Generalized Ordered Logit Models
Part II: Interpretation

Richard Williams
University of Notre Dame, Department of Sociology
rwilliam@ND.Edu
Updated March 27, 2019
https://www.nd.edu/~rwilliam/

Violations of Assumptions

- We previously talked about violations of the parallel lines/ proportional odds assumption. Parallel lines isn't too hard to understand – but what does proportional odds mean?

- Here are some hypothetical examples

Example of when assumptions are not violated

Model 0: Perfect Proportional Odds/ Parallel Lines

```
            |                    attitude
    gender  |       SD          D          A         SA  |     Total
------------+------------------------------------------------+----------
       Male |       250        250        250        250  |     1,000
     Female |       100        150        250        500  |     1,000
------------+------------------------------------------------+----------
      Total |       350        400        500        750  |     2,000
```

	1 versus 2, 3, 4	1 & 2 versus 3 & 4	1, 2, 3 versus 4
Odds$_M$	750/250 = 3	500/500 = 1	250/750 = 1/3
Odds$_F$	900/100 = 9	750/250 = 3	500/500 = 1
OR (Odds$_F$ / Odds$_M$)	9/3 = 3	3/1 = 3	1/ (1/3) = 3
Gologit2 Betas	1.098612	1.098612	1.098612

Gologit2 χ^2 (3 d.f.)	176.63 (p = 0.0000)
Ologit χ^2 (1 d.f.)	176.63 (p = 0.0000)
Ologit Beta (OR)	1.098612 (3.00)
Brant Test (2 d.f.)	0.0 (p = 1.000)
Comment	If proportional odds holds, then the odds ratios should be the same for each of the ordered dichotomizations of the dependent variable. Proportional Odds works perfectly in this model, as the odds ratios are all 3. Also, the Betas are all the same, as they should be.

Examples of how assumptions can be violated

Model 1: Partial Proportional Odds I

```
           |                    attitude
   gender  |     SD          D          A         SA  |      Total
-----------+---------------------------------------------+----------
     Male  |    250        250        250        250  |     1,000
   Female  |    100        300        300        300  |     1,000
-----------+---------------------------------------------+----------
    Total  |    350        550        550        550  |     2,000
```

	1 versus 2, 3, 4	1 & 2 versus 3 & 4	1, 2, 3 versus 4
Odds$_M$	750/250 = 3	500/500 = 1	250/750 = 1/3
Odds$_F$	900/100 = 9	600/400 = 1.5	300/700 = 3/7
OR (Odds$_F$ / Odds$_M$)	9/3 = 3	1.5/1 = 1.5	(3/7)/(1/3) = 1.28
Gologit2 Betas	1.098612	.4054651	.2513144

Gologit2 χ^2 (3 d.f.)	80.07 (p = 0.0000)
Ologit χ^2 (1 d.f.)	36.44 (p = 0.0000)
Ologit Beta (OR)	.4869136 (1.627286)
Brant Test (2 d.f.)	40.29 (p = 0.000)
Comment	Gender has its greatest effect at the lowest levels of attitudes, i.e. women are much less likely to strongly disagree than men are, but other differences are smaller. The effect of gender is consistently positive, i.e. the differences involve magnitude, not sign.

Examples of how assumptions can be violated

Model 2: Partial Proportional Odds II

```
             |                    attitude
    gender   |       SD           D          A          SA |     Total
-----------+----------------------------------------------+----------
     Male   |       250         250        250         250 |     1,000
   Female   |       100         400        250         250 |     1,000
-----------+----------------------------------------------+----------
    Total   |       350         650        500         500 |     2,000
```

	1 versus 2, 3, 4	1 & 2 versus 3 & 4	1, 2 3 versus 4
Odds$_M$	750/250 = 3	500/500 = 1	250/750 = 1/3
Odds$_F$	900/100 = 9	500/500 = 1	250/750 = 1/3
OR (Odds$_F$ / Odds$_M$)	9/3 = 3	1/1 = 1	(1/3)/(1/3) = 1
Gologit2 Betas	1.098612	0	0

Gologit2 χ^2 (3 d.f.)	101.34 (p = 0.0000)
Ologit χ^2 (1 d.f.)	9.13 (p = 0.0025)
Ologit Beta (OR)	.243576 (1.275803)
Brant Test (2 d.f.)	83.05 (p = 0.000)
Comment	Gender has its greatest – and only – effect at the lowest levels of attitudes, i.e. women are much less likely to strongly disagree than men are. But, this occurs entirely because they are much more likely to disagree rather than strongly disagree. Other than that, there is no gender effect; men and women are equally likely to agree and to strongly agree. The ologit estimate underestimates the effect of gender on the lower levels of attitudes and overestimates its effect at the higher levels.

Examples of how assumptions can be violated

Model 3: Partial Proportional Odds III

```
           |                    attitude
    gender |      SD          D          A          SA |     Total
-----------+--------------------------------------------+----------
      Male |     250        250        250        250 |     1,000
    Female |     100        400        400        100 |     1,000
-----------+--------------------------------------------+----------
     Total |     350        650        650        350 |     2,000
```

	1 versus 2, 3, 4	1 & 2 versus 3 & 4	1, 2, 3 versus 4
Odds$_M$	750/250 = 3	500/500 = 1	250/750 = 1/3
Odds$_F$	900/100 = 9	500/500 = 1	100/900 = 1/9
OR (Odds$_F$ / Odds$_M$)	9/3 = 3	1/1 = 1	(1/9)/(1/3) = 1/3
Gologit2 Betas	1.098612	0	-1.098612

Gologit2 χ^2 (3 d.f.)	202.69 (p = 0.0000)
Ologit χ^2 (1 d.f.)	0.00 (p = 1.0000)
Ologit Beta (OR)	0 (1.00))
Brant Test (2 d.f.)	179.71 (p = 0.000)
Comment	The effect of gender varies in both sign and magnitude across the range of attitudes. Basically, women tend to take less extreme attitudes in either direction. They are less likely to strongly disagree than are men, but they are also less likely to strongly agree. The ologit results imply gender has no effect while the gologit results say the effect of gender is highly significant. Perhaps the current coding of attitudes is not ordinal with respect to gender, e.g. coding by intensity of attitudes rather than direction may be more appropriate. Or, suppose that, instead of attitudes, the categories represented a set of ordered hurdles, e.g. achievement levels. Women as a whole may be more likely than men to clear the lowest hurdles but less likely to clear the highest ones. If men are more variable than women, they will have more outlying cases in both directions. Use of ologit in this case would be highly misleading.

- Every one of the above models represents a reasonable relationship involving an ordinal variable; but only the proportional odds model does not violate the assumptions of the ordered logit model
- FURTHER, there could be a dozen variables in a model, 11 of which meet the proportional odds assumption and only one of which does not
- We therefore want a more flexible and parsimonious model that can deal with situations like the above

Unconstrained gologit model

- Unconstrained gologit results are very similar to what we get with the series of binary logistic regressions and can be interpreted the same way.
- The gologit model can be written as

$$P(Y_i > j) = \frac{\exp(\alpha_j + X_i\beta_j)}{1 + [\exp(\alpha_j + X_i\beta_j)]}, \, j = 1, 2, ..., M - 1$$

- The ologit model is a special case of the gologit model, where the betas are the same for each j (NOTE: ologit actually reports cut points, which equal the negatives of the alphas used here)

$$P(Y_i > j) = \frac{\exp(\alpha_j + X_i\beta)}{1 + [\exp(\alpha_j + X_i\beta)]}, \text{j} = 1, 2, ..., \text{M} - 1$$

Partial Proportional Odds Model

- A key enhancement of gologit2 is that it allows some of the beta coefficients to be the same for all values of j, while others can differ. i.e. it can estimate partial proportional odds models. For example, in the following the betas for X1 and X2 are constrained but the betas for X3 are not.

$$P(Y_i > j) = \frac{\exp(\alpha_j + X1_i \beta1 + X2_i \beta2 + X3_i \beta3_j)}{1 + [\exp(\alpha_j + X1_i \beta1 + X2_i \beta2 + X3_i \beta3_j)]}, \, j = 1, 2, ..., M-1$$

- Either mlogit or unconstrained gologit can be overkill – both generate many more parameters than ologit does.
  - All variables are freed from the proportional odds constraint, even though the assumption may only be violated by one or a few of them

- gologit2, with the *autofit* option, will only relax the parallel lines constraint for those variables where it is violated

```
Generalized Ordered Logit Estimates              Number of obs   =        2293
                                                 LR chi2(10)     =      338.30
                                                 Prob > chi2     =      0.0000
Log likelihood = -2826.6182                      Pseudo R2       =      0.0565

 ( 1)   [1SD]white - [2D]white = 0
 ( 2)   [1SD]ed - [2D]ed = 0
 ( 3)   [1SD]prst - [2D]prst = 0
 ( 4)   [1SD]age - [2D]age = 0
 ( 5)   [2D]white - [3A]white = 0
 ( 6)   [2D]ed - [3A]ed = 0
 ( 7)   [2D]prst - [3A]prst = 0
 ( 8)   [2D]age - [3A]age = 0
------------------------------------------------------------------------------
      warm |     Coef.    Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
1SD          |
      yr89 |     .98368    .1530091     6.43   0.000     .6837876    1.283572
      male |  -.3328209    .1275129    -2.61   0.009    -.5827417   -.0829002
     white |  -.3832583    .1184635    -3.24   0.001    -.6154424   -.1510742
       age |  -.0216325    .0024751    -8.74   0.000    -.0264835   -.0167814
        ed |   .0670703    .0161311     4.16   0.000     .0354539    .0986866
      prst |   .0059146    .0033158     1.78   0.074    -.0005843    .0124135
      _cons |    2.12173    .2467146     8.60   0.000     1.638178    2.605282
-------------+----------------------------------------------------------------
2D           |
      yr89 |    .534369    .0913937     5.85   0.000     .3552406    .7134974
      male |  -.6932772    .0885898    -7.83   0.000    -.8669099   -.5196444
     white |  -.3832583    .1184635    -3.24   0.001    -.6154424   -.1510742
       age |  -.0216325    .0024751    -8.74   0.000    -.0264835   -.0167814
        ed |   .0670703    .0161311     4.16   0.000     .0354539    .0986866
      prst |   .0059146    .0033158     1.78   0.074    -.0005843    .0124135
      _cons |   .6021625    .2358361     2.55   0.011     .1399323    1.064393
-------------+----------------------------------------------------------------
3A           |
      yr89 |   .3258098    .1125481     2.89   0.004     .1052197       .5464
      male |  -1.097615    .1214597    -9.04   0.000    -1.335671   -.8595579
     white |  -.3832583    .1184635    -3.24   0.001    -.6154424   -.1510742
       age |  -.0216325    .0024751    -8.74   0.000    -.0264835   -.0167814
        ed |   .0670703    .0161311     4.16   0.000     .0354539    .0986866
      prst |   .0059146    .0033158     1.78   0.074    -.0005843    .0124135
      _cons |  -1.048137    .2393568    -4.38   0.000    -1.517268   -.5790061
------------------------------------------------------------------------------
```

Interpretation

- Once we have the results though, how do we interpret them???

- There are several possibilities.

Interpretation 1: gologit as non-linear probability model

- As Long & Freese (2006, p. 187) point out "The ordinal regression model can also be developed as a nonlinear probability model without appealing to the idea of a latent variable."

- Ergo, the simplest thing may just be to interpret gologit as a non-linear probability model that lets you estimate the determinants & probability of each outcome occurring. Forget about the idea of a y*

- Other interpretations, such as we have just discussed, can preserve or modify the idea of an underlying y*

Interpretation 2: The effect of x on y depends on the value of y

- Our earlier proportional odds examples show how this could plausibly be true

- Hedeker and Mermelstein (1998) also raise the idea that the categories of the DV may represent stages, e.g. pre-contemplation, contemplation, and action.

- An intervention might be effective in moving people from pre-contemplation to contemplation, but be ineffective in moving people from contemplation to action.

- If so, the effects of an explanatory variable will not be the same across the K-1 cumulative logits of the model

Working mother's example

- Effects of the constrained variables (white, age, ed, prst) can be interpreted pretty much the same as they were in the earlier ologit model. For yr89 and male, the differences from before are largely just a matter of degree.
  - People became more supportive of working mothers across time, but the greatest effect of time was to push people away from the most extremely negative attitudes.
  - For gender, men were less supportive of working mothers than were women, but they were especially unlikely to have strongly favorable attitudes.

- Substantive example: Boes & Winkelman, 2004:

  "Completely missing so far is any evidence whether the magnitude of the income effect depends on a person's happiness: is it possible that the effect of income on happiness is different in different parts of the outcome distribution? Could it be that "money cannot buy happiness, but buy-off unhappiness" as a proverb says? And if so, how can such distributional effects be quantified?"

Interpretation 3: State-dependent reporting bias - gologit as measurement model

- As noted, the idea behind y* is that there is an unobserved continuous variable that gets collapsed into the limited number of categories for the observed variable y.

- HOWEVER, respondents have to decide how that collapsing should be done, e.g. they have to decide whether their feelings cross the threshold between "agree" and "strongly agree," whether their health is "good" or "very good," etc.

- Respondents do NOT necessarily use the same frame of reference when answering, e.g. the elderly may use a different frame of reference than the young do when assessing their health

- Other factors can also cause respondents to employ different thresholds when describing things
  - Some groups may be more modest in describing their wealth, IQ or other characteristics

- In these cases the underlying latent variable may be the same for all groups; but the thresholds/cut points used may vary.
  - Example: an estimated gender effect could reflect differences in measurement across genders rather than a real gender effect on the outcome of interest.

- Lindeboom & Doorslaer (2004) note that this has been referred to as state-dependent reporting bias, scale of reference bias, response category cut-point shift, reporting heterogeneity & differential item functioning.

- If the difference in thresholds is constant (index shift), proportional odds will still hold
  - EX: Women's cutpoints are all a half point higher than the corresponding male cutpoints
  - ologit could be used in such cases

- If the difference is not constant (cut point shift), proportional odds will be violated
  - EX: Men and women might have the same thresholds at lower levels of pain but have different thresholds for higher levels
  - A gologit/ partial proportional odds model can capture this

- If you are confident that some apparent effects reflect differences in measurement rather than real differences in effects, then

  ▫ Cutpoints (and their determinants) are substantively interesting, rather than just "nuisance" parameters

  ▫ The idea of an underlying y* is preserved (Determinants of y* are the same for all, but cutpoints differ across individuals and groups)

- Key advantage: This could greatly improve cross-group comparisons, getting rid of artifactual differences caused by differences in measurement.

- Key Concern: Can you really be sure the coefficients reflect measurement and not real effects, or some combination of real & measurement effects?

- Theory may help – if your model strongly claims the effect of gender should be zero, then any observed effect of gender can be attributed to measurement differences.

- But regardless of what your theory says, you may at least want to acknowledge the possibility that apparent effects could be "real" or just measurement artifacts.

Interpretation 4: The outcome is multi-dimensional

- A variable that is ordinal in some respects may not be ordinal or else be differently-ordinal in others. E.g. variables could be ordered either by direction (Strongly disagree to Strongly Agree) or intensity (Indifferent to Feel Strongly)

- Suppose women tend to take less extreme political positions than men.
  - Using the first (directional) coding, an ordinal model might not work very well, whereas it could work well with the 2nd (intensity) coding.
  - But, suppose that for every other independent variable the directional coding works fine in an ordinal model.

- Our choices in the past have either been to (a) run ordered logit, with the model really not appropriate for the gender variable, or (b) run multinomial logit, ignoring the parsimony of the ordinal model just because one variable doesn't work with it.
- With gologit models, we have option (c) – constrain the vars where it works to meet the parallel lines assumption, while freeing up other vars (e.g. gender) from that constraint.

For more information, see:

https://www.nd.edu/~rwilliam/gologit2

http://www.stata-journal.com/article.html?article=st0097

https://www.tandfonline.com/doi/full/10.1080/0022250X.2015.1112384