

institut
für medizinische
soziologie

Lifehistory data - computer workshop

Morten Wahrendorf / Christian Deindl

17th October 2018

Centre for Health and Society,
Institute for Medical Sociology, University Düsseldorf

Harmonized Life-history data

- Five "histories" (from 20 to 70):

- | | |
|---------------------------|---|
| (1) Children: | childnstate* / caren* / youngn*
total plus children-age specific |
| (2) Partnership: | partnerstate* |
| (3) Accommodation: | residstate* |
| (4) Employment: | workstate* |
| (5) Health: | healthstate* |
-

Lab session / Exercise I.

Basics and first sequence exploration

(1) Open Stata

(2) Get Gateway data and do-file from the following source:

`net from http://www.lifecourse.de/sequences/`

(3) Open data and explore different state variables (e.g. `des workstate*`)

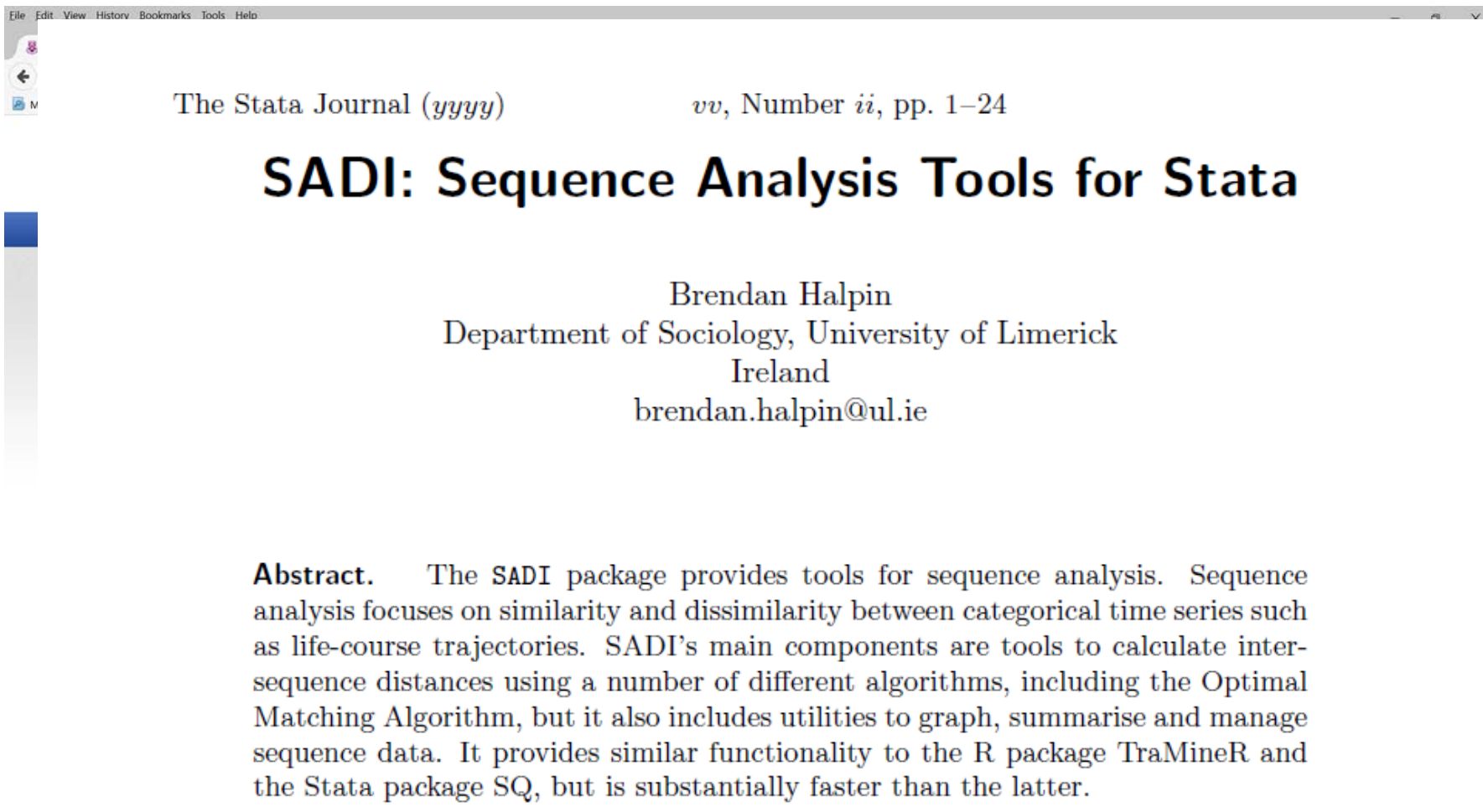
(4) How does the employment situation at age 40 vary by gender?

(5) Does the proportion of men/women in full-time employment vary by number of children at the age of 40?

(6) Does the proportion of men/women in full-time employment vary by partnership situation at the age of 40?

(7) How does the employment situation for women at age 40 vary between countries?

The SADI-Package.



The screenshot shows a web browser window with a menu bar (File, Edit, View, History, Bookmarks, Tools, Help) and a toolbar with a back button and a search icon. The page content is centered and includes the journal information, title, author, and abstract.

The Stata Journal (*yyyy*) *vv*, Number *ii*, pp. 1–24

SADI: Sequence Analysis Tools for Stata

Brendan Halpin
Department of Sociology, University of Limerick
Ireland
brendan.halpin@ul.ie

Abstract. The SADI package provides tools for sequence analysis. Sequence analysis focuses on similarity and dissimilarity between categorical time series such as life-course trajectories. SADI's main components are tools to calculate inter-sequence distances using a number of different algorithms, including the Optimal Matching Algorithm, but it also includes utilities to graph, summarise and manage sequence data. It provides similar functionality to the R package TraMineR and the Stata package SQ, but is substantially faster than the latter.

Lab session / Exercise II.

Summary measures

- (1) Install the package "sadi" by typing "ssc install sadi" (same for "moremata")
- (2) Look at the help of the sadi package by typing "help sadi"
- (3) Restrict the sample to respondents aged 70 or older with complete employment histories (to focus on histories from age 20 to 70).
- (4) Create a string-variable that describes employment sequence between age 20 and 70 (with each state represented by a different symbol).
- (5) Create a variable whether respondents had an episode of long-term unemployment (at least 3 years). In which country was this most prevalent?
- (6) Create variables with information on time spent in each of the employment states (in years and as a proportion of the whole sequence).
- (7) How many years did women on average spend with "home and family work" in Sweden compared with Spain and England?
- (8) Does the number of years spent in "home or family work" for women differ by cohort and by country at the age of 10?
- (9) Estimate a linear regression using years spent with "home or family work" as outcome and gender, cohort, country and country as predictors.

Lab session / Exercise III.

Visualizations (chronograms)

(1) Create a chronogram for employment histories by survey.

(2) Create a chronogram for employment histories for Spain, Sweden and England by sex.