

Computational Social Sciences using R

Ajit Kumar¹

Neeraj Hatekar²

1. Department of Applied Mathematics, Institute of Chemical Technology, Mumbai, INDIA 400 019

Email: ajit72@gmail.com

2. Department of Economics, University of Mumbai, INDIA 400 098

Email: neeraj.hatekar@gmail.com

Keywords: CSS, R

The computational social science programme (CSS) at the University of Mumbai envisages using R as a large scale data mining tool as well as for statistical modelling of previously unpublished data. The first section of the paper examines the various applications of R in the interdisciplinary CSS programme. One of the objectives of the CSS programme is to computerise large unpublished data bases that are of relevance to social scientists and develop software tools that will enable practitioners to query these data on various dimensions. In particular, Western India has a large number of unpublished data sources on demographic variables like age at marriage, fertility, birth spacing, educational attainments for women that stretch over nearly a century and a half. These have been only rarely used (Hatekar, Mathur, Rege (1997)) as they are not available in a computerised form, and since appropriate tools do not exist in order to query them. The CSS programme has been digitising these data bases and using R for statistical modelling. The current paper presents one such application on modelling the fertility transition that took place in the 1920's in upper caste households in Western India. The fertility shift is modelled and shown to be associated with larger cultural shifts in the ideas of ideal households and housewives.

In the second section we deal with various issues regarding use of R in social sciences and economics in India. We also touch upon advantages and disadvantages of using R, some of challenges and difficulties in using R and how to overcome these challenges.

References

(Hatekar, Mathur, Rege (1997): "Social Reform by Legislation: The Strange Case of the Sarda Act", Economic and Political Weekly, January 1-7, 2007.

Kleiber C. and A. Zeileis (2008): Applied Econometrics with R, Springer.

R-project home page <http://www.r-project.org/>