(Bird, 2019). Third, list combination should consider attributes other than list size; for example, some non-governmental organizations are set up to help child victims; others, adult females. Fourth, when assessing the UK's extent of modern slavery, it matters whether the analysis concerns 'potential victims' who have passed the UK's initial 'reasonable grounds' test or 'conclusive decision victims'. As only 55% of the UK's potential victims receive a positive decision (eventually), the UK's MSE analyses to date have related to 'potential victims'. Fifth, Professor Silverman alluded to the difficulty of matching across lists. In consultation with victims, and sensitively, we should consider whether there is a role for DNA matching to reunite trafficked victims from the same family.

Finally, a public health perspective on the care for, and follow-up of, rescued victims (Bird, 2019) should address their (and others') morbidity and mortality—*en route* to the UK; during enslavement; and for at least 2 years after rescue—so that we learn how better to care.

Paul A. Smith (*University of Southampton*)
I found this paper a fascinating insight into the workings of model fitting in multiple-systems estimation. I want to make two points. First, in Fig. 5(a), I am interested in the models which give rise to the model curves with the highest likelihoods. So I examined the five models with the largest maxima in their likelihoods for the five-source case, all visible in Fig. 5(a). Four of these models have their maxima in the lower mode, and one in the upper mode (Table 18).

There is a noticeable separation in that all the models on the left-hand side of Table 18 include multiple interactions with the general public (GP) ($\equiv 5$) source, whereas the model on the right-hand side has only a single modifying interaction for this source. This suggests that relying on the main effects for GP leads to these large estimates, which fits with Professor Silverman's view that GP is an 'outlier' source. Perhaps this kind of approach can lead to an evaluation of the quality of the different sources. I can imagine less clear-cut cases where the decision on whether or not to include a source would benefit from some evidence of internal coherence, though an alternative explanation might be that a discrepant source is better because it provides new information.

This leads to my second point, which is about the quality of the input data, and more specifically the linking (distinguishing *links* derived by using a statistical process from *matches* which are the unobservable true homologies). Let me say at once that I understand the limitations of the sources and applaud the efforts to derive sensible estimates from sparse data. But, while taking great care of the model uncertainty, I find only a little comment in the paper on the uncertainty in the underlying data. Multiple-system estimation is sensitive to the linking, and, if there are missed or false links because the original sources contain rather little identifying information, this is likely to have a substantial effect on the resulting estimates. Is there any linkage metadata, or assessment of linkage quality, to go with the data sets that have helpfully been made available? Use of such information has been extended to multiple systems by Zult *et al.* (2019). I suspect (without doing the work) that adding or removing one link may have quite an effect on the chosen model with such sparse data. Perhaps the apparent discrepancy in GP as a source would be reduced in one of these scenarios, and, if indeed a coherence statistic can be constructed, we could examine what changes in linkage would most increase the coherence of the sources.

Table 18. Five models with the highest maxima in their likelihoods, in descending order of likelihood†

Lower mode	Maximum	Upper mode	Maximum
[125][135][45]	0.00279	[123][34][25]	0.00192
[125][135][245]	0.00072		
[12][25][35][45]	0.00059		
[125][35][45]	0.00051		

†1, LA; 2, NG; 3, PFNCA; 4, GO; 5, GP. Terms grouped together in square brackets indicate that the interaction between all the terms and all lower order interactions involving them are included in the model.