Reviewer’s report

Title: Girl child marriage, socioeconomic status, and undernutrition: Evidence from 35 countries in sub-Saharan Africa

Version: 2 Date: 06 Nov 2018

Reviewer: Arijit Nandi

Reviewer's report:

I appreciate the effort from the authors to respond to my earlier comments. I will comment on the two main areas of revision, first the estimation of effects on the absolute (risk difference/probability) scale and, second, the inclusion of mediators in the series of regression adjusted models.

I greatly appreciate the transition from odds ratios to risk differences. Aside from sidestepping the methodological issues with the odds ratio that I alluded to earlier, the risk differences facilitates the interpretation of the main study findings. That said, I do have some concerns about the results and whether they are being estimated and/or interpreted accurately. One obvious red flag is the sheer magnitude of the estimates reported, which in some cases exceed theoretical limits/bounds on the risk difference. For example, on page 6 (Results) there is the following statement: "Conditional on primary education and EA-fixed effects, women who married before age 18 had 200 percentage points increased probability of giving birth before age 18 (risk difference = 1·97, 95% CI [1·94, 1·99], p<0·01)...." Clearly this (i.e., 200 percentage-point increase) cannot be the case—it would be like saying the percentage of women who married before age 18 went from 50% to 250%. The probability is bounded between 0 and 1 and therefore the maximum percentage-point change that can be attributed to any intervention is 100 percentage-points. We see other implausibly large effects reported elsewhere. Continuing just further down on the page: "Girl child marriage was associated with a 78-percentage point reduction in the likelihood of completing secondary education compared to those who married as adults (risk difference = -0·78, 95% CI [-0·82, -0·73]), p<0·01)". With roughly 9% of adult married women completing secondary or higher in the sample (Table 1), this result would imply that child marriage reduces the probability of completing secondary education from 10% to 68%. The maximum reduction that can be observed in this instance is of course 10 percentage-points.

There are a few things that could be going on here, though it is difficult to diagnose from the paper itself. First, it's possible the effects are just being misinterpreted. Maybe we're talking about 2 and 0.78 percentage-point changes, rather than 200 and 78 percentage-point differences; in other words, the decimal place has been shifted two digits to the right. Second, these could still be multiplicative/relative effects rather than additive/absolute effects. The revised text indicates that the estimates were obtained through Poisson regression. The text states: "We estimated Poisson regression models, rather than odds ratios...." (an aside, but this sentence should be edited since one is a type of regression and the other a measure of association/effect—one could obtain several effect measures from the same generalized model with a Poisson
distribution). However, if this was not done with an identity link the estimates might be ratios rather than differences. When I look at Appendix H, for example, the estimates look more like ratios than differences, although the caption suggests the estimates are from a linear regression. One method of diagnosis would be to run some different models for estimating risk differences or changes in probabilities. Although you have a binary outcome, you could run a linear probability model. Epidemiologists do not typically recommend this and generally prefer inverse logit or log-binomial models to output risk differences, but the estimates for a fairly common outcome like underweight will be similar whether you use a linear probability model, logit model, or other means to get the risk difference. The nice thing about the linear probability model is the ease of interpretation. No transformation is necessary. You can compare the estimates you get to those from your Poisson approach and see if they match. If not, this could very well be the issue. You should not get anything close to a 200 percentage-point change from the linear probability model; if you do, there must be some mistake with the code or serious model misspecification. In addition to the linear probability model, you could have estimated the risk differences from a logit model, followed by the margins command to get risk differences. This is what I had recommended earlier. The logit approach is unlikely to have the convergence issues of the log-binomial or log-Poisson approach. This is also slightly preferred to the linear probability model since it's designed for binary outcomes. I'll end here by saying that all of the estimates in the text and tables should be carefully vetted to prevent errors from making it into the published literature.

With respect to mediation, I'll be more concise. I agree that it is useful to look at secondary outcomes, including potential mediators such as early and multiple childbearing, as well as secondary education. One approach would be to examine the association between early marriage and these secondary outcomes in a model that does not include the primary outcome of underweight, as in Table 3. This can help to elucidate whether child marriage is associated with these factors. However, I would not use the Baron & Kenny product method approach to mediation that has been so thoroughly debunked. Thus, I still find the series of models (M2-M5) in Table 2, which regress the primary outcome on child marriage plus the hypothesized intermediaries without consideration of the implicit assumptions to be problematic. I see that the authors have now included a note that estimates from these models are more reflective of direct effects than total effects, but the language is still unclear. For example, M2-M5 are described as adding more "controls" until results are "fully-adjusted" in Model 5. These terms are typically used to discuss potential confounding, not mediation. If formal mediation analyses are beyond the scope of this paper, then I would simply remove the informal mediation analyses. If you plan to pursue these analyses in earnest, please refer to the work of VanderWeele (who happens to be at the authors' institution), which I cited earlier, rather than Sobel-Goodman, which you cited in your response.

Are the methods appropriate and well described?
If not, please specify what is required in your comments to the authors.

No

Does the work include the necessary controls?
If not, please specify which controls are required in your comments to the authors.
No

Are the conclusions drawn adequately supported by the data shown? If not, please explain in your comments to the authors.

No

Are you able to assess any statistics in the manuscript or would you recommend an additional statistical review? If an additional statistical review is recommended, please specify what aspects require further assessment in your comments to the editors.

I am able to assess the statistics

Quality of written English
Please indicate the quality of language in the manuscript:

Acceptable

Declaration of competing interests
Please complete a declaration of competing interests, considering the following questions:

1. Have you in the past five years received reimbursements, fees, funding, or salary from an organisation that may in any way gain or lose financially from the publication of this manuscript, either now or in the future?

2. Do you hold any stocks or shares in an organisation that may in any way gain or lose financially from the publication of this manuscript, either now or in the future?

3. Do you hold or are you currently applying for any patents relating to the content of the manuscript?

4. Have you received reimbursements, fees, funding, or salary from an organization that holds or has applied for patents relating to the content of the manuscript?

5. Do you have any other financial competing interests?

6. Do you have any non-financial competing interests in relation to this paper?

If you can answer no to all of the above, write 'I declare that I have no competing interests' below. If your reply is yes to any, please give details below.

I declare that I have no competing interests

Statement on potential review bias
Please complete a statement on potential review bias, considering the following questions:
1. Did you co-author any publication with an author of this manuscript in the last 5 years?
2. Are you currently or recently affiliated at the same institution as an author of this manuscript? If you can answer no to all of the above, write 'I declare that I did not publish with these authors in the last 5 years and also meet the affiliation criteria”. If your reply is yes to any, please give details below.

I declare that I did not publish with these authors in the last 5 years and also meet the affiliation criteria

I agree to the open peer review policy of the journal. I understand that my name will be included on my report to the authors and, if the manuscript is accepted for publication, my named report including any attachments I upload will be posted on the website along with the authors' responses. I agree for my report to be made available under an Open Access Creative Commons CC-BY license (http://creativecommons.org/licenses/by/4.0/). I understand that any comments which I do not wish to be included in my named report can be included as confidential comments to the editors, which will not be published.

I agree to the open peer review policy of the journal