# The R package modslavmse for multiple systems analysis

Bernard W. Silverman*
Rights Lab
Faculty of Social Sciences
University of Nottingham

March 26, 2018

**Abstract**

The package **modslavmse** contains various routines for the application of multiple systems estimation in conjunction with other routines. This note contains details of the package and its implementation in R.

## 1   Introduction

Multiple systems estimation (MSE) is a generalisation of the mark-recapture approach to estimating the size of a hidden population. This approach is commonly credited to Petersen in the late 19th century, but it is probably considerably older; see Goudie and Goudie (2007). The classic mark-recapture approach depends on two 'catches' from the population of interest; MSE generalises this to a larger number of capture occasions or lists of subjects. MSE has been used in a number of contexts where the population is hard to count, such as counting casualties in armed conflicts (Manrique-Vallier et al., 2013) and numbers of injecting drug users (King et al., 2013). In the context of Modern Slavery, MSE was used to obtain an estimate of the numbers of victims in the United Kingdom based on the 2013 National Crime Agency Strategic Assessment (NCA, 2014). Details are given by Silverman (2014) and Bales et al. (2015). This note gives details of R programs contained in the package **modslavmse** that implement the approach set out in that work.

## 2   Installation

The package **modslavmse** is available from GitHub (Silverman, 2018). It can be loaded from within an R session as follows:

1. Install and load the package **devtools** from the CRAN package repository.

---

*Sir Bernard Silverman holds a range of part-time and honorary appointments. For details see `http://www.bernardsilverman.com`. To contact him please use `bernard.silverman` 'at' `stats.ox.ac.uk`.

2. Install **modslavmse** by the command
   `install_github("bernardsilverman/modslavmse")`.

The main routine `MSEfit` also requires the installation of the package **Rcapture** (Baillargeon and Rivest, 2007) which is available from CRAN.

# 3  Routines within the package

The main routine `MSEfit()` calls a routine from the package **Rcapture** to perform the actual MSE, and uses the approach set out in Bales et al. (2015) to choose the parameters to fit. Interactions between the various lists are added stepwise, at each point choosing the interaction which makes the best improvement in the AIC measure of model fit, but stopping if the fitted interaction parameter is not significant.

Suppose there are $k$ lists of subjects. The data input required for the estimation is a matrix $Z$ with $k + 1$ columns. Each row of $Z$ corresponds to a particular pattern of presence or absence on the various lists, with an entry of 1 in column $j$ denoting presence on list $j$, and 0 absence from that list. The $(k+1)^{\text{th}}$ column gives the count of cases with that pattern. The data used in Silverman (2014) is given in the matrix `UKdat` within the **modslavmse** package.

The output format from `MSEfit()` is the same as that of the function `closedpCI.t()` in the **Rcapture** package, and the user is referred to that package's documentation for full details.

One way that Silverman (2014) tested the robustness of the estimate was to repeat the analysis with various modifications of the original data, for example omitting some of the lists and coalescing others into a single list. These processes are facilitated by the routines `omitlists` and `mergelists`, which act on the original data matrix to produce a new data matrix with the desired list structure.

# 4  Future developments and conclusion

The package **modslavmse** also contains functions written for further methodological developments. For example, one aspect under current consideration is a Bayesian approach which, in essence, allows for averaging across possible models. This makes use of the package **MCMCpack** (Martin et al., 2011). See the **modslavmse** function `MCMCfit` for details.

Overall, it is hoped that the open distribution of R code and relevant data will make it possible for Multiple Systems Estimation to be deployed more widely in the Modern Slavery context, and for progress to be made towards standard approaches to model choice and fitting.

# References

Sophie Baillargeon and Louis-Paul Rivest. Rcapture: Loglinear models for capture-recapture in R. *Journal of Statistical Software, Articles*, 19(5):1–31, 2007. ISSN 1548-7660. doi: 10.18637/jss.v019.i05. URL `https://www.jstatsoft.org/v019/i05`.

Kevin B. Bales, Olivia Hesketh, and Bernard W. Silverman. Modern Slavery in the UK: How many victims? *Significance*, 12(3):16–21, 2015. doi: 10.1111/j.1740-9713.2015.00824.x.

I. B. J. Goudie and M. Goudie. Who captures the marks for the Petersen estimate? *Journal of the Royal Statistical Society, Series A*, 170:825–839, 2007. doi: 10.1111/j.1467-985X.2007.00479.x.

R. King, S. M. Bird, A. M. Overstall, G. Hay, and S. J. Hutchinson. Injecting drug users in Scotland, 2006: Number, demography, and opiate-related death-rates". *Addiction Research and Theory*, 21:235–246, 2013.

Daniel Manrique-Vallier, Megan E. Price, and Anita Gohdes. Multiple systems estimation techniques for estimating casualties in armed conflicts. In Taylor Seybolt, Baruch Fischhoff, and Jay Aronson, editors, *Counting Civilian Casualties. An Introduction to Recording and Estimating Nonmilitary Deaths in Conflict*, pages 77–93. Oxford University Press, New York, 2013.

Andrew D. Martin, Kevin M. Quinn, and Jong Hee Park. MCMCpack: Markov chain monte carlo in R. *Journal of Statistical Software*, 42(9):22, 2011. URL `http://www.jstatsoft.org/v42/i09/`.

NCA. *NCA Strategic Assessment: the nature and scale of human trafficking in 2013*. National Crime Agency, 2014. URL `https://tinyurl.com/mj9z5rm`.

Bernard W. Silverman. *Modern Slavery: an application of Multiple Systems Estimation*. Home Office, UK, 2014. URL `https://tinyurl.com/ydegfjaw`. Published in conjunction with the UK Government Modern Slavery Strategy.

Bernard W. Silverman. *modslavmse: Multiple Systems Estimates for estimating the prevalence of Modern Slavery*, 2018. URL `https://github.com/bernardsilverman/modslavmse`. R package, also available from `https://tinyurl.com/y6ujlwnk`.