

gologit2: Generalized Logistic Regression/ Partial Proportional Odds Models for Ordinal Dependent Variables

Richard Williams
Department of Sociology
University of Notre Dame
July 2005
<http://www.nd.edu/~rwilliam/>

Key features of gologit2

- Backwards compatible with Vincent Fu's original gologit program – but offers many more features
- Can estimate models that are less restrictive than ologit (whose assumptions are often violated)
- Can estimate models that are more parsimonious than non-ordinal alternatives, such as mlogit

Specifically, gologit2 can estimate:

- Proportional odds models (same as ologit – all variables meet the proportional odds/ parallel lines assumption)
- Generalized ordered logit models (same as the original gologit – no variables need to meet the parallel lines assumption)
- Partial Proportional Odds Models (some but not all variables meet the pl assumption)

Example 1: Proportional Odds Assumption Violated

- (Adapted from Long & Freese, 2003 – Data from the 1977 & 1989 General Social Survey)
- Respondents are asked to evaluate the following statement: “A working mother can establish just as warm and secure a relationship with her child as a mother who does not work.”
 - 1 = Strongly Disagree (SD)
 - 2 = Disagree (D)
 - 3 = Agree (A)
 - 4 = Strongly Agree (SA).

- Explanatory variables are
 - yr89 (survey year; 0 = 1977, 1 = 1989)
 - male (0 = female, 1 = male)
 - white (0 = nonwhite, 1 = white)
 - age (measured in years)
 - ed (years of education)
 - prst (occupational prestige scale).

Ologit results

```
. ologit warm yr89 male white age ed prst
```

Ordered logit estimates

Log likelihood = -2844.9123	Number of obs = 2293
	LR chi2(6) = 301.72
	Prob > chi2 = 0.0000
	Pseudo R2 = 0.0504

warm	Coef.	Std. Err.	z	P> z	[95% Conf. Interval]
yr89	.5239025	.0798988	6.56	0.000	.3673037 .6805013
male	-.7332997	.0784827	-9.34	0.000	-.8871229 -.5794766
white	-.3911595	.1183808	-3.30	0.001	-.6231815 -.1591374
age	-.0216655	.0024683	-8.78	0.000	-.0265032 -.0168278
ed	.0671728	.015975	4.20	0.000	.0358624 .0984831
prst	.0060727	.0032929	1.84	0.065	-.0003813 .0125267

_cut1	_cut2	_cut3
-2.465362	-.630904	1.261854
.2389126	.2333155	.2340179

(Ancillary parameters)

Interpretation of ologit results

- These results are relatively straightforward, intuitive and easy to interpret. People tended to be more supportive of working mothers in 1989 than in 1977. Males, whites and older people tended to be less supportive of working mothers, while better educated people and people with higher occupational prestige were more supportive.
- But, while the results may be straightforward, intuitive, and easy to interpret, are they correct? Are the assumptions of the ologit model met? The following Brant test suggests they are not.

Brant test shows assumptions violated

```
. brant
Brant Test of Parallel Regression Assumption
Variable |      chi2  p>chi2  df
-----+-----
All |      49.18   0.000   12
-----+-----
yr89 |      13.01   0.001    2
male |      22.24   0.000    2
white |       1.27   0.531    2
age |       7.38   0.025    2
ed |       4.31   0.116    2
prst |       4.33   0.115    2
-----+-----
```

A significant test statistic provides evidence that the parallel regression assumption has been violated.

How are the assumptions violated?

• brant, detail

Estimated coefficients from j-1 binary regressions

	y>1	y>2	y>3
yr89	.9647422	.56540626	.31907316
male	-.30536425	-.69054232	-1.0837888
white	-.55265759	-.31427081	-.39299842
age	-.0164704	-.02533448	-.01859051
ed	.10479624	.05285265	.05755466
prst	-.00141118	.00953216	.00553043
_cons	1.8584045	.73032873	-1.0245168

- This is a series of binary logistic regressions. First it is 1 versus 2,3,4; then 1 & 2 versus 3 & 4; then 1, 2, 3 versus 4
- If proportional odds/ parallel lines assumptions were not violated, all of these coefficients (except the intercepts) would be the same except for sampling variability.

Dealing with violations of assumptions

- Just ignore it! (A fairly common practice)
- Go with a non-ordinal alternative, such as mlogit
- Go with an ordinal alternative, such as the original gologit & the default gologit2
- Try an in-between approach: partial proportional odds

Interpretation of the gologit/gologit2 model

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

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

- Note that the logit model is a special case of the gologit model, where $M = 2$. When $M > 2$, you get a series of binary logistic regressions, e.g. 1 versus 2, 3 4, then 1, 2 versus 3, 4, then 1, 2, 3 versus 4.
- The ologit model is also 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)

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

- 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.

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

gologit2/ partial proportional odds

- Either mlogit or the original 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

gologit2 with autofit

```
. gologit2 warm yr89 male white age ed prst, auto lforce
```

```
-----  
Testing parallel lines assumption using the .05 level of significance...
```

```
Step 1: white meets the pl assumption (P Value = 0.7136)  
Step 2: ed meets the pl assumption (P Value = 0.1589)  
Step 3: prst meets the pl assumption (P Value = 0.2046)  
Step 4: age meets the pl assumption (P Value = 0.0743)  
Step 5: The following variables do not meet the pl assumption:  
       yr89 (P Value = 0.00093)  
       male (P Value = 0.00002)
```

If you re-estimate this exact same model with gologit2, instead of autofit you can save time by using the parameter

```
pl(white ed prst age)
```

- gologit2 is going through a stepwise process here. Initially no variables are constrained to have proportional effects. Then Wald tests are done. Variables which pass the tests (i.e. variables whose effects do not significantly differ across equations) have proportionality constraints imposed.

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

```
( 1) [SD]white - [D]white = 0  
( 2) [SD]ed - [D]ed = 0  
( 3) [SD]prst - [D]prst = 0  
( 4) [SD]age - [D]age = 0  
( 5) [D]white - [A]white = 0  
( 6) [D]ed - [A]ed = 0  
( 7) [D]prst - [A]prst = 0  
( 8) [D]age - [A]age = 0
```

- Internally, gologit2 is generating several constraints on the parameters. The variables listed above are being constrained to have their effects meet the proportional odds/ parallel lines assumptions
- Note: with ologit, there were 6 degrees of freedom; with gologit & mlogit there were 18; and with gologit2 using autofit there are 10. The 8 d.f. difference is due to the 8 constraints above.

		Coef.	Std. Err.	z	P> z	[95% Conf. Interval]
SD						
warm						
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
D						
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
A						
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

- At first glance, it appears there are just as many parameters as before – but 8 of them are duplicates because of the proportionality constraints that have been imposed.

Interpretation of the gologit2 results

- 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 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.

Example 2: Alternative Gamma Parameterization

- Peterson & Harrell (1990) presented an equivalent parameterization of the gologit model, called the *Unconstrained Partial Proportional Odds Model*.
- Under the Peterson/Harrell parameterization, each explanatory variable has
 - One Beta coefficient
 - $M - 2$ Gamma coefficients, where M = the # of categories in the Y variable and the Gammas represent deviations from proportionality

-
- The difference between the gologit/ default gologit2 parameterization and the alternative parameterization is similar to the difference between running separate models for each group as opposed to having a single model with interaction terms.
 - The *gamma* option of gologit2 (abbreviated *g*) presents this parameterization

```

. gologit2 warm yr89 male white age ed prst, autofit lrforce gamma

Alternative parameterization: Gammas are deviations from proportionality
-----
      warm |      Coef.   Std. Err.      z    P>|z|      [95% Conf. Interval]
-----+-----
Beta
   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
-----+-----
Gamma_2
   yr89 |     -.4493111 .1465627    -3.07   0.002  -.7365686  -.1620533
   male |     -.3604562 .1233732    -2.92   0.003  -.6022633  -.1186492
-----+-----
Gamma_3
   yr89 |     -.6578702 .1768034    -3.72   0.000  -1.004399  -.3113418
   male |     -.7647937 .1631536    -4.69   0.000  -1.084569  -.4450186
-----+-----
Alpha
  _cons_1 |      2.12173   .2467146     8.60   0.000   1.638178   2.605282
  _cons_2 |      .6021625   .2358361     2.55   0.011   .1399323   1.064393
  _cons_3 |     -1.048137   .2393568    -4.38   0.000  -1.517268  -.5790061
-----

```

Advantages of the Gamma Parameterization

- ❑ Consistent with other published research
- ❑ More parsimonious layout – you don't keep seeing the same parameters that have been constrained to be equal
- ❑ Alternative way of understanding the proportionality assumption – if the Gammas for a variable all equal 0, the assumption is met for that variable, and if all the Gammas equal 0 you have the ologit model
- ❑ By examining the Gammas you can better pinpoint where assumptions are being violated

Example 3: Imposing and testing constraints

- Rather than use *autofit*, you can use the *pl* and *npl* parameters to specify which variables are or are not constrained to meet the proportional odds/ parallel lines assumption
 - Gives you more control over model specification & testing
 - Lets you use LR chi-square tests rather than Wald tests
 - Could use BIC or AIC tests rather than chi-square tests if you wanted to when deciding on constraints
 - *pl* without parameters will produce same results as *ologit*

-
- Other types of linear constraints can also be specified, e.g. you can constrain two variables to have equal effects (neither *ologit* nor *logit* currently allow this, so if you want to impose constraints on these models you could use *gologit2* instead)
 - The *store* option will cause the command *estimates store* to be run at the end of the job, making it slightly easier to do LR chi-square contrasts
 - Here is how we could do tests to see if we agree with the model produced by *autofit*:

LR chi-square contrasts using gologit2

```
. * Least constrained model - same as the original gologit
. quietly gologit2 warm yr89 male white age ed prst, store(gologit)

. * Partial Proportional Odds Model, estimated using autofit
. quietly gologit2 warm yr89 male white age ed prst, store(gologit2) autofit

. * Ologit clone
. quietly gologit2 warm yr89 male white age ed prst, store(ologit) pl

. * Confirm that ologit is too restrictive
. lrtest ologit gologit

Likelihood-ratio test                LR chi2(12) =    49.20
(Assumption: ologit nested in gologit)  Prob > chi2 =    0.0000

. * Confirm that partial proportional odds is not too restrictive
. lrtest gologit gologit2

Likelihood-ratio test                LR chi2(8) =    12.61
(Assumption: gologit2 nested in gologit)  Prob > chi2 =    0.1258
```

Example 4: Substantive significance of gologit2

- gologit2 may be “better” than ologit – but substantively, how much should we care?
 - ologit assumptions are often violated
 - Substantively, those violations may not be that important – but you can’t know that without doing formal tests
 - Violations of assumptions can be substantively important. The earlier example showed that the effects of gender and time were not uniform. Also, ologit may hide or obscure important relationships. e.g. using nhanes2f.dta,

	health	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]	
poor							
	female	.1212723	.0975363	1.24	0.223	-.0776543	.3201989
	_cons	2.940598	.0957485	30.71	0.000	2.745317	3.135878
fair							
	female	-.1833293	.0640565	-2.86	0.007	-.3139733	-.0526852
	_cons	1.682043	.058651	28.68	0.000	1.562424	1.801663
average							
	female	-.1772901	.0545539	-3.25	0.003	-.2885535	-.0660268
	_cons	.2938385	.0402766	7.30	0.000	.2116939	.3759831
good							
	female	-.2356111	.05914	-3.98	0.000	-.356228	-.1149943
	_cons	-.8493609	.0382026	-22.23	0.000	-.9272756	-.7714461

- Females are less likely to report poor health than are males (see the positive female coefficient in the poor panel), but they are also less likely to report higher levels of health (see the negative female coefficients in the other panels), i.e. women tend to be less at the extremes of health than men are. Such a pattern would be obscured in a straight proportional odds (ologit) model.

Other gologit2 features of interest

- The predict command can easily compute predicted probabilities
- Stata 8.2 survey data estimation is possible when the *svy* option is used. Several *svy*-related options, such as *subpop*, are supported

-
- The *v1* option causes *gologit2* to return results in a format that is consistent with *gologit* 1.0.
 - This may be useful/necessary for post-estimation commands that were written specifically for *gologit* (in particular, the Long and Freese *spost* commands currently support *gologit* but not *gologit2*).
 - In the long run, post-estimation commands should be easier to write for *gologit2* than they were for *gologit*.

-
- The *lforce* option causes Stata to report a Likelihood Ratio Statistic under certain conditions when it ordinarily would report a Wald statistic. Stata is being cautious but I think LR statistics are appropriate for most common *gologit2* models
 - *gologit2* uses an unconventional but seemingly-effective way to label the model equations. If problems occur, the *nolabel* option can be used.
 - Most other standard options (e.g. *robust*, *cluster*, *level*) are supported.