Reviewer’s report

Title: Is any job better than no job at all? Studying the relations between employment types, unemployment and subjective health in Belgium.

Version: 0 Date: 06 Apr 2017

Reviewer: Fredrik Norstrom

Reviewer's report:

General comment

The submitted manuscript is within an area where more research is required. It is highly important to get a better understanding on how unemployed, precarious employment and stable employment are related to each other. Both to get ideas about the importance of a more stable employment for those with an employment and how important it is to get an employment for the unemployed. If a precarious employment is unfavourable attempts to move the unemployed to a precarious employment might be a poor strategy.

The results of the article is showing somewhat what to expect but nevertheless are the results of high value as what to expect and the magnitude of it needs to be "confirmed". However, there are some major weaknesses of the article that needs to be improved as especially the magnitude of the association between health measures and types of employment (as well as unemployment then) is not trustworthy in current manuscript. The main problem is that the analyses does not well explain how the estimates can be considered to how low bias. This is especially obvious from the different of some of the estimates between model 2 and model 3 and 4. The authors must be able to explain which estimates are trustworthy and why. Current text are written in such a way that both the estimate from model 2 and 3 are considered trustworthy and considering that the odds ratio can go from 3.41 to 1.97 this is not a reasonable conclusion. The second most important issue to deal with is the composition of the four precarious employment groups (one of them the reference group "Standard jobs"). It needs to be better explained and reasoned on why they are good definitions and also a reference for why it is reasonable to use the approach from the authors, i.e. to create variables based on a cluster analysis and to use these variables in the logistic regression based on probabilities. For most of the above mentioned issues are collinearity issues, i.e. correlated exposure variables, a very important thing to deal with. I suggest that the authors are looking into papers that have evaluated important criteria that need to be fulfilled for logistic regression, e.g. Bagley et al as mentioned in one of my comments below, and evaluate whether these criteria are well enough dealt with in the article besides collinearity.

The way the logistic regression is set up is the most important issue to solve among the comments given below but there are also other revisions that are required before the manuscript can be submitted. Despite my criticism towards the analysis is my overall judgement that the
manuscript should have a very good possibility of being accepted for publications. The analyses might not have to be redone, but there must at least be a good reasoning on why the results can be used and what their limitations are from the authors.

Major Compulsory Revisions

Abstract:

1. Results from the study should be presented in the abstract. The background part of the abstract can be much reduced and give space for this. It would also be valuable to have a clearer description of what the health measures are in the method part of the abstract.

Background:

2. Row 154-156: The statement at either needs to be clearer so that the uniqueness in the manuscript is obvious or it needs to be revised. There are previous studies that have include precarious employment and unemployment in the analyses, e.g. Virtanen et al. "Health inequalities in the workforce: the labour market core-periphery structure" and Flint et al. "Do labour market status transitions predict changes in psychological well-being?". My interpretation from current text is that the submitted manuscript consider itself to be the first to include both precarious employment and unemployment in their analyses, which is not the case. Maybe is the solution to be clearer with that employment quality is referring to a combination of 5 different work-related properties instead of the traditional definition of precarious employment and that this is unique.

3. Row 177: It is written "the second main objective". The first objective is not obvious from the paragraph. I strongly recommend to write two-three sentences in an own paragraph where the objectives are clearly written. If not, it must at least be obvious what the first objective is.

Methods:

4. There is no information on how self-perceived general health is measured.

5. The "Center for Epidemiological Studies Depression Scale" is used for self-rated mental health. Seven of twenty items are used in the manuscript but the rationale behind this reduction is not explained. Such explanation should be added to the manuscript.
6. The probabilities to be in each of the four groups are used for four variables that will be highly correlated. Why this unorthodox way of defining variables is working in a logistic regression needs to be explained as the collinearity between the newly created variables could highly bias estimates. A reference to a manuscript that is advising this creation of variables and their use in a logistic regression (or similar statistical method) might handle the issue. It intuitively feels like a reasonable choice to create variables but that does not mean that it will work well in the analysis and an explanation is therefore required.

Results:

7. The presentation of results need to be improved. To start with is the focus on the "uncontrolled" model too large. You are controlling for other variables because you consider there to be important confounders that needs to be included in the analysis. The crude odds ratio is therefore of very low, if any, value for your results. The crude odds ratio would benefit from being presented as crude odds ratio or crude model in Table 2 and 3. I suggest that you rename other models from model 2-4 to model 1-3 and change Model 1 to either crude odds ratio or crude model. It is important to show the crude odds ratio in the table to explain how the odds ratio is affected by the confounding variables. It gives an indication of how important the more advanced models are. It is still, of course, very important to evaluate whether the extended models are improving the estimates of the odds ratio, i.e. decreasing the bias, but this should be dealt with in the discussion where the analysis method are recommended to be commented on. Keep in mind that the crude odds ratio could potentially have the lowest bias even if this unlikely to be the case and when performing the analysis it needs to be understood also if model 2 are presenting better estimates than the crude odds ratios.

The result section would also benefit a lot from a more compact reporting. I suggest that the result section concludes for what labour market positions that there is a significantly poorer health. As it is significant results for all or no model for all labour market positions it might be sufficient to only specify the range for the ORs of current model 2-4.

Discussion:

8. Row 387-394: The discussion should not summarize what has been done. It is sufficient to mention this information in the method part (and the aim). This information should be removed from the discussion section.

9. Row 396-397: When concluding results in the discussion you should rely on your analysis method and it should not be commented in detail how variables where controlled for unless
there are some important message that need to be given. Now the same message is repeated when it is written that it holds also for the more extensive model.

10. Row 398-402: This information are not relevant for this part of the article. It should either be mentioned in the method section why these analysis are performed or in the discussion of methodological concerns why your choice of analysis is relevant. Now it is presented in relation to your results where it is surplus.

11. Row 404-409: As in previous comment, you need to rely on your analysis. Now you are arguing about how additional variables are affecting the estimates and not about which estimate is most reliable. You must not be able to judge which one of your estimates is the most reliable one but you cannot assume that both are somewhat correct, which current phrasing are.

The true estimates of the effect from different labour market positions is not affected by which variables are in the model. It is simply a parameter value that our statistical model to the best of its ability tries to solve. How small the bias is depend on the set-up of variables in the model and the lowest bias is of course for the best set-up of variables which we cannot confirm as the true parameter remains unknown. The different model set-ups need to be validated as far as possible and there should be a good explanation. If the estimate is strongly affected by collinear variables then you need to make sure that collinear variables are not part of the analysis. That some estimates are so highly affected by the addition of household variables needs to be better taken care of in the analyses and the discussion of them. An alternative way of dealing with this is to present stratified estimates and to check whether the estimates are the same for all outcomes for the controlling variable. If not it is a sign of the household variables explaining that the effect on health from e.g. unemployment are different on individual level and it might for such reason be needed to present things differently. However, the possibility to present results on group level are often difficult to do in a trustworthy way due to small sample sizes and might not work for your study for such reason.

12. Row 423-427: In line with the comments above, this needs to better elaborated. My interpretation of this is that the estimate is mainly affected by collinearity between household variables and labour market status variables where some of these are partly derived from similar information in the cluster analysis.

13. The cluster analysis created four groups of individuals. Why these should be considered as valid groups needs to be discussed as this is a very central part of the content of the article.
Minor Essential Revisions

Abstract:

1. Row 30-32: Health inequalities are brought up first in abstract while it is nowhere mentioned in the remaining part of the article. It feels like a good idea to bring it up in the remaining part of the article. Either do so or delete this argumentation from the abstract as it cannot be important in only the abstract.

Background:

2. Row 124: "Find" does not work in this sentence as you are referring to the results of their result.

3. Row 127-129: Grammar needs to be improved. The scale is not what is related to poor health. It is employment quality that is related to poor health.

4. Row 179-181: The text has no value for the article. It only informs about the standard practice in research articles in Archives of Public Health and should therefore in my opinion be deleted. This explanation of what to expect next in the text also feels surplus on row 111-113.

Methods:

5. There is no information about the number of invited in the GGS (only that 7163 participated). This information should be added.

6. On row 207 it is written: "assuming missing at random (MAR)". It is unclear what this means. Are values imputed randomly based on other values or is the best predicted value replacing missing value? The rationale behind the procedure must be rewritten so that it is understandable how values are imputed.

7. Row 233: AIC, BIC and CAIC are mentioned. These short forms have not been explained previously in the text. Revise the manuscript accordingly.

8. Row 274-279: Material deprivation is mentioned. It is unclear if this definition has been used before and whether the questions are part of a validated instrument. Especially item 5 (eating fish, meat or chicken) feels a bit controversial as this might not be due to the lack of possibility to eat these food dishes, e.g. a vegetarian would skip them for obvious reason.
Also in the method section, a clarification of the definition of material deprivation is required.

9. Material deprivation and perceived financial situation are likely to measure very similar things. Please explain how they could both be used in your analysis. It feels obvious that there is a collinearity problem between variables.

10. The social support cut-off value surprises me. From the description, the scale seems to be from 0-60 (6 items with scale 0-10). Is it really as few as 10% that scores 5 or lower or have I misunderstood the calculation of scores for social support? Even if this is correct it needs to be specified what a high score corresponds to. It is confusing that the first question (availability of someone to talk with) is indicating a positive situation while the second question (emptiness) is indicating a negative situation. Please clarify how the score for this measurement is calculated and also check if the cut-off value is correct.

11. Row 330-331: Write "household situation variables" because if I understood it correctly are you using four variables that are indicators of household situation.

Results:

12. The coefficients for controlling variables should according to experts (see e.g. Bagley et al. "Logistic regression in the medical literature: Standards for use and reporting, with particular attention to one medical domain") be presented in articles. They should therefore be added in Table 2 and 3 as they give highly relevant information for your statistical model.

13. "Odds ratios and 95% confidence intervals" should not be part of the title of a table. They should be given either in the table or as a footnote to the tables. Also the number of participants should not be given in the title. As you are using a sample, "n" should be used and not "N". "N" is referring to the population size according to the commonly agreed standard.

Discussion:

14. Policy implications from the article needs to be discussed in the article. The authors are expected to have ideas on how the content of the article can be used in decision making or at least if there is a requirement of further analyses before there can be some advices to give to the decision makers. It would also be valuable to have a suggestion on what future research to recommend based on current study.
References:

15. Reference 13 and 23 are referring to "Self-reference removed". This needs to be corrected.

16. The surname is Paul in reference 17, i.e. author name is Karsten Paul and not Paul Karsten.

Discretionary Revisions

1. The background would be improved if it was more compact and clearer written with a focus on precarious employment in regard to the definition you have decided to use. The background would also benefit from being shortened. However, I don't consider it necessary to rewrite it to fulfil this suggestion.

2. The manuscript would benefit from more references in peer-reviewed journals. At least the references 2, 4, 5, 6, 10, 14, 15, 16, 24, 26, 27, 28, 30 and 31 constitutes of references to books, reports and similar.

Level of interest
Please indicate how interesting you found the manuscript:

An article of importance in its field

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' below.

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

Do you want to get recognition for reviewing this manuscript? Add a record of this review to Publons to track and showcase your reviewing expertise across the world’s journals. Signing up is quick, easy and free!

No