Reviewer’s report

Title: Long term unemployment, income, poverty, and social public expenditure, and their relationship with self-perceived health in Spain (2007-2011)

Version: 0 Date: 22 Jun 2017

Reviewer: Christos Grigoroglou

Reviewer's report:

This is a novel study that explores the effect of unemployment length, poverty and other individual and regional characteristics on self-perceived health. The paper is interesting and the methods are explained in a clear way. The paper is of direct interest to policy and adds to the literature of self-perceived health.

Major

Please provide a justification and some background for your model selection. The fixed-effects multilevel logistic regression is introduced arbitrarily in the manuscript. As you are targeting a world-wide audience, bear in mind that some of your readers may not be familiar with the methodology that you followed. I would suggest a justification of your chosen modelling method, why you chose to follow this path and why (if so) your model works better in this instance in comparison to a standard regression model. In general I think that this is a good use of multilevel modelling. I am not an expert in this family of models but the general assumptions, even though they are strong, seem to be satisfied. Nevertheless, I would like to see a defensive approach and justification towards the selection of the multilevel model instead of a standard OLS model with fixed effects to account for regional differences. I suggest this as there may be some scepticism in the audience that will read the manuscript in regards to model selection.

I am not familiar with the dataset that you use but you need to clarify the rationale behind dichotomisation. Is this variable represented as a dichotomous variable in the original dataset or is ordinal? If the variable is ordinal then there is a great loss of information due to dichotomisation that needs to be discussed in the limitations.

Even though, I understand the use of chronic illness; I was a bit unclear when I first read the manuscript about the use of the chronic illness dichotomous variable to control for the bidirectionality of the association between unemployment and self-perceived health. It might be worth explaining this a bit more. Furthermore, I am not sure that you can say that the causal hypothesis prevails over the selection hypothesis. In my opinion, the use of the dichotomous
chronic illness variable favors indeed the causal hypothesis but does not necessarily rule out the selection hypothesis. Is there any other way to test these hypotheses and thus make inference more clear?

I am wondering whether you had access to other variables (i.e. level of stress, smoking, physical activity, BMI, marital status) that are known to have a strong impact on self-perceived health. I am not familiar with your dataset and whether these variables exist in the dataset, but I feel that if you did not have access to any of other self-perceived health determinants this should be mentioned in the limitations.

I am a bit uncertain as to what the key message is in the introduction. Even though on page 7, lines 151-154 you define the main research question (although, in my opinion it needs rewording, please see previous comment) your introduction is structured heavily around the effect of long-term unemployment on health. I feel that the introduction could be shortened as your main research question is not the effect of long-term unemployment on health (at least according to your title and abstract). For example, on pages 5 and 6 there is a lot of information about the effect of long-term unemployment (which would also be good to define) on health but minimal information about the effect of length of unemployment on health. Of course long term unemployment is relevant but I feel that it is discussed extensively in the introduction.

Minor

In the abstract nowhere you mention the causal and selection hypotheses before. I would suggest removing the statement "the causal hypothesis prevails over the selection hypothesis in periods of deep financial crisis with high rates of unemployment". This is irrelevant to the key message of the manuscript and is not explained before.

I am curious as to what was the motivation behind the exclusion of the autonomous cities of Cauta and Melilla? Why these cities in particular were excluded whereas other autonomous areas in Spain were not?

Page 5, line 91: A reference is needed for the statement "bidirectional relationship between unemployment and perceived health"

Page 5, line 92: I would suggest to remove the term causal, as causality itself is an "loud" term that is not used so frequently in the literature. Also I believe a reference is needed for this statement.

Page 7, line 154: I would remove anything that refers to causality and replace these terms with terms such as association, correlation etc. Causality is a very "loud" term.
Page 7, line 151: This is a bit mixed up as a research question. What exactly is your association of interest? I am a bit confused as to which one is the research question. Please rephrase this by focusing on the main research question and say while controlling for area and individual characteristics such as etc.

Page 12, line 236: I am curious about the use of MEQRLOGIT instead of MELOGIT command in Stata? What was the rationale behind this?

Page 18, line 366: I agree with this statement but you could justify this a bit more based on the huge theoretical background that is available out there in terms of the reliability of self-perceived health.

Page 18, line 373: I would also suggest talking about the loss of information due to the dichotomisation of self-perceived health. Even though you justify it at the beginning of the manuscript, it remains a limitation which should be mentioned.

Even though secondary, you can also report p-value in table 3. Although it's not necessary it might facilitate some readers.

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

Yes

Does the work include the necessary controls?
If not, please specify which controls are required in your comments to the authors.

Yes

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

Yes

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'

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