Discussion of 'Multiple-systems analysis for the quantification of modern slavery: classical and Bayesian approaches' by Bernard Silverman

Serge Aleshin-Guendel*, Mauricio Sadinle, Jon Wakefield

University of Washington, Seattle

We thank the author for this paper, which will undoubtedly prompt methodological discussions around an important application of multiple-systems estimation (MSE). We point out issues with the proposed approach and connections with existing literature.

Uncertainty from Model Selection. The proposed "Bayesian" thresholding approach leads to an understatement of uncertainty. Inferences on the dark figure are still conditioned on the selected model, which is subject to sample variability, so the approach will almost surely have poor operating characteristics (Regal and Hook, 1991; Whitehead et al., 2019).

Connection with Spike and Slab Priors. The thresholding approach in the paper is justified as, "an approximation to a prior ... which is a mixture of an atom of probability at zero and some other distribution." Such priors are well-known as *spike and slab* priors (see e.g. Rockova et al. (2012)) and have already been employed in MSE, where they are presented as a model averaging approach that is equivalent to using spike and slab priors (King and Brooks, 2001; Overstall and King, 2014). While intuitively appealing, the spike and slab approach is not feasible in general, as it does not scale to a large number of lists, especially when considering higher-order interactions. As an alternative, we are currently working on continuous shrinkage priors (see e.g. Bhadra et al. (2019)) for exploring the space of log-linear models.

Robustness and Stability. The author relies on notions of "robustness" and "stability" to evaluate MSE approaches. These concepts refer to procedures that lead to similar estimates, regardless of the selected model. While the author acknowledges that following these guidelines suggests fitting only main effects models, his proposed approach still focuses on searching over a narrow set of models. At play here is a bias-variance trade-off: the author is favoring a potentially very biased approach in favor of lower variance. It is preferable to have a

*aleshing@uw.edu

procedure that provides us with honest assessments of uncertainty and thereby avoids misleading and overconfident results. For illustration, the dark figure is essentially known for the Kosovo data (it is around 6,001, see Manrique-Vallier (2016)). None of the 95% intervals obtained with the thresholding approach include this value, but fitting a log-linear model with all three-way interactions gives a 95% interval of $[4,922, 31,584]$, which includes the known value. This indicates that a method that searches over a larger model space, yet still encourages parsimony, could be beneficial.

Matching. Recent advances in MSE with probabilistically linked data (fuzzy matching) include Tancredi and Liseo (2011) and Sadinle (2018). Incidentally, one can't help but wonder if under-matching might be partially responsible for the sparse tables that are presented.

References

Bhadra, A., Datta, J., Polson, N. G. and Willard, B. (2019) Lasso meets horseshoe: A survey. *Statistical Science*, **34**, 405–427.

King, R. and Brooks, S. (2001) On the Bayesian analysis of population size. *Biometrika*, **88**, 317–336.

Manrique-Vallier, D. (2016) Bayesian population size estimation using Dirichlet process mixtures. *Biometrics*, **72**, 1246–1254.

Overstall, A. and King, R. (2014) conting: An R package for Bayesian analysis of complete and incomplete contingency tables. *Journal of Statistical Software*, **58**, 1–27.

Regal, R. R. and Hook, E. B. (1991) The effects of model selection on confidence intervals for the size of a closed population. *Statistics in Medicine*, **10**, 717–721.

Rockova, V., Lesaffre, E., Luime, J. and Löwenberg, B. (2012) Hierarchical Bayesian formulations for selecting variables in regression models. *Statistics in Medicine*, **31**, 1221–1237.

Sadinle, M. (2018) Bayesian propagation of record linkage uncertainty into population size estimation of human rights violations. *Annals of Applied Statistics*, **12**, 1013–1038.

Tancredi, A. and Liseo, B. (2011) A Hierarchical Bayesian Approach to Record Linkage and Size Population Problems. *Annals of Applied Statistics*, **5**, 1553–1585.

Whitehead, J., Jackson, J., Balch, A. and Francis, B. (2019) On the Unreliability of Multiple Systems Estimation for Estimating the Number of Potential Victims of Modern Slavery in the UK. *Journal of Human Trafficking*, **0**, 1–13.