

Sequence Analysis as a Tool for Family Demography

Nicola Barban, Department of Sociology and Nuffield College, University of Oxford

nicola.barban@nuffield.ox.ac.uk

Maria Sironi, Department of Social Science, University College London

m.sironi@ucl.ac.uk

Introduction

In this work we will describe the development of sequences analysis (SA) techniques to investigate the process of family formation and dissolution. Family structure has changed substantially in the past decades, and family trajectories are more heterogeneous than they used to be in the past. The age at leaving home and at first marriage have increased in many, if not all, western societies; Cohabitation has become a very standard stage in people's family formation; Divorce rates rose considerably since the 1970s, but their growth slowed down and even halted in some countries; Childbearing has been postponed and stepfamilies have become more and more common (Amato et al., 2008; Baizan, Aassve, & Billari, 2003; F C Billari, Philipov, & Baizan, 2001; Iacovou, 2010; Schoen, Landale, & Daniels, 2007; Stevenson & Wolfers, 2007). All these trends imply not only changes in timing of events, but also changes in the sequencing and the duration of events.

Many aspects of family trajectories have been analysed individually, without taking into account the interrelation among different events. However, it is necessary to look at the process of union formation and the subsequent family pathways from a holistic point of view. Sequence analysis is therefore the appropriate tool to analyse family histories, taking into account timing, sequencing, and duration of events. In this paper we will discuss the way in which sequence analysis has been used so far in family demography, and we will illustrate the most relevant developments and innovative procedures relative to this technique (e.g., multichannel sequence analysis and dyad analysis). Moreover, family change has been very different in different contexts – with varying degrees of postponement and de-standardization in life course trajectories. Hence, in the second part we will use data from the European Social Survey to illustrate an empirical application of sequence analysis, and we will describe family trajectories across European countries.

Family Change Over Time

This section will summarize the most relevant changes in the process of family formation and dissolution, and therefore show the need for innovative ways to analyse family trajectories.

A Formal Representation for Life Course Trajectories

Life course trajectories can be described as the observation, over the course of an individual's age, of a number of events representing specific family roles. The concept of trajectory derives from the interdisciplinary systematisation of the life course paradigm proposed by Elder (1985), in which life course trajectories usually refer to the joint occurrence of events in multiple life domains. For example, one may want to have a representation of the evolution of family formation, childbearing and work history. Trajectories can be analysed by representing the original data (i.e. each individual's life course) as a sequence of states. Life course analysis looks at the entire development of family history, i.e., the whole trajectory, taking into account the interrelation among events. Under this perspective, characteristics such as type, number, and duration of unions, or the order of events may have an effect on later outcomes in individuals' life course (Peters & Liefbroer, 1997).

Sequence analysis for social sciences (Abbot & Tsay, 2000; Abbott, 1995; Aisenbrey & Fasang, 2010) is a set of techniques that focus on the analysis of entire trajectories rather than the occurrence of single events. In the sequence analysis, each life course trajectory is represented by a string of characters, which resembles the one used to code DNA molecules in biological sciences. Hence, every trajectory is made up of a number of values that correspond to the number of years (or months) each individual is observed. Trajectories can be analysed by representing the data as a categorical time series. Each individual i can be associated to a variable s_{it} indicating the individual's life course status at time t . In other terms, trajectories can be represented as strings or sequences of characters, with each character denoting one particular state that denotes a particular family role.

Identifying Typical Trajectories

Sequence analysis is associated to a family of algorithms used to quantify dissimilarities between life course trajectories. Optimal Matching algorithm (OM) is the most known technique that has been applied to social science. The development of OM started in the seventies and the technique has been described in details by Kruskal (1983). Basically, OM expresses distances between sequences in terms of the minimal amount of effort, measured in terms of edit operations (insertion; deletion and substitution), that is required to change two sequences such that they become identical. Abbott

(1995) adapted OM to social sciences assigning to three elementary operations different costs, based on the ‘social’ differences between states (Lesnard, 2006). The choice of the operations’ costs determines the matching procedure and influences the results obtained. Sequence analysis algorithms identify differences in trajectories due to changes in *timing* (when events happen), *quantum* (what and how many transitions), and *ordering* (in what order) of life course events (Francesco C. Billari, Fürnkranz, & Prskawetz, 2006; Francesco C. Billari & Piccarreta, 2005).

Sequence analysis and OM are often used in conjunction with cluster analysis to identify patterns in the data and highlight typical life course trajectories (Abbot & Tsay, 2000; Aisenbrey & Fasang, 2010; Barban, 2013; Barban & Billari, 2012).

Uses of Sequence Analysis in Family Demography Literature

This section will give an overview of how SA has been used so far in the demography literature, and it will also highlight the most recent and innovative developments of this technique (e.g. multichannel sequence analysis and dyad analysis)

Comparative Research Using Sequence Analysis

Given the fact that family structure and life course trajectories have changed differently in different countries, it is important to take a cross-country comparative approach. This section discusses the use of sequence analysis as a tool for comparative research in family demography. Sequence analysis can be used to identify country differences in family trajectories in a holistic way. Rather than analysing the median age at different life course events (e.g. age at marriage and children progression), sequence analysis can be employed to describe the occurrence and the timing of events together with the sequencing (i.e. the order) of different family events.

We apply this analytical framework to investigate the relevance of different theories of convergence/divergence of family behavior (Billari & Wilson, 2001) in different European countries, and study how family trajectories differ across these countries.

Data

To illustrate the use of comparative sequence analysis for family demography, we use the European Social Survey (ESS). The ESS is a multi-country survey that started in 2002, and that counts seven rounds, one every two years (from 2002 to 2014). Its main aim is to outline the attitudes of the different regions towards religion, politics, and moral issues, while also depicting their social habits

and how they are changing over time. We focus in particular on Round 3 of the ESS collected in 2006, which includes the module ‘timing of life’. This module, on top of the core questions included in each round, collects detailed information on various family-related events such as: age at leaving home; age at first non-marital cohabitation; age at first (successive) marriage; age at first (successive) child; age at marital dissolution; age at grandparenthood. The third round includes 23 countries (Austria, Belgium, Bulgaria, Switzerland, Cyprus, Germany, Denmark, Estonia, Spain, Finland, France, Great Britain, Hungary, Ireland, Netherlands, Norway, Poland, Portugal, Russia, Sweden, Slovenia, Slovakia, Ukraine) and 43,000 individuals born between 1905 and 1992.

Methods

- We identify typical trajectories for men and women based on these family-related event: age at leaving home; age at first non-marital cohabitation; age at first (successive) marriage; age at first (successive) child; age at marital dissolution; age at grandparenthood. We use sequence analysis based on Optimal Matching to identify groups of family trajectories.
- We explore cluster distribution by gender, countries and birth cohorts in order to divide family trajectories into typologies.
- We demonstrate the use of sequence analysis for the analysis of cohort dynamics of family formation and dissolution in different countries.

References

- Abbot, A., & Tsay, A. (2000). Sequence Analysis and Optimal Matching Methods in Sociology: Review and Prospect. *Sociological Methods & Research*, 29(1), 3–33.
<http://doi.org/10.1177/0049124100029001001>
- Abbott, A. (1995). Sequence Analysis: New Methods for Old Ideas. *Annual Review of Sociology*, 21, 93–113. <http://doi.org/10.1146/annurev.soc.21.1.93>
- Aisenbrey, S., & Fasang, A. F. (2010). New Life for Old Ideas: The “Second Wave” of Sequence Analysis Bringing the “Course” Back Into the Life Course. *Sociological Methods & Research*, 38(3), 420–462. <http://doi.org/10.1177/0049124109357532>
- Amato, P. R., Landale, N. S., Havasevich-Brooks, T. C., Booth, A., Eggebeen, D. J., Schoen, R., & McHale, S. M. (2008). Precursors of young women’s family formation pathways. *Journal of Marriage and Family*, 70(5), 1271–1286. <http://doi.org/10.1111/j.1741-3737.2008.00565.x>
- Baizan, P., Aassve, A., & Billari, F. C. (2003). Cohabitation, marriage, and first birth: The interrelationship of family formation events in Spain. *European Journal of Population/Revue Européenne de Démographie*, 19(2), 147–169. Retrieved from [file:///Users/Maria/Papers/2003/Baizan/European Journal of Population/Revue européenne de Démographie 2003 Baizan.pdf](file:///Users/Maria/Papers/2003/Baizan/European%20Journal%20of%20Population/Revue%20européenne%20de%20Démographie%202003%20Baizan.pdf)
- Barban, N. (2013). Family Trajectories and Health: A Life Course Perspective. *European Journal of Population / Revue Européenne de Démographie*, 29(4), 357–385. <http://doi.org/10.1007/s10680-013-9296-3>
- Barban, N., & Billari, F. C. (2012). Classifying life course trajectories: A comparison of latent class and sequence analysis. *Journal of the Royal Statistical Society. Series C: Applied Statistics*, 61(5), 765–784. <http://doi.org/10.1111/j.1467-9876.2012.01047.x>
- Billari, F. C., Fürnkranz, J., & Prskawetz, A. (2006). Timing, sequencing, and quantum of life course events: A machine learning approach. *European Journal of Population*, 22(1), 37–65.
<http://doi.org/10.1007/s10680-005-5549-0>
- Billari, F. C., Philipov, D., & Baizan, P. (2001). Leaving home in Europe: The experience of cohorts born around 1960. *International Journal of Population Geography*, 7(5), 339–356. Retrieved from <papers://e09fda77-1450-4449-8ecf-5a9bb72f5b0a/Paper/p493>
- Billari, F. C., & Piccarreta, R. (2005). Analyzing Demographic Life Courses through Sequence Analysis. *Mathematical Population Studies*, 12(2), 81–106.
<http://doi.org/10.1080/08898480590932287>

- Billari, F. C., & Wilson, C. (2001). Convergence towards diversity? Cohort dynamics in the transition to adulthood in contemporary Western Europe. *Max Planck Institute for Demographic Research, Working Paper WP2001-039*. Retrieved from papers://e09fda77-1450-4449-8ecf-5a9bb72f5b0a/Paper/p887
- Elder, G. H. (1985). Perspectives on the life course. *Life Course Dynamics: Trajectories and Transitions, 1*(2007), 23–49. <http://doi.org/10.1111/j.1751-9020.2007.00004.x>
- Iacovou, M. (2010). Leaving home: Independence, togetherness and income. *Advances in Life Course Research, 15*(4), 147–160. <http://doi.org/10.1016/j.alcr.2010.10.004>
- Kruskal, J. B. (1983). An Overview of Sequence Comparison: Time Warps, String Edits, and Macromolecules. *SIAM Review, 25*(2), 201–237. <http://doi.org/10.1137/1025045>
- Lesnard, L. (2006). Optimal Matching and Social Sciences. *Centre de Recherche En Économie et Statistique Working Paper Series, 2006*(1). Retrieved from <http://www.crest.fr/doctravail/document/2006-01.pdf>
- Peters, A., & Liefbroer, A. C. (1997). Beyond Marital Status: Partner History and Well-Being in Old Age Well-Being in Old Age. *Journal of Marriage and Family, 59*(3), 687–699. <http://doi.org/10.2307/353954>
- Schoen, R., Landale, N., & Daniels, K. (2007). Family transitions in young adulthood. *Demography, 44*(4), 807–820. <http://doi.org/10.1353/dem.2007.0044>
- Stevenson, B., & Wolfers, J. (2007). Marriage and divorce: Changes and their driving forces. *Journal of Economic Perspectives, 21*(2), 27–52. Retrieved from papers://e09fda77-1450-4449-8ecf-5a9bb72f5b0a/Paper/p628