

## Class 3 – Problem set

Due on Canvas on or before Sep 27

Submit a single document with all responses in it

Summary: You will write three separate “mini-papers,” one each about a different regression model. The first paper will feature a mediation, the second a dichotomous moderator, and the third a continuous moderator.

Notes on how to do it:

1. In real research HARKing and fishing are bad. But in this exercise feel free to do both. I’d like you to have evidence of statistically significant mediation and moderation, so you may need to do a little some searching.

FYI, some of you seem to be having fun finding your own variables and concocting theories to explain them. That’s great, and you should keep doing it; you’ll learn more that way. At the same time, I want to respect my new understanding of how highly varied the levels of comfort with R and with stats are in the group. For folks who are not having fun finding variables and/or who are getting lost in the minutia of R, here are a few hints:

- a. JobLearn mediates the relationship between WorkVary and JobSat
  - b. FindNew and Female interact when predicting JobSat
  - c. FindNew and EDUC interact when predicting JobSat
2. Your “theoretical” explanations may be highly tenuous. That’s fine, so long as you don’t say anything that is logically impossible (e.g., absent time travel, current income cannot cause childhood income, because the temporal order is wrong).
  3. The explanations may also be quite short. That’s fine too. But I do want to see an actual explanation of why.
  4. If you use the GSS data, each of your models can use at most one of the variables I used in the class example (i.e., Wordsum, Childwealth, Educ, SEI, Age, Married).
  5. Beyond the limits of #4, it’s fine to reuse variables (e.g., it’s okay if you have the same IV and DV in all three papers).
  6. Please take the “paper” framework seriously. We have three goals here. In order of descending importance, they are – (a) learn to understand regression models, (b) learn to use and report the statistics of regression, and (c) learn to use R. In the first two problem sets, the emphasis was lightest on the second goal, but it really is important. You’re not

taking this course to become a data scientist, you're taking this course to become an empirical researcher, so you need to practice the reporting norms of the profession.

For each "paper:"

- 1) Choose three appropriate variables from the GSS data or data of your own
  - i) If you choose to use your own data, be sure that you give me enough explanation to understand it.
  - ii) If you needed to use any of your own cleaning code, please include it in the Appendix.
- 2) Please write highly abbreviated versions of all the parts of a paper: abstract, intro, theory/background, methods, results, discussion, bibliography (if needed), tables, and figures (when needed). Also include an appendix that has all the R code you used for the paper.
  - i) The "pretend paper" format should be used throughout your submission. Unlike previous problem sets, do not include R code in your main text, only in the Appendix.
  - ii) Likewise, do not organize your submission around these numbered steps. You need to complete all the steps, but what you submit must look like a journal submission.
  - iii) At the same time, you are not writing a 30-page paper. The abstract may be only 2-3 sentences, the entire intro is likely one short paragraph, you should probably only have 1 paragraph explaining each hypothesis, etc. Only the results section should be written in full detail.
- 3) State explicit hypotheses in the format used in journal articles
  - a) The mediation paper should have two Hs: (i) how IV affects DV and (ii) the mediation
  - b) Each moderation paper should also have two Hs: (i) IV to DV and (ii) the moderation
  - c) Be sure to include at least the direction/valance in the Hs. (If you can include effect sizes, so much the better)
- 4) Create a "Table 1" summary of your data.
- 5) Be sure that your results section includes text references to p-values, confidence intervals, and effect sizes.
- 6) Regarding results
  - a) In the mediation paper, report bootstrap results and the extent of mediation.
  - b) In the dichotomous moderation paper, include a simple slopes plot
  - c) In the continuous moderation paper, include a plot of the entire interaction, not simple slopes
  - d) If Johnson-Neyman analysis gives you an interesting and meaningful result, include it. Otherwise, it doesn't need to be reported.

- 7) No need for model diagnostics. They often don't get reported in papers. Plus, we know that they're likely to be bad because of the limitations of the GSS data. So, like the HARKing, we'll allow some bad practice since this work is not for actual submission.

As a concrete example, perhaps look at this paper

Kim, Y. J., & Kim, J. (2020). Does negative feedback benefit (or harm) recipient creativity? The role of the direction of feedback flow. *Academy of Management Journal*, 63(2), 584-612.

It is from a recent AMJ issue. Study 1 reports both moderation and mediation, so you can see what such a report looks like in practice. I chose this article only because it was the first one I came across while flipping through the journal; you'll find many other examples. (Interesting aside: these authors report a hierarchical regression -- which you know I don't favor -- but then they do what I think is the right thing; they ignore the early stage models and only discuss the final ones. That's good. I don't know, but I'd guess that a reviewer who didn't know any better insisted they include the hierarchical regression, and they did so but kept their reporting focused on the correct parts.)

Note that Kim and Kim (2020) do not report the extent of mediation, though you should in your submission. Regarding that, you might look at

Yang, F., Liu, J., Huang, X., Qian, J., Wang, T., Wang, Z., & Yu, H. (2018). How supervisory support for career development relates to subordinate work engagement and career outcomes: The moderating role of task proficiency. *Human Resource Management Journal*, 28(3), 496-509.

Section 4.2 has a simple report of mediation that includes reference to proportions of mediation.