Answers to

Report on Social Policy Targeting and Binary Information Transfer between Surveys

"This paper can be published with the following modifications.

(i) Many of the statements in the first few paragraphs in the Introduction section need to be authenticated by giving appropriate references."

See footnotes 2 and 3.

(ii) The readers may sometimes skip reading footnotes. So any statement made on the basis of remarks in the foot note should be avoided. E.g., the abbreviation GC should be expanded.

We fixed footnote 14.

(iii) The formula $H_{Pop} = \sum_i nn_i + under20_i \times \frac{nn_i}{over20_i}$ seems to be incorrect. The

correct formula is $H_{Pop} = \sum_i nn_i + \sum_i under20_i \times \frac{nn_i}{over20_i}$

or $\qquad H_{Pop} = \sum_i [(nn)_i + under20_i \times \frac{nn_i}{over20_i})$.

This was a typo. We fixed it. See p. 9. Of course the calculations were done correctly.

The author may check the following formula for estimating the same which seems to be more robust than the above formula.

$$(\sum_i [[nn]_i ([over20]_i] + [under20]_i))/(\sum_i [nn]_i \times$$
.

As can be seen from the referee's suggested formula, this would not provide an absolute population number, which is of course what we're looking for, but rather some kind of ratio. Our formula simply sums up all the weights of the (actually sampled) persons of 20 years and over and adds to them the weighted number of the persons below age 20 (which we actually calculated from the more detailed file provided to us than the typical public-use-file). However, since no such weights (for the under20) are available in the original survey, the statistician responsible in the Israeli Central Bureau of Statistics suggested to calculate the

relevant weights as the average weight of the sampled persons aged over20, since some of them might be living in the same household. At every step in our actual calculations we double-checked our results with actual age-specific population data and found them to be well in line with the actual overall distribution.

(iv) It is not clear why the "Optimal BIT" and "Education based model" estimates are always biased downwards. If the estimates are always biased in one direction then there must be a method for improving the estimates.

As for the "education based model" the reason for the downward bias is straightforward: As described in the paper, the Hassidic oriented Ultra-orthodox put much less emphasis on life-long learning of the men than the other Haredi (particularly the Lithuanian Haredi) since they try to enter the labor market without visiting such Bible-study institutes. Therefore the reliance on an education based affiliation necessarily creates a downward bias.[1] This bias is strongly reduced in our method of identifying the Haredi, as can be seen from table 5 in the paper.

Optimality of course does not imply getting rid of all biases, since any such method may still suffer from insufficient information. The method claims to be optimal only in the sense that (1) it takes better advantage of the information provided in the various surveys, (2) we devise a method for minimizing errors in the "source survey", and for optimally carrying over these conclusions to the "target survey".

---

[1] If many non-Haredi were to study in such institutes of high Bible studies as their last school the conclusion might be different. But of course hardly any non-orthodox tends to study at such an institute.