

Substance:

- This is a difficult research topic, both because it can be approached from many points of view and because relevant data are always likely to be difficult to obtain.
- The authors' specific research questions are clearly stated on p 4.
- There is a literature review, but it is somewhat shorter and haphazard. It seems to me unlikely that this is a representative or sufficiently comprehensive review of the relevant literature. In particular, it is not very well matched to the research questions in this project (i.e. what is already known about those questions).
- The discussion in the end is reasonably well matched to the original research questions (even if the conclusions end up being largely flat, which is not in itself a problem).

Methods:

- It is nice that the authors have been able to find any relevant data on this topic. A better description of these data should have been given. Where do they come from? How were the respondents selected/sampled? Note also that all of the respondents had purchased some luxury items recently. This immediately changes the meaning of the results: instead of the general population, we can only aim to answer the research questions for those who have made some luxury purchases.
- As noted in the report, data are only available in the form of three-way contingency tables of country x a demographic group x response to a survey question. This is a severe limitation, because it rules out multivariate analysis of more than three such variables at a time. The authors do their best with this, but in the end this setting puts serious constraints on what can be learned from the analysis.
- Summary statistics of the data would have been useful. This would have included at least country x age and country x gender tables. Some aggregate percentages are now reported in passing in the introduction, but they would have been better placed in the analysis section or an appendix.
- The main method of analysis is (negative binomial) log-linear modelling for these contingency tables. This is something that the authors have learned for themselves as part of the project. That is a strength of this project. The description is broadly correct, but there are also some weaknesses or misunderstandings:
 - The model specification is inconsistently stated. At least three versions are mentioned: (1) a saturated model, including the three-way interaction (p 15); (2) model with all two-way interactions (p 17), and (3) model with group x response as the only two-way interaction (p 17 and appendix E). Of these, (2) would be most natural for this analysis. Specification (3) is not appropriate, because it constrains the model unnecessarily by not allowing anything the distributions of groups and responses to vary by country. Which one was actually used?
 - The discussion in Table 1 shows good care and attention to detail. However, some of it is also a bit beside the point. The items labelled "already satisfied" are indeed inherent features of the model set-up, so they need not be discussed as "assumptions". Similarly, "lack of multicollinearity" is satisfied by design, because the "explanatory variables" in a log-linear model like this are determined purely (and deterministically) by the structure of the contingency table.
- The results (group differences) are so clearly and consistently non-significant that it seems almost surprising. I wonder (but cannot of course be sure) if there is something in the analysis itself that contributed to this result. One possibility is the specification of a negative binomial model, which inevitably inflates uncertainty in the estimates

(compared to the Poisson model). There is little information about the overdispersion parameter in contingency table data, so allowing for it may make estimates unstable.

- Since no differences are significant by conventional criteria, the discussion of them (from p 23 onwards) could strictly speaking have been omitted or at least shortened. If this is discussed, it should be done in terms of informative parameters. The interaction parameters (lambdas in the model notation) are not useful for this purpose. Instead, associations from log-linear models are most often summarised in terms of relevant odds ratios, which are functions of the lambda-terms.

Presentation

- The report is generally well written and fairly neatly formatted.
- The information given in the appendix is helpful.

Comments from the second marker:

I agree with the comments from the first marker. The research questions in this project address an interesting problem, but the approach to the statistical analysis of contingency tables can be improved. Below some additional comments:

Strengths:

- The authors identify and acknowledge the limitations of the dataset used in this project.
- Modelling strategy is sensible and carefully chosen. The description of the log-linear model is informative given the relative “novelty” of the approach (considering the audience).
- The guidelines for coefficient interpretation are useful and help the reader better understand the results from the estimated models.
- The presentation of the results for the estimated interaction coefficients, while somewhat wordy and potentially better presented in Table format, is detailed and is presented in accordance with the research questions stated in the Introduction.

Points for improvement:

- The literature review is very limited and scattered. There is no thread that helps the reader connect the ideas and infer what is the direction of the project.
- An important limitation of this study comes from the dataset. The data are aggregated, so no traditional regression analysis can be made to explore the research questions the authors propose. The authors do their best to
- Data description was done in the introduction, out of context. In Section 2, the approach is limited: readers do not have context of the dataset structure (e.g., individuals from each country, or income distribution in each country, etc.)
- The connection between Poisson regression and log-linear models is well documented, but it is out of context and somewhat unnecessary given the structure (and limitations) of the dataset.
- While overdispersion is an important matter in regression settings where the data are observed at the individual level, it might be less so in the context of log-linear models for contingency tables. The authors leave the explanation of why the Negative-binomial model would imply a better fit than the Poisson model in this case to the Appendix.