

## PEER REVIEW HISTORY

BMJ Open publishes all reviews undertaken for accepted manuscripts. Reviewers are asked to complete a checklist review form (<http://bmjopen.bmj.com/site/about/resources/checklist.pdf>) and are provided with free text boxes to elaborate on their assessment. These free text comments are reproduced below.

### ARTICLE DETAILS

<b>TITLE (PROVISIONAL)</b>	The Interrelationship among Economic Activities, Environmental Degradation, Material Consumption, and Population Health in Low-Income Countries: A Longitudinal Ecological Study
<b>AUTHORS</b>	Chuang, Ying-Chih; Huang, Ya-Li; Hu, Ching-Yao; Chen, Ssm-Ching; Tseng, Kuo-Chien

### VERSION 1 - REVIEW

<b>REVIEWER</b>	Vijayan K Pillai University of Texas at Arlington USA
<b>REVIEW RETURNED</b>	18-Oct-2014

<b>GENERAL COMMENTS</b>	This is a well organized study in general. The research questions are clear and the analysis is competently undertaken. However, I feel, the authors have not adequately informed the reader on the procedures used in this study for handling missing data. Please describe the pattern of missing data in general and a few sentences on how the problem was handled.
-------------------------	---

<b>REVIEWER</b>	Katherine Muldoon University of British Columbia, Vancouver, Canada
<b>REVIEW RETURNED</b>	13-Jan-2015

<b>GENERAL COMMENTS</b>	<p>The authors have put together a very interesting piece of research with data from standardized sources (i.e. UN, WDI etc) in a novel way informed by the ecological unequal exchange theory. I think this article makes an important contribution to the literature through the use of standardized data sources and on the associations between health, environment and economy. I have included technical revisions that apply to the methods and results section that are designed to improve the scientific contribution of the article.</p> <p>Methods Section Page 6, line 5 -These analyses can also be conducted on HIC. It would provide strength for the authors to add a justification sentence about why this analysis excludes high-income countries/core-countries.</p> <p>Page 6 , line 10 -The authors state that they have only included countries without missing data. Can you provide a sentence or two about how many countries were excluded because of missingness and where they were from. There is also a section in the methods on</p>
-------------------------	--

	<p>page 9 that describes some imputation of missing data. It also seems like the authors chose the linear mixed modeling because of its ability to deal with missingness. Please clarify.</p> <p>Page 6, line 19 – The groupings of the independent variables are important to the analysis. Using sub-heading would clarify the types of independent variables being used. For example: Dependent variable: IMR, U5MR. Independent variables: Economic, Environmental, Demographic. The use of the term ‘control variables’ does not seem appropriate in this explanatory model.</p> <p>Page 6, line 45-60 - The sentence beginning with ‘Dependency perspectives suggest that developed countries...’ and the rest of the paragraph seems out of place in the methods and might be better placed in the intro/discussion.</p> <p>Some additional technical information for each measure would be important to add. For example, IMR/CMR are rates and need a denominator (x per 100,000 live births). For the statistics that use percentages, please include these in the tables.</p> <p>Page 8, line 30 - It confusing to read the first sentence listing democracy, health services, female education, GDP per capita, and period effect. I would remove this sentence and just describe the variable as it is in each paragraph. This would be consistent with the format of the previous variables</p> <p>Page 8, line 30 - Please provide a reference for the correlation between fertility rate, GDP and female education</p> <p>Page 9, line 37 - There are analyses displayed in the results section that are not described in the analysis section. Please indicate the estimates that are used to display the linear mixed effects model. I am assuming they are beta’s. If possible, please log transform and use odds/risk ratios for more clear interpretation. It would clarify the results to see the estimate/magnitude of the association in addition to the direction of the relationship (positive/negative).</p> <p>Please include the analysis plan for each of the models 1-4. I was surprised to see this in the results section.</p> <p>Results</p> <p>Page 10, line 3 - If the sample size was limited, why did the authors conduct 3 stratified analyses? Please clarify. I would recommend including geographic region as an independent variable in place of stratifying all the analyses. Geographic region can also be used as a clustering variable in the mixed effects modeling.</p> <p>There are a total of 3 stratified regressions with 4 sub-analytic models on both IMR and U5MR, in addition to a meditational analysis. It is bordering on too much information. I would recommend restricting the number of sub-analyses (Model 1-4) or collapse the stratified analyses and include the geographic variable as an independent variable.</p> <p>Please describe the meditational analysis statistical procedure in the analysis section.</p> <p>Page 10, line 10- please provide estimates from the regression</p>
--	--

	<p>modeling, not just the direction</p> <p>Appendix 1 – very useful contribution. Can you please include a column including the breakdown of number of years contributed by each country that sums to 1497. I believe this value is a measure of country-years contributed to the analysis. Please clarify.</p> <p>Tables Table 1. Please include the number of countries included in each sub-analysis. For example – Sub-Saharan Africa (n=x), Latin America (n=x) etc. Including sub-headings within the table would clarify the variable groupings (as suggested above). Please include the units for each independent variable where applicable. For example, rural population (%). Can you please clarify the year that the variables from Table 1 are populated from. Is it cumulative, baseline? The title for tables 1 is vague.</p> <p>Tables 2-4 are very crowded and difficult to interpret. Please clarify what point estimates are being used. It is unclear which statistics are p-values. It is conventional to restrict the statistics to 2-3 decimal places. Either use the symbol* or the p-value, but not both.</p> <p>Typo Look for case consistencies with sub-Saharan Africa (lower case 'sub')</p>
--	--

### VERSION 1 – AUTHOR RESPONSE

Reviewer 1: Vijayan K Pillai

I feel, the authors have not adequately informed the reader on the procedures used in this study for handling missing data. Please describe the pattern of missing data in general and a few sentences on how the problem was handled.

We appreciate the reviewer's comments. We added more description about the pattern of missing data and how we used the linear mixed model with a spatial power covariance structure to handle the problem of missing cases.

We added the following paragraph in Methods:

Among 91 countries which met the selected criteria, 66 low-income countries with no missing data on any of the predictors used in this study were included in the analyses. Among the excluded countries, 3 were in sub-Saharan Africa; 8 were in Latin America; and 14 were in the other regions.

We added the following paragraph in the analysis section:

Because of the nature of the repeated measurements conducted in our study design, we used linear mixed models (SAS Proc Mixed) with a spatial power covariance structure and a correlation that decreased over time to accommodate the possible problem of a serial correlation of errors. In cross-national longitudinal studies, it is common to have missing assessment waves within or across countries. We used this statistic model because a spatial-power covariance structure accounted for uneven spacing of follow-up assessments. A spatial-power model is similar to an autoregressive model in that observations closer in time are more correlated than observations farther apart, but it does not require the time between observations to be equal. Therefore, this model can accommodate different missing data points within or across countries.

Reviewer: 2: Name Katherine Muldoon

### 1. Methods Section

Page 6, line 5 -These analyses can also be conducted on HIC. It would provide strength for the authors to add a justification sentence about why this analysis excludes high-income countries/core-countries.

We added a few sentences in the Methods section as follows to address this suggestion. We focused on low-income countries because prior studies have suggested that the impact of economic activities on population health is much greater in low-income countries than in high-income countries.

2. Page 6 , line 10 -The authors state that they have only included countries without missing data. Can you provide a sentence or two about how many countries were excluded because of missingness and where they were from. There is also a section in the methods on page 9 that describes some imputation of missing data. It also seems like the authors chose the linear mixed modeling because of its ability to deal with missingness. Please clarify.

We modified the section of sample description as follows:

We selected a sample of countries according to the classification system developed by Clark and Beckfield (2009).[11] We excluded core countries, according to their definition, and excluded countries classified as high-income countries based on the definition established by the World Bank.[5] Among 91 countries which met the selected criteria, 66 low-income countries with no missing data on any of the predictors used in this study were included in the analyses. Among the excluded countries, 3 were in sub-Saharan Africa; 8 were in Latin America; and 14 were in the other regions.

We added more description about the linear mixed models:

Because of the nature of the repeated measurements conducted in our study design, we used linear mixed models (SAS Proc Mixed) with a spatial power covariance structure and a correlation that decreased over time to accommodate the possible problem of a serial correlation of errors. In cross-national longitudinal studies, it is common to have missing assessment waves within or across countries. We used this statistic model because a spatial-power covariance structure accounted for uneven spacing of follow-up assessments. A spatial-power model is similar to an autoregressive model in that observations closer in time are more correlated than observations farther apart, but it does not require the time between observations to be equal. Therefore, this model can accommodate different missing data points within or across countries.

Regarding the imputation of values for the measurements of health services (page 9), this imputation procedure was conducted before begging all of the other regression analyses to ensure a larger number of cases included in the study.

3. Page 6, line 19 – The groupings of the independent variables are important to the analysis. Using sub-heading would clarify the types of independent variables being used. For example: Dependent variable: IMR, U5MR. Independent variables: Economic, Environmental, Demographic. The use of the term 'control variables' does not seem appropriate in this explanatory model.

Thanks for the reviewer's suggestions. We modified the tables according the suggestions.

4. Page 6, line 45-60 - The sentence beginning with 'Dependency perspectives suggest that

developed countries...’ and the rest of the paragraph seems out of place in the methods and might be better placed in the intro/discussion.

We modified the sentences and moved them to the Introduction section.

5. Some additional technical information for each measure would be important to add. For example, IMR/CMR are rates and need a denominator (x per 100,000 live births). For the statistics that use percentages, please include these in the tables.

We included the information both in Methods and the tables.

6. Page 8, line 30 - It confusing to read the first sentence listing democracy, health services, female education, GDP per capita, and period effect. I would remove this sentence and just describe the variable as it is in each paragraph. This would be consistent with the format of the previous variables

We removed the sentence as requested.

7. Page 8, line 30 - Please provide a reference for the correlation between fertility rate, GDP and female education

We added the references for the correlation between fertility rates, GDP, and female education.

8. Page 9, line 37 - There are analyses displayed in the results section that are not described in the analysis section. Please indicate the estimates that are used to display the linear mixed effects model. I am assuming they are beta’s. If possible, please log transform and use odds/risk ratios for more clear interpretation. It would clarify the results to see the estimate/magnitude of the association in addition to the direction of the relationship (positive/negative).

We are sorry that we did not make it clear in the previous submission. The number in each table cell without parentheses is the regression coefficient for each variable. The number within the parenthesis is the standard error for that regression coefficient. In Table 2, 3, 4, we log transformed the outcome variables. We did not log transform most of the independent variables because independent variables are not required to have normal distribution in a regression model. Since the outcomes are continuous variables, we decided to keep the original beta’s. It is easier for readers to recalculate and examine the results of this study.

9. Please include the analysis plan for each of the models 1-4. I was surprised to see this in the results section.

We included our analysis plan following the description about mediation analysis.

In our analyses, models were firstly fitted with economic variables and demographic variables. Then, models were fitted with economic variables, environmental variables, and demographic variables. Last, we examined the relationship between the mediator and the outcomes by regressing environmental variables on economic variables and demographic variables.

## 10. Results

Page 10, line 3 - If the sample size was limited, why did the authors conduct 3 stratified analyses? Please clarify. I would recommend including geographic region as an independent variable in place of stratifying all the analyses. Geographic region can also be used as a clustering variable in the mixed effects modeling.

We appreciate the reviewer's suggestion. We did conduct mixed effects modeling using the whole sample and the geographic region as an independent variable. According to the results, several economic variables were significantly associated with infant mortality rates and under-five mortality rates. However, when we added interaction terms of region and economic variables, we found the some of the direction of the relationships were reversed, which indicated that it is necessary to stratify the analyses by region because the directions of the relationship between economic factors and IMR may vary across region. Therefore, we kept our original stratified analyses.

11. There are a total of 3 stratified regressions with 4 sub-analytic models on both IMR and U5MR, in addition to a meditational analysis. It is bordering on too much information. I would recommend restricting the number of sub-analyses (Model 1-4) or collapse the stratified analyses and include the geographic variable as an independent variable.

To simplify the tables, we deleted Model 1 for each region in Table 2 and Table 3.

12. Please describe the meditational analysis statistical procedure in the analysis section.

We added the following description:

Our mediation analyses followed the approach outlined by Baron and Kenny (1986) to assess whether environmental variables (ecological footprints and CO2 emissions) were the mediators between the independent variables (economic characteristics) and the dependent variables (infant or under-five mortality rates). The first step is to regress the dependent variable on the independent variables to confirm that the independent variable is a significant predictor of the dependent variable. The second step is to regress the mediator on the independent variable to confirm that the independent variable is a significant predictor of the mediator. The third step is to regress the dependent variable on both the independent variable and the mediator. If the relationship between the independent variable and the dependent variable disappears, then a mediating relationship is established.

13. Page 10, line 10- please provide estimates from the regression modeling, not just the direction

We included the estimate from the regression models in the Results.

14. Appendix 1 – very useful contribution. Can you please include a column including the breakdown of number of years contributed by each country that sums to 1497. I believe this value is a measure of country-years contributed to the analysis. Please clarify.

Thanks for the reviewer's suggestion. We added a column including number of years contributed by each country in the Appendix.

## 15. Tables

Table 1. Please include the number of countries included in each sub-analysis. For example – Sub-

Saharan Africa (n=x), Latin America (n=x) etc. Including sub-headings within the table would clarify the variable groupings (as suggested above). Please include the units for each independent variable where applicable. For example, rural population (%). Can you please clarify the year that the variables from Table 1 are populated from. Is it cumulative, baseline? The title for tables 1 is vague.

We included the number of countries in each sub-analysis in the tables. We specified the units of variables, changed the title, and included the year of data collection in Table 1.

16. Tables 2-4 are very crowded and difficult to interpret. Please clarify what point estimates are being used. It is unclear which statistics are p-values. It is conventional to restrict the statistics to 2-3 decimal places. Either use the symbol\* or the p-value, but not both.

We deleted model one and included three models for each region to simplify the tables. The number in each table cell without parentheses is the regression coefficient for each variable. The number within the parenthesis is the standard error for that regression coefficient. We use the symbol\* to indicate the significant relationship at the p value<.05. We restricted the statistics to 3 decimal places if the number is not too small and can be allowed to 3 decimal places.

17. Typo- Look for case consistencies with sub-Saharan Africa (lower case 'sub')

We corrected this typo.

Editors Comments:

1. Please strengthen the conclusions of your abstract so there is an emphasis on health here

We changed the conclusion to:

This study suggested that environmental protection and economic development are both important for reducing infant or under-five mortality rates in low-income countries.

2. Please include the study design in your title

We changed to title to : "The Impact of Trade and Economic Activities on Environmental Degradation, Material Consumption, and Population Health in Low-Income Countries: A Longitudinal Ecological Study".

#### VERSION 2 – REVIEW

<b>REVIEWER</b>	Vijayan K Pillai University of Texas at Arlington, USA
<b>REVIEW RETURNED</b>	02-Mar-2015

<b>GENERAL COMMENTS</b>	The author has adequately responded to the queries I had at the first review
-------------------------	--

<b>REVIEWER</b>	Katherine Muldoon University of British Columbia, Canada
<b>REVIEW RETURNED</b>	19-Feb-2015

<b>GENERAL COMMENTS</b>	<p>This is a very interesting analysis and the authors have addressed all my previous reviews. I have included a small list of minor revisions that the authors are trusted to make.</p> <p>Inconsistencies with terms – poor countries, low income countries, less-developed countries, high-income countries, core countries</p> <p>Introduction, paragraph 3 – ‘Various studies have tested the ecological unequal exchange theory.’ – Can you include the citations for the reader.</p> <p>Methods, paragraph 9 – there is still confusion regarding missing data. In the abstract and paragraph 1 of the methods section it states that 66 countries with no missing data on any predictors were included in the analysis. Paragraph 9 describes a process for dealing with missing data. The analysis section describes the assessment of missingness. The authors have developed processes for dealing with missing data, which adds strength to the analysis. Is it just missingness based on the outcome? I would recommend removing the statements that countries were excluded because of missing data.</p> <p>Methods, measures section – It would clarify the analysis plan to include a sub-heading for the dependent variables and the independent variables. These terms are used further down in the methods section when describing the mediation analysis. This would add clarity and consistency with terms.</p>
-------------------------	--

### VERSION 2 – AUTHOR RESPONSE

1. Impact' still gives the impression that cause-and-effect is being looked at

We changed the title to “The Interrelationship among Economic Activities, Environmental Degradation, Material Consumption, and Population Health in Low-Income Countries: A Longitudinal Ecological Study”. We also revised the statements, which used the term “impact”.

1. Inconsistencies with terms – poor countries, low income countries, less-developed countries, high-income countries, core countries

We used “low-income countries” to ensure consistency with terms. We only kept one other term “core countries” because it was defined and used specifically in the theories.

2. Introduction, paragraph 3 – ‘Various studies have tested the ecological unequal exchange theory.’ – Can you include the citations for the reader.

We provided the reference as the reviewer suggested.

3. Methods, paragraph 9 – there is still confusion regarding missing data. In the abstract and paragraph 1 of the methods section it states that 66 countries with no missing data on any predictors were included in the analysis. Paragraph 9 describes a process for dealing with missing data. The analysis section describes the assessment of missingness. The authors have developed processes for dealing with missing data, which adds strength to the analysis. Is it just missingness based on the

outcome? I would recommend removing the statements that countries were excluded because of missing data.

We handled the missing data based on all of the variables in the analyses. To avoid the confusion regarding missing data, we deleted the statements that countries were excluded because of missing data.

4. Methods, measures section – It would clarify the analysis plan to include a sub-heading for the dependent variables and the independent variables. These terms are used further down in the methods section when describing the mediation analysis. This would add clarity and consistency with terms.

We added the sub-heading for the variables as the reviewer suggested.