

Analysis of individual data from historical parish registers for the study of demographic behavior in the past – case study on the population born from 1650s to 1830s in Jablonec, Czech lands

Ludmila Fialová¹, Klára Hulíková², Barbora Kuprová³

Historical data of population of the Czech lands have been studied only through the traditional methods of historical demography (e.g. family reconstitution method according to Henry, 1980) to our study (e.g. Fialová, 1994, 1998, 2004). However, these data (individual records from parish registers) offer many possibilities of application of other methods, which are commonly used in demographic analysis of contemporary data (e.g. the survival analysis, the Cox regression). The individual historical data may be even more suitable for these methods, using the individual life histories it is possible to analyze the way of life of person. It is often not possible for the contemporary data in many developed countries (incl. the Czech lands) because of the protection of personal data.

The aim of the paper is to describe the reproductive behavior of people born from the beginning of the 18th century to 1830s and to show possibilities of usage of the survival analysis and the Cox regression in historical demography. Moreover, it will be shown, how the problem of incomplete records could be solved. Data used in the study are individual records from Jablonec, city in the northern part of the Czech lands, which are excerpt from the parish registers.

Jablonec was a village typical for the manufacture of glass products in the second half of the 18th century. It was promoted to the township on the 21th April 1808. The title meant for Jablonec that it received the right to hold markets. It meant that Jablonec received the place where its products could be sold. Manufacture of textile products developed in Jablonec in the first half of the 19th century. So Jablonec was one of the industrial townships in that time. Jablonec had over the 3 thousand people in 1830.

The studied time period was very special thanks to the social and economic changes in the Czech lands. We can date a number of important changes in different areas in the reporting period (significant shift in health and social services, the patent of tolerance etc.), some of them probably had a significant impact on the demographic and social behavior of the population of the Czech lands. The historical data could be used for illustration of the influence of all these historical facts to the demographic behavior during the demographic transition.

For the analytical part of our study we set several research questions related to studied period and demographic processes. They could be briefly summarized as follows:

1. How the ongoing process of demographic transition could be characterized through the basic demographic processes (mortality, nuptiality, fertility) and their timing?
2. Could be the founded changes in intensities as well as timing of demographic processes directly related to or explained by the historical political acts or social and economic changes?
3. What are the most important factors influencing the intensity and timing of the studied demographic processes during the studied time period. Which are the most significant ones?

¹ Charles University in Prague, Faculty of Science, Department of Demography and Geodemography (ludmila.fialova@natur.cuni.cz)

² Charles University in Prague, Faculty of Science, Department of Demography and Geodemography (klara.hulikova@natur.cuni.cz)

³ Charles University in Prague, Faculty of Science, Department of Demography and Geodemography (bara.kuprova@gmail.com)

In accordance to the aim of the paper and the stated questions, demographic processes and their intensity and timing are studied separately for groups of cohorts and changes can be described not only from period point of view (comparison in time) but also from the cohort one (comparison of cohorts and differences among them). We focused on birth cohorts as well as on marriage cohorts.

For the analysis the dataset of individual records from parish registers was used. Database is based on the date of wedding. Subsequently we searched date of birth and death of the engaged couple and their children if it was possible. We included only records about first marriages to the analysis and due to the sparse records on the beginning and the end of studied time period we limited the births cohort by years 1700 and 1799. Selected cohorts are included in the Table 1.

Table 1: Cohorts (birth and marriage cohorts) included to the analysis

	Cohorts included to the analytical part	Total number of records
Births – Males	1700–1799	1064
Births – Females	1700–1799	963
Marriages	1650–1830	1350

In the paper, the main analytical method is the survival analysis, which enables to find the answers for some of the above stated research questions. Survival analysis is an analytical tool contemporary used and very popular everywhere, where the time durations are the subject of study, especially in epidemiology, medicine, biology, demography, sociology, econometrics, etc. (Aalen et al., 2010). The respond variable of interest is the time until a studied event occurred (SAS Institute Inc., 2009). This method was developed for longitudinal data (Allison, 2010) what fully corresponds with the historical data set. One of the most important advantages of the survival analysis is the possibility to work also with incomplete data. That means, in the survival analysis we are able to cope with the problem of censoring (Aalen et al., 2010).

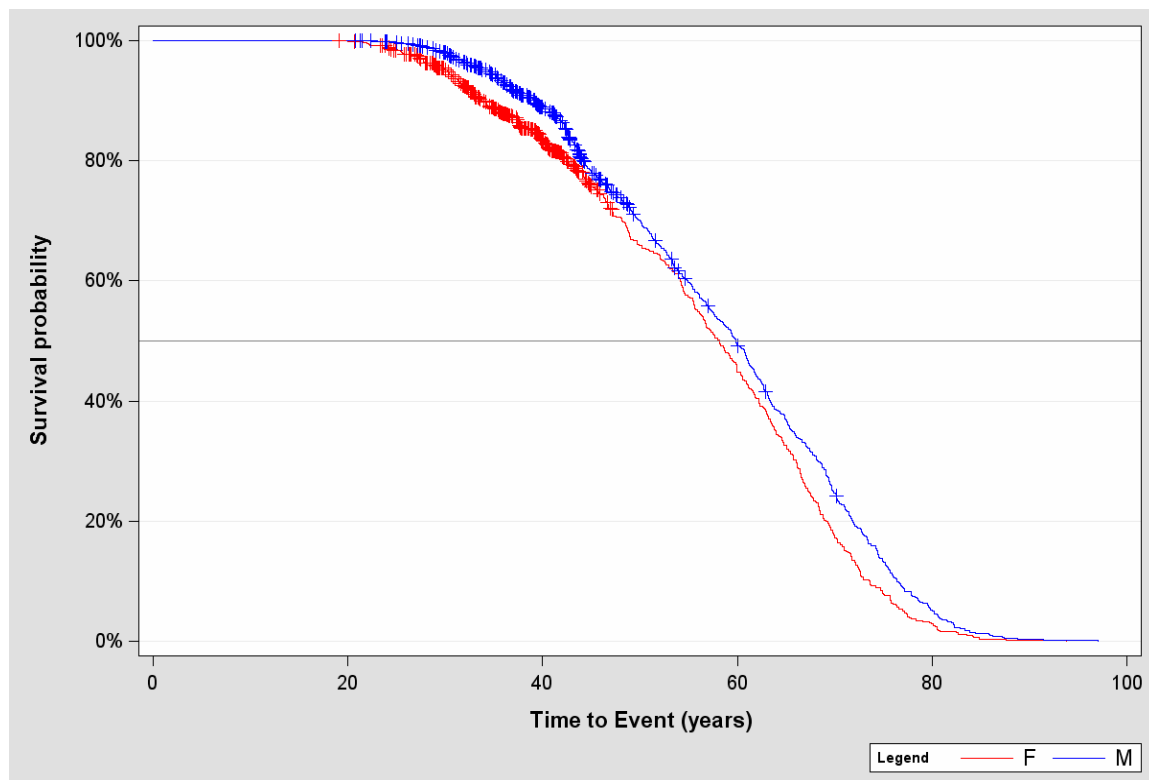
For the survival analysis it was possible to use the exact lengths of life or time duration from initial event (e.g. birth, marriage) to the studied one (e.g. marriage, birth of child). This length was calculated in days as the difference of the date of the initial event and the studied one.

The simplest survival analysis which was done dealt with the time durations from birth to death of the analyzed persons (in Figure 1 represented by the survival curves). It is clear, that the survival probability was higher for males than for females at almost all ages. The reasons could be found in higher mortality of adult women associated with the motherhood – the maternal mortality. The survival probability approached for both sexes at the age of about 50 years, after the end of the reproductive period of women. Censoring in this type of analysis was present in cases where the date of death was not available – in such cases it was necessary to find any other event (date of the event) in individual life histories, when the studied person had to be alive. In our studied data set the last information about the person (the last date where the person was still alive) was the date of birth of the last child. This date was used instead of the date of death in the cases of consored observations. That means, that we do not know the exact time duration, but to the analysis we can involve the information, that the time duration cannot be shorter than the time from the initial event to the birth of the last child.

The reproductive behavior which was analyzed for example by using the time duration from birth to the marriage was done for various birth cohorts so as the differences among cohorts could be traced (Figure 2). The family behavior was studied through the time durations form marriage to the birth of

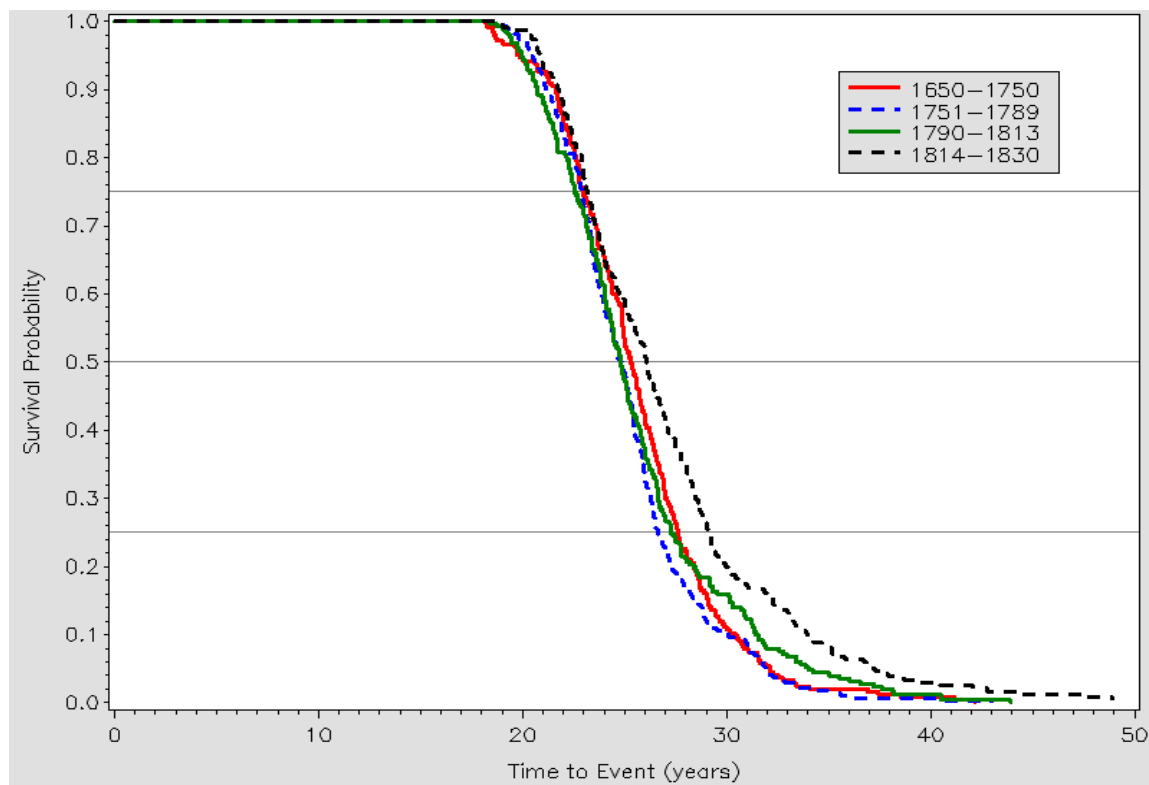
the first child. All the analyses could be done according to birth cohorts of parents or according to marriage cohorts (groups of people who entered to marriage in similar time period).

Figure 1: Survival probability from birth to death, males and females who got married



Note: *M = males; F = females; Output from the SAS software*

Figure 2: Survival probability from birth to marriage, males (according to marriage cohorts)



Note: *In the legend, there are marriage cohorts – i.e. the analysis was done separately for groups of marriages according to years of the marriage; Output from the SAS software*

Only persons who married at least once entered into the analysis of nuptiality. The constant values at the beginning of the curve for both men (Figure 2) and women correspond to the probability of the studied population to survive to the marriage, which is equal to one (as stated above, we analyzed only persons who entered to marriage at least once). Change in the timing of marriage was more visible for men (Figure 2) than for women. This change was probably related with the regulation from 1815 which imposed to submit a certificate of the earnings and place of residence when applying for permission to get married. So the possibility of marriage was more complicated for the poorer stratum of the population. However, this could be taken only as one of more possible hypotheses, and for its evaluation and correct description of the reality it is necessary to know the detailed history of Jablonec.

The Cox regression enables to study the processes of demographic reproduction more in detail and to quantify the effect of several explaining factors. The Cox regression is a semi-parametric function, where the base comes from the simple survival analysis (see above). Through this method we can estimate the hazard function for a particular individual (group of individuals with the same studied characteristics) relative to a baseline hazard (reference level) at a particular point in time. In our study the Cox regression was used for detailed analysis of the mortality process – time duration from birth to death or from marriage to death (study and quantification of the effect of sex, age, marriage cohort, age-difference of bride and groom, number of children in case of females) and for the analysis of nuptiality and fertility process – time duration from birth to marriage or to the birth of the first child, time duration from marriage to the first child (study and quantification of the effect of sex, age, marriage cohort, age-difference of bride and groom).

All the performed analyzes not only helped to answer the set research questions but also confirmed the usefulness of contemporary analytical approaches also in the field of historical demography. Thanks to that it is possible to learn more about the life in the past and gain the most from the unique data source as (individual records from) parish registers are.

References

- AALLEN, O.O. – BORGAN, O. – GJESSING, H.K. 2010. *Survival and Event History Analysis : A Process Point of View*. 2010. Springer Science+Business Media, LLC. ISBN 978-1-4419-1909-0.
- ALLISON, P.D. 2010. *Survival Analysis Using SAS ®: A Practical Guide*, Second Edition. Copyright © 2010, SAS Institute Inc., Cary, North Carolina, USA.
- SAS Institute Inc. 2009. *Survival Analysis Using the Proportional Hazards Model Course notes*. 2009 SAS Institute Inc. Cary, North Carolina, USA.
- FIALOVÁ, L. 1994. Changes of Nuptiality in Czech Lands and Slovakia. *Journal of Family History*, 1994, 19, č. 2, s. 107-115.
- FIALOVÁ L. 1998. On the Population development through population change in the Chech Lands in the 18th century. *Second European Social Science History Conference*. Amsterdam, 5–7 March 1998.
- HENRY, L. *Technique d'analyse en démographie historique*. Paris: INED, 1980. 177 s.