (Johansen 2002). Danish municipal population registers became mandatory from 1925.

With the introduction of municipal registers throughout the country, it was mandatory from 1946 to send messages about migrants from municipality to municipality. In the 1946 Act on Population Registration, anyone who moved over a municipal boundary was required to obtain a mobility certificate for himself and his family and deliver it to the local Population Register no later than eight days after arrival to the new residence. These certificates were forwarded four times each year to Statistics Norway. This material was aggregated experimentally in 1947 and 1948, and from 1949 Statistics Norway computed detailed statistics about migration each year or for five-year periods (Statistics Norway 1956 page 7). While previous statistics relate to net migration (from birthplace to place of residence in the census), the country now for the first time had an overview of gross migration, i.e. in principle all migration that crossed a municipal boundary. Thus, Norway obtained a type of migration statistics that Sweden had already made during the 19th century on the basis of the catechismal records (Vikström 2003). Thus, with the introduction of population registers throughout the country, Statistics Norway from 1950 could identify individual step migration based on migration reports from the municipal registers. They nevertheless warned that the registration of the moving population was not perfect and that some migrants failed to report all stays over four months, so that a certain number of migration notes were missing (Statistics Norway 1965 184). This explains why *Historical Statistics 1994*
only covers gross migration since 1951 (Statistics Norway 1995 Table 3.29 page 94). At the same time, Statistics Norway dropped producing statistics on domestic net migration in the 1950 census, since it was only four years since the previous census and that gross statistics was considered a good replacement, including aggregates about immigrants. However, it is a pity that we do not have a net migration overview of the extent to which people left their birthplaces, after such a shattering period as the Second World War. Fortunately, the ongoing transcription of the 1950 census will soon allow us to present aggregates about net migration after the war.

Thus, the hundred-year period covered by this article is delimited by both empirical and source-based considerations. From 1865 to 1960, the population statistics based on the censuses can provide overviews of the number of net migration, while the introduction of the Central Population Register from 1967 provides easier access to data on gross migration. This is related to the introduction of electronic data processing in 1952 in Statistics Norway, which was prepared with electromechanical equipment from 1950. From 1960, people could buy a new car without a special permit, indicating that the post-war period was over, and allowing them to settle further away from the workplace, thus stressing the importance of detailed migration statistics.

**Net migration between place of birth and residence in Norway**

Statistics Norway summarized developments regarding permanent change of residence since 1865 in the publication of aggregates from the 1920s (Statistics Norway 1923). This summary was related to the publication of the overview of Norwegian overseas emigration (Ministry of Social Affairs, 1921), a context emphasized in publications on internal migration. The definition of residence was, of course, not the same from census to census. From 1875, Statistics Norway distinguished between permanent and temporary places of residence, while its forerunner, the Table Office in 1865 only counted the people at their formal (de jure) residence (Thorvaldsen 2004). In the latter case, all statistics was based on the resident population – those temporarily present were not enumerated. For unknown reasons, Statistics Norway provided population numbers for the present (de facto) population as the basis of the migration statistics from the census of 1875 to 1910, while in 1920 again chose the de jure or resident population. In general, the difference between the number of people formally and actually present was not big at the national level, a half or a couple of percentage points in 1875 and
1920 respectively. But the difference was larger in some municipalities, for example with large fishing populations. Also, the definition of birthplace municipality has weaknesses. Controls against the church books show deviations between the place where a person was baptized and the place of birth specified in the census, and the deviations can only be explained to some extent by disparities between the parish and municipal boundaries. The many municipality border changes during the period may affect the indication of birthplace to the extent that persons were registered with a municipality that had ceased to exist at the time of the census and might have been merged with the municipality of residence after his or her birth (Thorvaldsen 1996a). Statistics Norway returned to the analysis of “migrations” at the national level after the transition to statistics based on gross migration (1965).

The trend in the evolution of migration at the national level, as shown in Figure 2, is nevertheless feasible. The figure distinguishes between urban and rural municipalities according to the formal urban status criterion that was used throughout the period, indicating for each municipality type how many inhabitants who were born in the municipality of residence, in other Norwegian municipalities and abroad. The latter category was still small; it was typically less than a tenth of the migrants and was falling from the turn of the century following a slight increase in the late 1800s. The urban population grew from about ¼ to
over one million inhabitants throughout the period, more due to migration than several small towns receiving urban status during the period. The proportion of in-migrants increased in both urban and rural municipalities, and the relative increase was greatest in rural municipalities. In addition, the internal migration had an indirect fertility effect because many of the migrants were young and a short period after moving, gave birth to children.

**Conclusion**

This article attempts to summarize important aspects of internal migration in Norway from 1865 to 1960. Earlier than this period, mapping internal migration is difficult because the censuses lacked a column for birthplace, and the in-/out-migration lists included in the church books from 1812 only contain a small part of the actual migrants. From the 1960s, the construction of a Central Population Register facilitated the computing of gross migration statistics, meaning that the unit counted was every move across an administrative boundary resulting in a new permanent address. This was attempted already from the late 1940s, but the municipality based population registers missed much migration. Thus, from 1865 to 1960 the main statistical series was about net migration, using the birthplace and place of residence information in the decadal censuses to measure net migration. These aggregates show that men more often than women were migration pioneers, especially over longer distances. But by 1920 it was as usual for women as for men to be internal migrants in Norway. The censuses are also the main sources for tracking specific flows of migration, especially from landlocked parishes in southern Norway to less densely settled areas in the north, particularly in the valleys of Målselv and Bardu south-east of Tromsø. These flows can be regarded as forerunners of migration from the eastern part of Finnmark to the Kola Peninsula in the decades around 1900, which of course must be classified as emigration.

More dramatic cases of geographic mobility took place during World War II, when the Nazi authorities launched several cases of forced migration. The population of Telavåg south-west of Bergen was punished in 1942 by resettling women and children to the fjords further east, while the adult men were sent to concentration camps in Germany and their houses burnt. Hundreds of teachers were at the same time sent into forced labour at Kirkenes by the Russian border to make them accept membership in the corporate Nazi trade union. They accepted in
order to return home, but this remained a formality with no practical consequences. Most catastrophic was the evacuation of nearly 50 000 persons from the northernmost provinces of Finnmark and Troms when the Germans used scorched earth tactics to stop the advancing Red Army in the autumn of 1944. These more special cases of migration cannot be studied in detail in censuses or church records, we rather depend on interviews and publications by the people involved in such geographic mobility.

References


Hägerstrand, T. (1953). *Innovationsförloppet ur korologisk synpunkt* [The process of innovation from a corological point of view]. Lund, Lunds universitets geografiska institution.


Niemi, Einar (1994). *Den lange, lange sti over myrene og ind i skogene, hvem har trakket op den?* [The long, long path over the bogs and into the woods, who has created it?] in *Drivenes, Hauan og Wold* (red.), *Nordnorsk kulturhistorie*. Oslo, Gyldendal.


Tysdal, Olav. (2013). Demografiske konsekvenser etter kollapsen i vår-sildfisket i 1870 [Demographic Consequences after the Collapse of the Spring Fisheries in 1870.” Heimen 50, 1: 3–24


Chapter 10.
DEATH IN THE CITY: A VIEW FROM 19TH CENTURY CHURCH REGISTERS IN NORWAY

Hilde L. Sommerseth,
Evelien C. Walhout

Introduction

This paper explores the causes of death in church book burial registers from Trondheim city in Norway during the late 19th century. Whereas most studies on trends in cause specific mortality have focused on aggregate data, we shall provide a more in-depth analysis by using individual-level causes of death for the period 1866-1895. According to Norwegian Official Statistics (NOS), the crude death rate in Norway’s three largest cities, Kristiania/Oslo, Bergen, and Trondheim drifted around 25 per thousand, and over time we see a stagnation in the mortality rate until the late 1890s at which time a marked and steady decline set in, only interrupted by the Spanish flu epidemic of 1918/1919. Trondheim was a quite deadly place for infants, with an infant mortality rate fluctuating around 150 to 160 infant deaths per 1000 live births until the late 1890s. The breakthrough of a consistent decline came at the turn of the 20th century.

In this paper we are particularly interested in reporting practices and trends in cause of death registration during a period characterized by urbanisation and industrialisation. Previous research has shown that individual causes of death significantly varied in terms of time and place regarding the way causes of death were recorded by priests or doctors as well as categorised by the authorities (Reid et al. 2015; Walhout 2010). Variations in the way the causes of death were recorded can be affected by the medical and church personnel responsible for reporting and recording, the likelihood of a doctor seeing a patient before

---

1 This paper is a part of an ongoing research project about individual causes of death in Norway and the Netherlands. Part of the text, in which the dataset and the coding are described, corresponds with sections in our paper currently under review in Social History of medicine, entitled The Gendering of Disease. A century of classification practices regarding male and female causes of death in Norway and the Netherlands, 1839-1938.
death, medical knowledge and local terminology. This study provides a first careful evaluation of the causes of death through examination of local recording practices done by the priests and of age-specific profiles of individual causes. The explanations for the findings are considered in the context of medical infrastructure.

In the following pages, we first present the dataset and methods, followed by a description of the setting and results. The paper concludes with a short discussion and suggestions for further research.

Sources

The Trondheim database consists of individual level data transcribed from the church book register containing burials. It covers the period 1866 to 1895 from three different churches, namely Vor Frues kirke, Domkirken and Ila kirke. The 30-year period includes 11,146 deceased individuals (stillbirths are excluded), where 8,474 (76%) are registered with causes of death by the priest. In addition, the dataset provides information about sex, age, street address and occupation of the deceased. Occupations are only given for deceased men, or the father/husband of the deceased.

From 1820, the church register included a specific entry where the priest had to report, “if the death was caused by an infectious disease or in case of an accident”. An extensive revision of the register rules was introduced in 1877, stating that all causes of death had to be registered along with information about whether or not the doctor had visited the deceased prior to or upon death. Clearly, the church register had a two-fold agenda. In addition to being a register for sacral events dating back to the late 17th century, an early 19th century act instructed the church to report yearly numbers of births, marriages, and deaths to the authorities as the basis for vital population statistics. From 1866 onwards, these aggregates were replaced with nominative lists, providing a more accurate basis for statistical purposes. Causes of death were all along a part of the medical statistics – provided not only by the priest, but from 1853 also increasingly by the medical profession. By comparing the statistical reports from medical doctors and priests, an immediate impression is that priests to a much higher extent were present to determine the cause of death of the deceased, at least in the 19th century. Around 1860 only about 40 percent of all deaths in Norway were represented in the medical profession’s reports, increasing to around 80 percent in 1900 (Pedersen 2007).
Without doubt, these figures hide significant regional variations – and perhaps profound rural-urban differences. Previous studies on infant mortality in typical Norwegian 19th century rural areas reveal that doctors visited patients rarely, and that the registration of causes of death was for the priests to decide (Fure 2004; Jåstad 2003). This was mainly due to long distances to travel for the doctors. Around 1860, Norway had 315 doctors and the country was divided into 79 urban and rural districts. About 40 years later, the population had increased by 40 percent and 1138 doctors covered 158 districts (Schiøtz and Skaset 2003, p. 143). Still, in the rural areas, a medical doctor often had to travel several days if someone summoned medical help. In cities things were different. First, distance did not matter that much, and secondly, cities employed more doctors. For example, in 1880 the ratio of medical doctors per person in the rural districts around Trondheim was 1:4,700 compared to 1:1,200 in the city. With relatively easier access to medical help, we need to ask to what extent our causes of death echo the priest’s medical knowledge, or if it was merely a duplicate of the medical doctors’ death certificates.

A second matter, which pinpoints this question, is related to the Norwegian Health Law of 1860. This law directed every municipality to have a health commission in charge of health politics and campaigns. In rural areas, the main purpose was to educate people about better personal hygienic standards, while in the cities the focus was to eliminate the risk factors, which by contemporary observers was seen as critical. Hence, in the rural areas the appointed commission members consisted more or less of the same people that had been working on health related issues before the Law was introduced, that is to say the members from the civil service (teachers, police, priests and doctors). The composition in the cities underwent major changes. Here the commission’s members reflected specialization within health and city planning and engineering, priests slowly losing their health related role (Schiøtz and Skaset 2003, pp. 45-46). Consequently, it is reasonable to believe that the medical profession in the cities at some point during the 19th century effectuated control over the registration of causes of death, and increasingly had a say on how they should be registered in the church registers. Such a development may be seen in Bergen (the second largest city in Norway) where causes of death rapidly disappeared from the church register from around 1880. Interestingly, this did not happen in Trondheim.

To what degree can we rely on the cause of death registration? Today we take for granted that medical tests or autopsies are the basis
for defining the cause of death, usually with great accuracy. A similar accuracy cannot be expected when interpreting historical records. The doctor or the priest might not have seen the deceased, and had to rely on descriptions given by relatives and neighbours, or for instance rely on the probability that the epidemic diseases that currently swept the area had taken yet another victim. In retrospect, we know that the Norwegian priests’ registration of causes of death was gradually taken away from them. Some priests might have been satisfied with less responsibility, others might have been disappointed. The medical profession was in continuous development during the 19th century, and this can for example be viewed by the changing classification schemes for causes of death, the nosologies. Several scholars have discussed how changing nosologies affected the labelling of the causes of deaths.\(^2\) In what way did the priest or doctor comply with these changes? In addition, the fact that both the priest’s and the medical doctor’s documentation of death causes had a statistical purpose, leads us to ask to what extent the development of classification schemes for death causes was influenced by the different areas of epidemiology.

**Coding the causes of death**

For coding and classifying the causes of death, we used a classification scheme entitled Cph1876+ (version 1.0 June/2016, cf Revuelta-Eugercious and Løkke (2016)). This classification is hierarchically structured, and concerns a rework of a system developed by medical historian Josep Pep Bernabeu (2003). As opposed to the standard tool for classifying diseases in current-day societies, the International Classification of Disease (ICD), our coding and classification tool leaves more room for specifically historical causes of death and older rationales.\(^3\) The 141 categories are clustered following an aetiology or mode of infection and transmission. It distinguishes between six main groups at the highest-level: (1) infectious diseases; (2) non-infectious diseases; (3) external caus-

---

\(^2\) For an overview, see Reid et al. (2015).

\(^3\) See codes of ICD-10 on the webpage of the World Health Organization: http://www.who.int/classifications/icd/en/ (visited 3 December 2017). Currently, a historical version of ICD is developed, see also: Alice Reid et al., ‘Coding and classifying Individual Cause of Death Data from Scotland: the nineteenth century and beyond’ and Maria Hiltunen and Sören Edvinsson, ‘Classifying Literate Cause-of-Death information originating from Swedish Historical Parish Registers’, papers presented at SHiP-workshop, 27 March 2018, Cambridge.
es; (4) older rationales; (5) ill-defined, and (6) not reported/blank. The first group includes five sub categories. First, infectious diseases transmitted mainly by air and direct human contact (summarised as airborne diseases) include infections typical of childhood, such as smallpox and whooping cough, infections in the respiratory system such as bronchitis, tuberculosis and other acute respiratory diseases such as influenza and pneumonia. The second group of infectious diseases includes causes of death indicating ‘water- and food-borne infectious diseases’. This group includes cholera, typhus and typhoid fever, acute digestive diseases such as diarrhoea and dysentery. The third group of infectious diseases refers to ‘vector transmitted infectious diseases (including human vectors)’ in which causes like puerperal fever, exanthematic typhus and yellow fever are placed. The fourth group consists of infectious diseases related to more vaguely described causes of death, for which it is impossible to determine to what extent the cause of death is related to one of the above-mentioned categories. Finally, a group was defined as ‘other specific infections’ including anthrax and tetanus.

The second main group of ‘non-infectious diseases’ includes sixteen subgroups, mainly subdivided into causes related to different parts of the body, such as the brain, lungs, stomach, the genito-urinary system, but also classifications such as deficiency diseases, perinatal pathologies and cancer. The third main group comprises all external causes of death: accidents, burnings, suicides, homicides, et cetera. The fourth main category, ‘older rationales’, refers to causes that are usually assigned as ‘ill-defined’ causes, like ‘malnutrition’, ‘sudden death’, ‘eclampsia’, ‘lack of breastfeeding’, ‘teething’, ‘oedema’ and ‘old age’. The importance of including these types of causes was emphasised in recent research, where the authors show that the increase in deaths caused by cardiovascular diseases in fact was a result of a rebranding of deaths from ‘old age’, and not a real increase (Reid et al. 2018, p. 330). Besides the four main categories of disease we also regrouped all causes of death that were either badly specified, ill-defined, illegible or ‘unknown’ causes. A final sixth group consists of not reported causes of death - when the original source contains a blank field.

**Trondheim city**

Depicting the 19th century, Trondheim had long been an administrative and religious centre for the north of Norway, the main seat of the County of Sør-Trønderlag and the Bishopric of Nidaros (from 990).
During the 19th century, the population in Trondheim increased more than fourfold: from 8,800 inhabitants in 1801 to 38,000 in 1900 (Danielisen 1958, p. 277–287). This increase included extensions of the city boundaries. Compared to population growth in the two other large cities of Norway (Bergen and Oslo), we see that Trondheim increased more than Bergen, but less than Oslo in relative terms (the latter from 9,000 in 1801 to 225,000 at the turn of the 20th century). A mix of birth surplus and immigration led to booming population growth. Consequently, at the end of the 19th century, one rarely found households where all members were born in the city.

The role of the town as an economic centre for the region was reinforced during the 19th century, when most of the goods entering and leaving the region passed through Trondheim. In 1880, the town was connected by railway southwards to Oslo and in 1882 eastwards to Sweden, and the opening of the Coastal Express ship route in 1893 made the optimists talking about making Trondheim “The Hamburg of the North”. Huge plans for extending the port facilities were on the drawing board. However, only parts of the plans were ever realized (Sandnes et al. 1995, p. 51–70).

Nevertheless, the growth in communication was accompanied by a marked increase in the town’s trade and industry. Originally, Trondheim had been a centre for exporting timber, fish and ore. Up to World War I fish exports increased particularly rapidly, but imports increased even more and both retail and wholesale trade flourished. In order to explain the background for these developments we must look beyond Trondheim. There was a major change at that time in the rural areas associated with the town. The agricultural system changed from a self-supporting economy to a cash economy and became dependent on having a market. The need for banking and goods such as fertilizer, seeds, farm machinery, building equipment and groceries increased rapidly. To a greater extent than before, Trondheim would send on its excess production, and at the same time it provided not only its own region but also parts of Northern Norway with goods and services (Sandnes et al. 1995, p. 55).

Already during the first part of the 19th century, great improvements were made concerning sanitary and hygienic matters. Water was piped under the town already in the 1850s, and from 1880 these were extended to also provide water inside houses. City development during this era of modernity, where political and economic ideas and prospects for future profit were decided and implemented at high speed, increased
social inequality and an increased awareness of such. One thing concerns access to water, where the majority of the city population had to wait until well into the 20th century, and the luxury of a warm bath and a shower was for the few to enjoy. Nevertheless, improvement of personal hygiene was a top priority for the Sundhedskommission (Health commission) during the last decades of the century, and related to their work, two public bathhouses were built. The bathhouses’ popularity can be viewed from the medical doctor’s report from 1903, where as much as 120,000 bath visits were registered in Østre- and Vestre folkebad during that year (Tretvik et al. 2007, p. 152). Another thing are the living conditions. The lower working class families clustered themselves in the eastern part of the city, their houses were small and draughty, and the living space sparse. Fertility control was still a future phenomenon, and from the 1900 census one gets the impression that about every house address in the city hosted lodgers.

The city center, where the families of the bourgeoisie (the patrician elite) and the upcoming middle classes lived, underwent changes too. While the patrician elite had built their fortune on export, the new wine in the city – the middle class – built their swelling assets through import. From 1880 a plan for disposal of sewage was carried out with sewage pipes, and many of the familiar ‘city smells’ disappeared. Having said that, in the last decades of the 19th century the city council received complaints about the stench when farmers carried the open wagons with human excrements from the city’s innumerable toilet bins (Sandnes et al. 1995, p. 37–50).

Despite economic growth and city expansion, the city census from 1885 reveals that no less than 215 pig houses and 97 barns existed within the city boundaries, with approximately 900 farm animals (Danielsen 1958, p. 14).

From 1860, with the Sunnhedslov (Law of health), yearly statistics are available on the population per medical doctor, midwife, dentist and assistant vaccinator, for the cities and the rural areas. In 1880, there were 19 medical doctors (1:1,200), 10 midwives (1:2,000) and two dentists (1:11,000) in Trondheim. It is clear, when looking at the figures for 1900, that the city authorities did not manage to follow up or improve consistently the medical coverage due to the enormous population growth that occurred during these two decades: 1:1,400 and 1:1,300 for medical doctors and midwives respectively.
Results

The Trondheim database contains 11,146 deceased individuals (stillbirths are excluded) from three different church book registers in Trondheim (Vor Frue Kirke, Ila Kirke and Trondheim Domkirke), where 8,474 (76%) of the registrations contain information about the cause of death. The raw data consists of 2,773 unique causes of death – including all types of spelling, which have been standardized and coded according to the Cph1876+ classification scheme. Of the 141 categories, 127 unique codes were used for the Trondheim data.

Figure 1 shows the yearly total numbers of death from two different sources, the NOS and the church book register for the period 1866-1895. In addition we have made a distinction between burials with and burials without a registered cause of death. As we do not know the population at risk we present the numbers of deceased individuals per year. The grey line shows the total numbers of deceased for whom we have information on the cause of death; the yellow line represents the total

![Graph showing yearly deaths, burials, and registered causes of death from 1866 to 1895.]

Figure 1. Total numbers of deaths from Norwegian Official Statistics (NOS), burials in Trondheim, and burials with registered causes of death, by year of death
number of Trondheim deaths per year, and the orange line denotes the total number of deaths in the official statistics. The gap between the total number of deceased in the NOS and the church book register is due to missing transcribed data from the Bakklandet parish, which became part of the city in 1863. The trends are similar over time, except for a slightly increased gap during the last couple of years. Although we do not have a 100 per cent registration of causes of death, we observe that the number of cause specific Trondheim deaths more or less follows the pattern of the overall number of deaths. The exceptions are for a missing peak in 1877 and a pronounced gap in the years 1884–1887, which can be linked to an absence of causes of death in one of the churches: the Trondheim Domkirke. However, we have a full coverage for the period 1866–1875 and 1881–1883.

The peak of deaths in 1875 is related to pulmonary phthisis (tuberculosis), pneumonia and measles, and the 1881–1883 peak is related to pulmonary phthisis, pneumonia and acute bronchitis. The peak in 1889-1891 is mostly related to diphtheria and pulmonary phthisis, being the two really big killers in this period.

![Figure 2. The distribution of individual classified causes of death according to Cph1876+ by year of death](image-url)
Figure 2 shows the distribution of classified causes of death by year of death. Overall, the category of infectious diseases dominates. The relative numbers of non-infectious diseases and deaths due to external causes (accidents being the largest group) does not change much during the 30 years covered, which may indicate that the vast majority of the ‘not reported’ cases is due to infectious diseases. We should, however, not rule out that a part of the ‘not reported’ category includes an increase in the number of deaths due to non-infectious diseases. As expected we see a decline in the use of ‘older rationales’ from about 1883.

Figure 3 shows the number of classified causes of death by age group. The distribution for the youngest age groups follows the expected pattern, with more infant deaths compared to children and adolescents. The number of deaths in age group 25–49 is high, but it is within an expected range since it comprises a wider age range than the younger age groups. The equally wide range age group of 50–75 forms the largest group. Interestingly, we see that some types of diseases are more pronounced in certain age groups than others.

Infectious diseases dominated child deaths (1–4 years), counting for about two-thirds of all deaths. A lower proportion of infec-
tious disease among infants indicates that infants were breastfed, and were to a much higher degree protected during the first year of life, but when the toddler was weaned, he or she became more exposed to infectious diseases. Among adults aged 25–49, infectious diseases were a major killer, causing about 45 per cent of all deaths. Some 25 per cent of all deaths in this age group were due to pulmonary phthisis.

Non-infectious diseases were most significantly related to adult deaths, and especially so for the age group of 50–74. The third group of age specific causes of death relates to ‘older rationales’, which will be discussed more thoroughly towards the end of the paper. Here we observe a clear clustering at both ends of the age distribution. Among infants, we find that 20 per cent of all infant deaths were registered with konvulsioner/kramper/eclampsia as the main cause of death. Among the elderly, the most frequently used older rationale was dying of ‘old age’. Assigning causes of death to the active part of society, the working population, apparently was a more straightforward task for the clergy.

Although deaths related to external causes formed a relatively small proportion of deaths in all age groups, we do find that this classification was more or less absent in ages up until 15 and for those who died at age 75 or older. This indicates that external causes often were related to a dangerous working environment, especially among men (Sommerseth and Walthout 2019). The fact that the proportion of ‘not reported’ causes of deaths is more or less evenly distributed across age groups, strengthens the assumption that the missing registration was not due to unknown causes, but rather that the priest decided not to register at all.

Figure 4 shows the nine (out of 36) most frequent causes of death by age group as percentage of total deaths per age group. The nine most frequent causes of death counted for about 75 percent of all deaths in age group 1–4 where airborne infectious diseases killed nearly 60 percent of the children. Airborne diseases were the main killer for ages 1–49. Cancer and tumours were the second most frequent causes of death in age group 50–74. Comparing the two oldest age groups, where ‘old age’ represents more than half of the nine most frequent causes of deaths (and 35 percent of all causes of death in this age group), the immediate impression is that the term ‘old age’ probably included cases of cancer/tumours as well as airborne infectious deceases.
Figure 4. The most frequent causes of death groups by age group.
Percentage of deaths in age group

Table 1
Total numbers of airborne infectious diseases causing death by age group

<table>
<thead>
<tr>
<th>Infections typical of childhood</th>
<th>0-1</th>
<th>1-4</th>
<th>5-14</th>
<th>15-24</th>
<th>25-49</th>
<th>50-74</th>
<th>75+</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Smallpox</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>2</td>
</tr>
<tr>
<td>Measles</td>
<td>23</td>
<td>42</td>
<td>9</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>76</td>
</tr>
<tr>
<td>Scarlet fever</td>
<td>23</td>
<td>161</td>
<td>53</td>
<td>6</td>
<td>3</td>
<td>0</td>
<td>0</td>
<td>246</td>
</tr>
<tr>
<td>Whooping cough</td>
<td>51</td>
<td>57</td>
<td>9</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>119</td>
</tr>
<tr>
<td>Diphteria</td>
<td>17</td>
<td>103</td>
<td>93</td>
<td>31</td>
<td>5</td>
<td>0</td>
<td>0</td>
<td>249</td>
</tr>
<tr>
<td>Croup</td>
<td>10</td>
<td>85</td>
<td>18</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>116</td>
</tr>
<tr>
<td>Infections - respiratory system</td>
<td>22</td>
<td>6</td>
<td>3</td>
<td>2</td>
<td>6</td>
<td>15</td>
<td>10</td>
<td>64</td>
</tr>
</tbody>
</table>
Chapter 10

<table>
<thead>
<tr>
<th>Disease</th>
<th>0-1</th>
<th>1-4</th>
<th>5-14</th>
<th>15-24</th>
<th>25-49</th>
<th>50-74</th>
<th>75+</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Throat infection</td>
<td>6</td>
<td>8</td>
<td>1</td>
<td>3</td>
<td>1</td>
<td>3</td>
<td>0</td>
<td>22</td>
</tr>
<tr>
<td>Rheumatic fever</td>
<td>0</td>
<td>0</td>
<td>2</td>
<td>7</td>
<td>9</td>
<td>1</td>
<td>0</td>
<td>19</td>
</tr>
<tr>
<td>Laryngitis</td>
<td>0</td>
<td>3</td>
<td>3</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>6</td>
</tr>
<tr>
<td>Bronchitis</td>
<td>150</td>
<td>130</td>
<td>13</td>
<td>5</td>
<td>18</td>
<td>107</td>
<td>75</td>
<td>498</td>
</tr>
<tr>
<td>Pneumonia</td>
<td>83</td>
<td>73</td>
<td>34</td>
<td>25</td>
<td>123</td>
<td>244</td>
<td>82</td>
<td>664</td>
</tr>
<tr>
<td>Pleurisy</td>
<td>4</td>
<td>2</td>
<td>2</td>
<td>4</td>
<td>4</td>
<td>5</td>
<td>1</td>
<td>22</td>
</tr>
<tr>
<td>Influenza</td>
<td>3</td>
<td>2</td>
<td>0</td>
<td>0</td>
<td>5</td>
<td>10</td>
<td>6</td>
<td>26</td>
</tr>
<tr>
<td>Tuberculosis</td>
<td>21</td>
<td>42</td>
<td>45</td>
<td>40</td>
<td>68</td>
<td>35</td>
<td>0</td>
<td>251</td>
</tr>
<tr>
<td>Tubercular meningitis</td>
<td>26</td>
<td>62</td>
<td>41</td>
<td>12</td>
<td>15</td>
<td>1</td>
<td>0</td>
<td>157</td>
</tr>
<tr>
<td>Pulmonary phthisis</td>
<td>9</td>
<td>31</td>
<td>84</td>
<td>256</td>
<td>497</td>
<td>259</td>
<td>27</td>
<td>1163</td>
</tr>
<tr>
<td>Total</td>
<td>448</td>
<td>809</td>
<td>412</td>
<td>394</td>
<td>756</td>
<td>682</td>
<td>201</td>
<td>3,702</td>
</tr>
</tbody>
</table>

Table 1 shows the total numbers of airborne infectious diseases causing death by age group. Among the infants, we see the typical childhood diseases, with measles, scarlet fever, whooping cough, diphtheria, and croup, though to a lesser degree than among older children. The main airborne disease that became fatal for infants was bronchitis. Once they survived the first year of life, we see an increase in the total number of children dying of airborne infections, scarlet fever and bronchitis being the two foremost causes of death. From age group 5–14 pulmonary phthisis (tuberculosis) was more than doubled compared to the younger age group, and the number of deceased rapidly increased until peaking in age group 25–49. The second largest airborne cause of death in this age group, was pneumonia. As previously mentioned, airborne infectious death causes experienced a relative decline in the oldest age group, compared to the young age groups, and we observe that the most substantial decline occurred for the pneumonia and pulmonary phthisis victims. Looking at the age pattern, and how the specific airborne infections are distributed throughout the life course, bronchitis shows a characteristic u-shape with its lowest level in age group 15–24. Pneumonia has a similar distribution, although not that steep. Tuberculosis, with no further specification, has a somewhat similar age distribution as pulmonary phthisis, with an increased number of deceased as
age increases until 25–49. A different pattern emerged among the tuberculosis meningitis victims, where the bacteria spread to the meninges, which are the membranes surrounding the brain and spinal cord. Here we observe a peak at age 1–4.

Figure 5. The total numbers of deaths in the three top causes of death classifications in the group ‘older rationales’, by five year periods 1866–1895. Norwegian in legend is explained in the text.

Figure 5 shows the total number of deaths from the top three causes of death in the group ‘older rationales’, namely ‘kramper’ or ‘convulsions/spasms/eclampsia’, ‘concerning infancy’ and ‘old age’ (93 percent), by 5 year periods. The classification system defined six historical terms that were used as causes of death ‘concerning infancy’. First, the broad category ‘concerning infancy’ is used as a code for terms like ‘nyfødt’ (new born). In Trondheim we only find one incidence where this code is used. Instead, 90 out of 103 were coded as ‘atrophia infantilis’. The Latin word ‘atrophia’ refers to either ‘atrophy’ or ‘wasting consumption’, a condition causing a person or a part of the body to become progressively weaker and emaciated. In Greek it means ‘lack of food’. The third code is ‘dentition’: teething. Some twelve infants were diagnosed with this cause of death. We found no cases of the following classification codes: ‘lack of breastfeeding’, ‘thrush’, and ‘lack of care’.
In the literature, ‘kramper’ (the most used term) is described as one of the main ‘old rationale’ for deaths in infancy. The concept describes the deathbed, with a condition that could occur without any other sign of disease. Contemporary medicine related ‘kramper’ to be caused by teething or if mothers breastfed during strong emotional disturbance (Løkke 1998). While about 38 percent of all infants’ causes of death were registered as ‘kramper’ until 1875, the expression gradually disappeared from the nomenclature towards the end of the period, although not entirely. About 9 per cent of infant deaths were registered as dying of ‘kramper’ in the two last periods, respectively.

Of the 625 individuals dying of ‘old age’, ‘alderdomssvakhets’ (old age weakness) was the registered term in 267 of the cases (43 percent) and ‘alderdom’ in 206 of the cases (32 percent). For Denmark, Anne Løkke finds that ‘alderdomssvakhets’ replaced the term ‘alderdom’ with the implementation of a new nomenclature in 1876, and suggests that age as such was no longer a sufficient explanation of death. Instead, the deceased had to be struck by a weakness that distinguished him/her from others of the same age, but on the other hand didn’t need any further explanation than age (Løkke 1998). We do not find a similar trend in Trondheim. Here, ‘alderdom’ and ‘alderdomssvakhets’ appears in the register from 1866, and are used throughout the period of study. The two terms even appear on the same page in the register, written by the same priest. The Latin term for old age, ‘senectus’, was used 59 times (9.5 percent). The use of a Latin term in comparison with a vernacular term, did not differ according to sex, age or social class. Widows on poor relief (fattiglem) died of ‘senectus’ as well as shipmasters.

Some 57 of the death causes classified as old age had an additional cause of death, though all of them had ‘alderdom’ or ‘alderdomssvækkelse’ as the primary cause of death. For example, a person could die of old age with an additional water or airborne disease, such as diarrhea or bronchitis.

Common for all three categories is the specific relation to age. While ‘kramper’, ‘Atrophia infantilis’ and ‘alderdomssvakhets’ refer to a physical ailment upon deathbed for a certain age group, ‘alderdom’ is used as a cause of death implying that death was perceived as something expected for the age group without any further explanation.
Some preliminary conclusions

It is still too soon to draw well-founded conclusions from this analysis of disease classification in Trondheim city, but the following observations seem to be justified.

The various individual-level causes of death data registered in the church book registers appear to provide a way of describing the disease environment and the manner in which they varied and changed over time and across age groups.

The disease environment in Trondheim during the last part of the 19th century was dominated by airborne infectious diseases. Mortality appears to have been accentuated in Trondheim because of the effects of crowding so that infections and contagion could spread easily among a large and mobile population. The relatively low incidence of water and foodborne infectious diseases, such as diarrhoea and dysentery, indicates that the water pipes implemented already in the 1850s and from 1880 also inside houses, had some effect, in addition to the top priority on personal hygiene addressed by the city’s Health Commission.

With access to individual-level causes of death data, we were able to relate specific disease patterns according to individual attributes. In this paper we have focused on the causes of death across age groups. We found that cause of death classification related to ‘older rationales’ had a clear clustering at both ends of the age line. In addition, airborne infectious diseases had a significant presence among children age 1–4, killing nearly 60 per cent.

Treated critically, the data generates interesting insights into the epidemiological history of a period where cause of death statistics from doctors was still scarce in most of the country. Medical doctors were mostly active in the cities, while 80 per cent of the population lived in rural areas where doctors rarely visited patients on their deathbed. This make cities excellent laboratories for comparative studies of causes of death provided by priests and doctors, which should be a task for future research.

References


Chapter. 11
INFANT MORTALITY IN THE LATE NINETEENTH AND EARLY TWENTIETH CENTURIES URALS: MACRO AND MICRO ANALYSES

Dmitrii Bakharev,
Elena Glavatskaya

Introduction

Infant mortality rate (IMR), which reflects the chances of survival for an infant under one year of age, is an important indicator of the society’s welfare. The IMR is also considered as one of the most accurate indicators of health care and the country’s overall socio-economic development and stability (Thorvaldsen). The IMR is particularly significant for studying transition periods, in our case – modernization of late Imperial Russia. In the late nineteenth and early twentieth century, the Russian Empire persistently occupied the first place in Europe in terms of infant mortality rates, which estimated 250 pro mille or higher while in Scandinavian countries it had already dropped to 70–80 pro mille (Kurkin, p. 6).

Demographic development in the Russian Empire after the abolition of Serfdom in 1861 attracted much scholarly attention both in Russia and abroad, however in most cases scholars focused their studies either on national or provincial levels-compatible in size with a small European country. Spatial and other differential analyses of infant mortality in Russia are underrepresented among international historical demography studies.

Our study analyses the late 19th – early 20th century Perm’ gubernia’s infant mortality trends and focuses on the urban population, which remains largely neglected by historical demographers. That was mainly because national demographic studies relied on the nineteenth and the early twentieth century zemstvo2 and state statistics. While the zemstvo presented mainly aggregated data on rural areas with cities’ population excluded, the state

---

1 The research was sponsored by the Russian Foundation for Basic Research grant No. 19-09-00292

2 Zemstvo – self-government elected sub-province level institution introduced in 1865 to manage local affairs, such as road building and maintenance, improvement of economic development, oversaw medical services and sanitation, public education and other socially important activities.
statistics in its turn did not distinguish between the rural and urban data. Thus, aggregated data made it difficult and in most cases impossible to separate urban statistics from the core of the official data. Meanwhile, the rural population accounted for up to 80% of the Empire’s population in the late 19th century and of infant mortality studies did not take into account a ‘hidden’ minority – urban infants and their mortality rates.

Researchers traditionally focused their studies on the western and central Russia or the main Russian cities – Moscow and St. Petersburg, with eastern parts and provincial cities of the Empire left neglected. Therefore, the study of regional and local differentials is of crucial importance, for it allows looking beyond national averages. While there have been several research efforts made on infant mortality in late 19th to early 20th century Russia (Andreev & Kvasha, 2002; Kvasha, 2003; Avdeev, 2010) and in its regions including the Urals (Golikova, 2012; Kornilov, 2014), none considered urban populations. The usual approach has been to analyse the aggregated data from the source material without considering individual characteristics.

The creation of the “Ural Population Project” (hereafter UraPP) with nominative data transcribed from the Ekaterinburg parish registers (metricheskie knigi) allowed conducting our study on the individual level. Our previous analyses showed that in the decade leading up to World War I infant mortality rates among the Ekaterinburg Catholics and Jews were rather low: 57‰ and 87‰ correspondently, compared to the IMR among the Orthodox majority at 352‰ (Glavatskaya, Borovik, Thorvaldsen). While that article was being prepared for publication, a part of the database that included the nominative data on the Russian Orthodox Church parish members was not ready yet, and we had to use tables prepared by the priests annually for comparison. Now, when the data was transcribed we can study how the IMR varied among the urban and rural population in the late 19th – early 20th century and its dynamics; what were possible reasons for the Perm’ gubernia’s (province) leading position in high IMR.

Perm’ province: geographic and medical setting

Perm’ province with its 12 uezds (sub-provinces): seven in the European part and five in the Asian, embraced a vast territory over 330 000 sq km stretching on both slopes of the Ural mountains. It was the fourth biggest gubernia in Russia with a population of 2 994 302, and only 179 339 of them (6%) were city dwellers in 1897. It was also one of the least densely populated provinces in the European part of Russian Empire with
about 13 persons per sq km. The two biggest cities were Perm’ – the administrative center of the province with a population of 45,205 and Ekaterinburg – a rival city due to its status of the Ural mining and metal production headquarter with a population of 43,239 according to the 1897 Census.

Russian official medical statistic stated that Perm’ province had extremely high infant mortality in the late 19th to early 20th century. According to the health reports collected and published by the local doctors, the infant mortality rate in Perm’ province reached 425‰ compared to a national mean of 270‰ in 1895. We should, however, be aware of the fact that the Russian national IMR calculations included the data only on 50 most western Russian gubernias, including contemporary Estonia, Lithuania, Latvia, Belarus and Ukraine, who had IMR similar to France; and excluding Siberia and Central Asia, whose rates were considerably higher. The deplorable infant mortality rates looked even gloomier as the Russian IMR dynamics compared with other European countries left much to be desired. In 1867–1881, the IRM in Russia, though very high, was better than in Württemberg and Bavaria. By 1901, however, these regions’ IMR decreased, like in other European countries; while in Russia it remained on the same level, which turned it into the European leader in this grim indicator (see Figure 1). A recent international overview ranked Russia with the highest European infant mortality rate also in 1910 (Klüsener et al., 2014).

<table>
<thead>
<tr>
<th></th>
<th>IMR, ‰</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1860-1880s</td>
</tr>
<tr>
<td>Württemberg</td>
<td>312</td>
</tr>
<tr>
<td>Bavaria</td>
<td>308</td>
</tr>
<tr>
<td><strong>European Russia</strong></td>
<td><strong>271</strong></td>
</tr>
<tr>
<td>Saxony</td>
<td>270</td>
</tr>
<tr>
<td>Austria (Cisleithania)</td>
<td>255</td>
</tr>
<tr>
<td>Italy</td>
<td>210</td>
</tr>
<tr>
<td>Prussia</td>
<td>208</td>
</tr>
<tr>
<td>Switzerland</td>
<td>195</td>
</tr>
<tr>
<td>Holland</td>
<td>193</td>
</tr>
<tr>
<td>France</td>
<td>166</td>
</tr>
<tr>
<td>Finland</td>
<td>165</td>
</tr>
<tr>
<td>England and Wales</td>
<td>149</td>
</tr>
<tr>
<td>Belgium</td>
<td>148</td>
</tr>
<tr>
<td>Denmark</td>
<td>138</td>
</tr>
<tr>
<td>Sweden</td>
<td>132</td>
</tr>
<tr>
<td>Scotland</td>
<td>122</td>
</tr>
<tr>
<td>Norway</td>
<td>105</td>
</tr>
</tbody>
</table>

Figure 1. Infant mortality rate dynamics in the late 19th and early 20th century
Figure 2. Infant mortality rate in European Russia, 1893–1896

*Source: (Nikitenko, p. 60–61).*
Naturally, there were also significant regional disparities in terms of the IMR even within the European part of the Russian Empire. As Figure 2 shows, Perm’ province had the worst IMR at that time. In 1896–1897, its IMR was 437 ‰, while the national average was 274 ‰.

Even if the Perm’ province’s IMR fell to 320 ‰ during the first decade of the twentieth century, it remained one of the highest in European Russia (see Figure 3).

![Figure 3. Dynamics of the IMR in Russia, 1867–1910, ‰](source: Rashin, p.195–196).

The growing awareness of the high infant mortality in Russia sparked debates in political and medical circles already in the late nineteenth century (Sokolov, Grebenschikov, p.1–25). Ural doctors regularly discussed the problem of high infant mortality in the region at national medical forums. Doctor Nikolai Russkikh employed by Ekaterinburg zemstvo, presented his paper ‘On Combatting Child Mortality’ at the Twelfth International Doctors’ Congress in Moscow in 1897 and cited the most horrifying figures of infant mortality in Okhansk uyezd, where IMR reached 600‰ in 1890s (Sokolov, Grebenschikov, p. 26).

In addition to regional differences in infant mortality rate varied depending on the type of settlements. The IMR was at its lowest in small towns, usually centers of uezds (sub-provinces), and much higher in rural areas. The highest IMR was recorded in large cities, centers of guberniyas (Gundobin, p. 8). Thus even within a given gubernia IMR had a fragmentary, mosaic-like character with extremely high and low figures, which aggregated together, resulted in an average similar to the metaphorical ‘temperature of patients across the hospital’. The Ural region, which for decades demonstrated high IMR levels, belonged to the type of ‘feverous patients’ and deserves more detailed research, which we shall start with this Ekaterinburg city case study.
Ekaterinburg: population, sanitation and medical services

Ekaterinburg was an important trade and industrial center, being a worthy rival to Perm’ – the administrative center of the province. As a booming center of business and culture, Ekaterinburg at that time was already given an unofficial title of the ‘Ural capital’. In the course of the Great Reforms of the 1860s, the government abolished Ekaterinburg’s status as a mining-factory center, which restricted in-migration and free business in the city. After the mining administration’s control was lifted, Ekaterinburg enjoyed considerable entrepreneurial growth, enhanced even more by the railroad, which was built through the city and connected European and Asian parts of the Empire. The city’s population increased dramatically due to in-migration: between 1873 and 1917, the number of city’ dwellers rose from 30,000 to 70,000. After the abolition of serfdom, thousands of peasants surged to the city in search of employment, others as the contemporary observer wrote, moved to Ekaterinburg in search of the urban life comforts and benefits: schools and colleges, hospitals and clinics, theatres and museums (Vesnovsky, 14–18). A steep population increase due to peasant in-migration and the lack of the necessary infrastructure aggravated the already inadequate sanitary conditions. Streets of the city were covered in litter, wells were polluted by sewage, trash and raw sewage were dumped into the river used by the city’s poor as a source of drinking water (Mikityuk, Yakhno, 132–140).

The Ascension Church parish of Ekaterinburg, whose vital events records we used for this study, was one of the city’s five Orthodox Church parishes, established in 1770. Originally, it had a wooden building but in the late eighteenth century it was rebuilt in stone and has survived to this day along with a few other church buildings in Ekaterinburg. In the early twentieth century, the parish was the second largest in the city with 3536 parishioners in 1909. In general, the parish combined all classes of Russian society with peasants and meschane (townsmen) composing its majority. It also included retired soldiers, a small number of clerics and merchants from the urban poor to the eminent millionaires’ family – the Zlokazovs.

Ekaterinburg’s medical services improved by the end of the 19th century due to the zemstvo’s activities. To target the high infant mortality, the zemstvo initiated the first city’s maternity home foundation in 1877, sponsored by local philanthropists. There were eight midwives officially practicing in the city and many private doctors to help during
delivery. Later they added an ambulance and a gynecology department to the maternity home, which made it more popular. According to local statistics, more than 15,000 women stayed there to deliver babies, 4,000 got treatment because of gynecological disorders and in all, 71,000 had doctors’ appointments within the first 25 years of its operation. Two newly founded schools trained 16 midwives annually (Mikitiuk & Iakhno, 2014, p. 92). In the beginning of the 20th century, the maternity home was reconstructed and got an additional building (see Fig. 4).

![Figure 4. The first maternity home in early 20th century Ekaterinburg](image)

Generally, the number of births conducted with doctors’ or midwives’ assistance steadily increased in the course of 1880–1910 (See Table 1).

<table>
<thead>
<tr>
<th></th>
<th>1890</th>
<th>1900</th>
<th>1910</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ekaterinburg uezd</td>
<td>819 (3.8%)</td>
<td>21397 (100%)</td>
<td>1678 (6.7%)</td>
</tr>
</tbody>
</table>

Table 1

Number of births conducted with doctors’ or midwives’ assistance

Doctor Nikolai Russkikh, employed by the Ekaterinburg zemstvo, attained national fame for fighting child and infant mortality. While
working as a doctor in Ekaterinburg, he studied European experiences from diminishing infant and child mortality and to promote this knowledge cofounded the Society for fighting child mortality in Russia. He also edited the journal focusing on protection of maternity and infants. In general, Ekaterinburg had an elaborate network of medical facilities that surpassed most other Russian cities in both quality and quantity (Glavatskaya, Borovik, Thorvaldsen).

**Sources and methods**

Despite the existence of relevant monographs on infant mortality in the Urals and Siberia, none of them allows easy comparison. The method of calculating rates has been different from the standard international approach, since they measure infant mortality as a fraction of all deaths, including adults. Instead, we use individual data transcribed from the Ekaterinburg Ascension Church’ parish register to compute infant mortality rate, and other city’s parishes’ data is on the way.

Tsar Peter the Great introduced vital events registration in parish register (metricheskie knigi) in 1722 and priests all over the Russian Empire registered baptisms, weddings and burials until the October Revolution in 1917, when the Bolsheviks conveyed this competence to the secular state office. However, some priests, as was the case of the Ekaterinburg Ascension Church, kept making records even a few years later. The metricheskie knigi had three parts – about baptisms, about weddings and about burials.

We found the Ekaterinburg’s Ascension Church parish registers in the State Archive of Sverdlovsk oblast’ (GASO) for the period of 1880–1919 relatively well preserved: only records made in 1904 are missing. We then transcribed them into the ‘UraPP’ (Ural Population Project) database, which so far includes 7786 baptism records and 8281 death records, among them the data on 3287 infants (See Table 2).

<table>
<thead>
<tr>
<th>Number of records analyzed</th>
</tr>
</thead>
<tbody>
<tr>
<td>Male</td>
</tr>
<tr>
<td>Female</td>
</tr>
</tbody>
</table>
The entries on burials in the parish registers provide names, death date, the age of the deceased and death cause. In the case of children aged under 16, there is also information about the parents: their names, social status/occupation, place of origin and marital status. In addition, death records contain data on priests conducted a funeral service and occasionally death certificate extracts verified by a doctor or police officer. Records on baptisms provide information on dates and places of birth and baptism, parents’ names, their social status/occupation, place of origin and marital status and the same information on godparents. In addition, it included names of priests conducted a baptism service and occasionally godparents’ signatures. In our previous research on ethno-religious minorities we checked the accuracy of the information on infants’ age as registered by the priests during the burial ceremony. We identified buried infants who were not entered in the baptism records, as can be expected in a rapidly growing city with heavy in-migration. With functioning railway transport this means that even a baby which died a month old could have been baptized on the way to Ekaterinburg. According to our controls, the infants’ ages reported by the city’s Rabbis, Old Believers’ leaders and Catholic priests were correct for 75% of the Jews, 86% of the Old Believers and 70% for the Catholics. The rest were correct to the nearest month, however, the corrections did not affect the results (Glavatskaya, Borovik, Thorvaldsen). Since record linkage of the Ascension Church parish register has not yet completed, we have not managed checking of how accurate the priests reported the infants’ ages. However, most Russian scholars agree that the quality of information provided by the late 19th – early 20th centuries Metricheskie knigi is satisfactory (Mironov, p. 101).

The registration of the stillborn varies from country to country (Klüsener et al., 2014). In the case of Ekaterinburg’s Catholic and Jewish parish registers we have found neither stillborn nor babies who died after a few hours or during the first day, which could have been interpreted as stillborn. Burial service was to be performed only for the baptized, which is why in case a baby died before baptism there was no records left in the church book. At this stage we do not take them into consideration.

**Infant mortality in the Urals: main findings**

As expected, the Ural IMR was significantly higher than the national figures both in rural and urban areas. Infant mortality in 1889–1917 Ekaterinburg with its 360‰, corresponded more to the big Russian city model, even though its population was smaller than that of
a conventional ‘big city’ category with a population over 100,000 at that time. We have, however, to take into consideration the industrial specialization of Ekaterinburg: the city encompassed a large number of enterprises and workshops, which was detrimental to the environment and the sanitary conditions of the city. In this respect, Ekaterinburg was closer to the level of pre-revolutionary Russian ‘megapolises’. This explains why in the late nineteenth century mortality among the Ascension Church parish’ newborn never went below the level that by half exceeded the average IMR in other cities of the Empire.

In distinction to most European countries, Urals’ urban infant mortality was lower than rural. Ekaterinburg IMR was also lower than IMR in the rural area around it, although the difference between the urban and rural IMR in the Urals was smaller compared to the rest of European Russia (see Figure 5).

These can be explained by the Ascension Church parish mixed – urban/rural character. Apart from the north-eastern part of the city, it also included parish members from the nearby villages Pyshma and Vladimirskoye, therefore, in 1900, almost every fourth parishioner (24%) was a peasant from one of these two villages. The general IMR decline in the late 19th – early 20th century Urals started with postneonatal infants (Golikova, 2012, p. 137–138). Babies in the neonatal pe-

---

3 For the Russian Empire urban/rural areas we used data on cities and uezds and for the Urals – Ekaterinburg uезд as a sample of rural case and Ascension Church parish – as a case of urban population.

---

Figure 5. Urban and rural IMR in the Russian Empire and the Urals3, 1890-1894, ‰

Source: (Pokrovsky, Richter, p. 97–98, 100–101; Population dynamics in Perm guberniya) UraPP.
period were exposed to the highest risk of dying; infants who managed to get through this period, given adequate treatment, stood much better chances of survival. Monthly analysis of the IMR has shown a decrease in postneonatal mortality in Ekaterinburg’s Ascension Church parish from 295‰ in 1891–1900 to 259‰ in 1901–1910. (see Fig. 6).

Figure 6. IMR in the Ekaterinburg’s Ascension Church parish according to the infant’s age in months, ‰. Source: UraPP.

The study of the IMR dynamics in Ekaterinburg Ascension Church parish in comparison with the synchronous indicators of Perm’ guberniya and Ekaterinburg uezd provide the main trends in the infant mortality rates decline (see Table 3).

<table>
<thead>
<tr>
<th></th>
<th>1891–1899, ‰,</th>
<th>1901–1910, ‰,</th>
<th>1911–1916</th>
</tr>
</thead>
<tbody>
<tr>
<td>Perm’ province</td>
<td>417</td>
<td>327</td>
<td>n/a</td>
</tr>
<tr>
<td>Ekaterinburg sub-province (rural)</td>
<td>405</td>
<td>305</td>
<td>n/a</td>
</tr>
<tr>
<td>Ekaterinburg (urban)</td>
<td>389 (N = 1024)</td>
<td>353**(N = 765)</td>
<td>347(N = 520)</td>
</tr>
</tbody>
</table>

*Ascension Church parish case
**without 1904 data
Source: (Potasnitel’nyi tekst, p. 7; UraPP).
Chapter 11

The table shows the more advantageous starting position of Ekaterinburg in terms of IMR compared the whole Perm’ province and rural sub-province subordinated to Ekaterinburg in the late 19th century. However, when both the province’s and sub-province’s IMR dropped down by more than 100‰ in just two decades, IMR in Ekaterinburg itself decreased only 36‰ on average. To a certain extent, this could be blamed on the anomalously high IMR of 1909 – 479‰ (see figure 7), which has not been explained yet. However, IMR in the three pre-war years (1911–1913), was 369‰ – equal to the average in the 1890s and 1900s. We unfortunately cannot compare IMR dynamics after 1910, due to the lack of the necessary data on the Perm’ province and Ekaterinburg sub-province rural area.

Reasons for high infant mortality rates in Russia and in the Urals

The majority of doctors and statisticians at the time explained the disastrous IMR in Russia by mass poverty, engagement of peasant mothers in field work, low level of hygiene, the lack of access to medical services and the lack of solid knowledge about infant care. The latter was particularly significant since the early weaning practice widely spread among Orthodox peasants, was harmful to babies’ immune systems and digestion and resulted in the low infant survival rate (Sokolov, Grebenschikov, 37–43; Novoselsky, 1911, 17).

The situation in the Urals, compared with European Russia, was exacerbated by the geography: the Perm’ province’s vast territory, comparable in size to a European country was far from densely populated. While the average population density in European Russia was approximately 28 people per sq.km, in Perm’ province was less than 13 people per sq.km (Statistical Yearbook, p.44, 49). That affected people’s access to doctors’ and health care, which diminished infants’ chances to survive. Furthermore, while inhabitants of European territories for decades were being enlightened by medical professionals and thus improved their infant care practices, the peasant in-migrants to the Urals and Ekaterinburg in particular brought their archaic folk pediatrics with them. Sufficient evidence comes from accounts on infant care among Ural peasantry recorded by contemporaries in the late 19th century (Golikova, 2005, 99–104).

Nikolai Gundobin hypothesized that the Urals’ high IMR was aggravated by the significant number of Old Believers in the region. Ac-
According to his hypothesis, in their daily life, Old Believers strictly adhered to numerous anti-sanitary religious practices, which were harmful to the health of both the mother and the child. Their pregnant women observed long fasts throughout the year; they often performed baptism ceremonies in cold chapels; their rituals included kissing deceased persons, who, for instance, possibly died from a highly infectious disease, etc. (Gundobin, p. 10). However, our analysis of the two Old Believer parish registers in Ekaterinburg did not confirm this hypothesis: their IMR was considerably lower than in the neighboring Orthodox parish (Glavatskaya, Borovik and Thorvaldsen). However, in this case we might need to make allowance for the fact that the Old Believer community was quite diverse: Gundobin wrote about the Old Believers in general, the majority of whom belonged to peasants or factory workers, while Ekaterinburg Old Believers were the urban dwellers.

**Conclusion**

The results of our study have confirmed our hypothesis that the level of infant mortality is closely connected with the type of settlement. It became evident when we compare the corresponding data of the uezd and the city parish. In the late nineteenth century Ekaterinburg had a moderate level of infant mortality, but in 1889 to 1917 it demonstrated only a slight decrease in the post-neonatal infant mortality rate. This decrease, however, was nullified by a slight increase in neonatal mortality: 1909, 1911 and the first year of war (1914) were grievous years for infants, which affected the uezd statistics (see Figure 7).

![Figure 7. IMR dynamics in the Ascension Church parish, 1889-1916, ‰*](image)

*instead of missing 1904 data we used aggregates. Source: UraPP.
The first epidemiological transition started to develop in Russia later than in most European countries and soon after the start was interrupted by socio-political disasters (Isupov). Even if slow, this process had a steady effect, which was noticed and reported by the doctors at once (Novoselsky, 1916, p.180–187). Although Perm’ guberniya had an extremely high level of IMR, when it entered the epidemiological transition, it soon became one of the leaders in terms of declining infant mortality rates. While from 1886–1897 to 1908–1910, the IMR declined on average at a pace of 21% across the country, the Perm’ gubernia IMR dropped down by 117‰. The comparative analysis of uezd and city dynamics shows that the IMR dropped down in rural areas while in the city it remained on the same level.

We believe that this effect was due to the zemvsto doctors’ activity, who focused predominantly on promoting knowledge and medical care in rural areas. This movement was particularly influential in the Urals, which had a large number of zavody – metal producing factories with a population around, generally more open to innovations. These settlements had a relatively developed medical network, system of district doctors (Shestova, p. 38) and nurseries (Golikova, Dashkevich). Spatial analyses of the IMR in Ekaterinburg uezd supports this hypothesis. The uezds’ subdivisions with zavod settlements in their territories had lower IMR, than agricultural units (Bakharev).

The situation in Ekaterinburg itself was influenced heavily by the mass influx of peasant in-migrants and numerous industrial and craftsman enterprises, which worsened the already poor environmental and sanitary conditions. The city administration was unable to provide urban dwellers with health care, access to information and sometimes even food, which resulted, among other things, in the high level of infant mortality rates, in which Ekaterinburg was lagging more and more behind the regional standards of that time. The role of ethnic and religious affiliation was a significant factor as it determined either a high level of education (Jews, Catholics, Lutherans) or strict hygiene rules adherence (Jews, Muslims). These factors, in their turn, determined the quality of infant care and increased the baby’s chances to survive in the city. Peasants who constituted the majority of Orthodox parishes had lower level of education and were suspicious of doctors, preferring folk healing practices or simply having no time or money to obtain professional medical assistance.

Finally, the UraPP data allowed us to find out that the Ascension Church parish members on average baptized their babies within first 3 days of their life (See table 4).
Table 4

Average baptism age of infants in the Ascension Church parish, 1889–1918, days

<table>
<thead>
<tr>
<th></th>
<th>1889–1899</th>
<th>1900–1910</th>
<th>1911–1918</th>
</tr>
</thead>
<tbody>
<tr>
<td>Average baptism age</td>
<td>3.7</td>
<td>4.6</td>
<td>5.2</td>
</tr>
</tbody>
</table>

This practice was based on the Orthodox Church strong belief that babies would not end up in paradise if they died before being baptized. This belief made both the parents and the priests to hurry up with baptism, whether a baby was well or not. Given the fact that the ritual required the baby’s complete immersion into a vessel full of water three times, it is easy to believe that the whole procedure could have affected the babies’ health. It is interesting to note, that Ekaterinburg Catholics, who generally believed in the very same idea that only baptized babies access the Heavens, on average baptized their newborn at the age of 41 days.

Further development of the UraPP and record linkage will allow us to check our findings made on the Ascension Church parish data and answer other research questions.

References


Chapter 11


Novosel’skiy, S. A. (1911). *O razlichiyakh v smertnosti sel’skogo i gorodskogo naseleniya Evropeyskoy Rossii* [Differences between mortality of urban and rural population of European Russia]. Moscow.


Rashin, A. G. (1956). *Naselenie Rossii za 100 let (1811–1913 gg.)* [The population of Russia for 100 years (1811–1913)]. *Statisticheskie ocherki* [Statistical essays]. Moscow: Gosstatizdat.


Tsentr’nyy statisticheskiy komitet MVD. (1911). *Statisticheskiy ezhegodnik Rossii 1910 g. (god sed’ moy)* [Statistics yearbook. Russia in 1910 (year seventh)]. Petrograd: Tsentr’nyy stat. kom. MVD.

Chapter 11


LIST OF THE CONTRIBUTORS

Dmitrii Bakharev is research fellow at the Institute of History and Archaeology of the Ural Branch of the Russian Academy of Sciences and International Demographic Unit of the Ural Federal University, Ekaterinburg, Russia
dmitry.bakharev@urfu.ru

Alexander Bobitsky is research fellow at the Institute of History and Archaeology of the Ural Branch of the Russian Academy of Sciences and International Demographic Unit of the Ural Federal University, Ekaterinburg, Russia
bobitskiy.alexander.101@yandex.ru

Gabriel Brea-Martínez is postdoctoral fellow at the Department of Economic History and at the Centre for Economic Demography of Lund University, Lund, Sweden
gabriel.brea_martinez@ekh.lu.se

Marius Eppel is lecturer, Babeș-Bolyai University, Department of Modern History, Archival Science and Ethnology, Cluj-Napoca, Romania and Research Professor, École des Hautes Études en Sciences Sociales, Centre de Recherches Historiques, Paris, France
mariuseppel@yahoo.fr

Georg Fertig is professor at the Martin Luther University, Halle-Wittenberg, Germany
georg.fertig@geschichte.uni-halle.de

Alicia Fornés is senior research fellow at the Computer Vision Center at the Autonomous University of Barcelona, Spain
afornes@cvc.uab.es

Elena Glavatskaya is professor at the Department of History, Ural Federal University and also a senior research fellow at the Institute of History and Archaeology of the Ural Branch of the Russian Academy of Sciences, Ekaterinburg, Russia
elena.glavatskaya@urfu.ru
Mark Gortfelder is junior researcher at the Estonian Institute for Population Studies, Tallinn University, Tallinn, Estonia
mark.gortfelder@tlu.ee

Siegfried Gruber is researcher at the Department of History, University of Graz, Austria
si.gruber@uni-graz.at

Josep Lladós is associate professor at the Computer Science Department, and director of the Computer Vision Center, at the Autonomous University of Barcelona, Barcelona, Spain
josep@cvc.uab.es

Maria A. Markova is associate professor, St. Petersburg State University, St. Petersburg, Russia m.markova@spbu.ru

Péter Őri is the deputy director of the Hungarian Demographic Research Institute, Budapest, Hungary
ori@demografia.hu

Joana Maria Pujadas-Mora is researcher and co-head of the Historical Demography area at the Center for Demographic Studies at the Autonomous University of Barcelona, and adjunct professor at the Department of Economic History at the University of Barcelona, Barcelona, Spain
jpujades@ced.uab.es

Hilde L. Sommerseth is professor of historical demography and director of the Norwegian Historical Data Centre at UiT – The Arctic University of Norway, Tromso, Norway
hilde.sommerseth@uit.no

Oana Sorescu-Iudean is graduate student at the Graduate School of East and South-East European Studies, University of Regensburg & Babeș-Bolyai University of Cluj-Napoca, and Research Assistant at the Babeș-Bolyai University of Cluj-Napoca, Romania
oana.soarescu@gmail.com
НАУЧНОЕ ИЗДАНИЕ

НОМИНАТИВНЫЕ ДАННЫЕ
В ДЕМОГРАФИЧЕСКИХ ИССЛЕДОВАНИЯХ.
ВОСТОК И ЗАПАД

Монография

На английском языке

Подписано в печать 20.06.2019. Формат 60×84 1/16.
Усл. печ. л. 12,79. Гарнитура Times.
Бумага офсетная. Тираж 100 экз. Заказ № 214.

Издательство Уральского университета
620000, Екатеринбург-83, ул. Тургенева, 4

Отпечатано в Издательско-полиграфическом центре УрФУ
620000, Екатеринбург-83, ул. Тургенева, 4
Факс: +7 (343) 358-93-06
E-mail: press-urfu@mail.ru
http://print.urfu.ru
Scientific publication

Nominative Data in Demographic Research in the East and the West

Monograph

Edited by:
Elena Glavatskaya, Gunnar Thorvaldsen, Georg Fertig, Mikolaj Szoltysk

Ural University Press
Turgeneva St. 4. Ekaterinburg, 620000, Russia
E-mail: press-urfu@mail.ru
http://print.urfu.ru
### Родившееся

<table>
<thead>
<tr>
<th>Место и число рождения</th>
<th>Место и число крещения</th>
<th>Место и число креации</th>
</tr>
</thead>
<tbody>
<tr>
<td>Весьма крепко родился</td>
<td>Весьма крепко внесённый</td>
<td>Весьма крепко внесённый</td>
</tr>
<tr>
<td>в деревне Рождествено</td>
<td>в деревне Победище</td>
<td>в деревне Победище</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>