Ordinal regression models: Problems, solutions, and problems with the solutions

German Stata User Group Meetings, June 27, 2008

Richard Williams, Notre Dame Sociology, rwilliam@ND.Edu
oglm support page: http://www.nd.edu/~rwilliam/oglm
gologit2 support page: http://www.nd.edu/~rwilliam/gologit2

Ordered logit/probit models are among the most popular ordinal regression techniques. These models often have serious problems, however. The proportional odds/parallel lines assumptions made by these methods are often violated. Further, because of the way these models are identified, they have many of the same limitations as are encountered when analyzing standardized coefficients in OLS regression, e.g. interaction terms and cross-population comparisons of effects can be highly misleading. This paper shows how generalized ordered logit/probit models (estimated via `gologit2`) and heterogeneous choice/location scale models (estimated via `oglm`) can often address these concerns in ways that are more parsimonious and easier to interpret than is the case with other suggested alternatives. At the same time, the paper cautions that these methods sometimes raise their own concerns that researchers need to be aware of and know how to deal with. First, misspecified models can create worse problems than the ones these methods were designed to solve. Second, estimates are sometimes implausible, suggesting that the data are being spread too thin and/or yet another method is needed. Third, multiple and very different interpretations of the same results are sometimes possible and plausible. Guidelines for identifying and dealing with each of these problems are presented.

Problem I: Heteroskedastic errors

Allison's example: Apparent differences in effects across groups may be an artifact of differences in residual variability

Table 1: Results of Logit Regressions Predicting Promotion to Associate Professor for Male and Female Biochemists

	Men		*Women*		*Ratio of*	*Chi-Square*
Variable	*Coefficient*	*SE*	*Coefficient*	*SE*	*Coefficients*	*for Difference*
Intercept	-7.6802***	.6814	-5.8420***	.8659	.76	2.78
Duration	1.9089***	.2141	1.4078***	.2573	.74	2.24
Duration squared	-0.1432***	.0186	-0.0956***	.0219	.67	2.74
Undergraduate selectivity	0.2158***	.0614	0.0551	.0717	.25	2.90
Number of articles	*0.0737****	*.0116*	*0.0340***	*.0126*	*.46*	*5.37**
Job prestige	-0.4312***	.1088	-0.3708*	.1560	.86	0.10
Log likelihood	-526.54		-306.19			

p* < .05, ***p* < .01, * *p* < .001

Reprinted from Allison (1999, p. 188)

Allison's solution: Add delta to adjust for differences in residual variability

Table 2: Logit Regressions Predicting Promotion to Associate Professor for Male and Female Biochemists, Disturbance Variances Unconstrained

	All Coefficients Equal		*Articles Coefficient Unconstrained*	
Variable	*Coefficient*	*SE*	*Coefficient*	*SE*
Intercept	-7.4913***	.6845	-7.3655***	.6818
Female	-0.93918**	.3624	-0.37819	.4833
Duration	1.9097***	.2147	1.8384***	.2143
Duration squared	-0.13970***	.0173	-0.13429***	.01749
Undergraduate selectivity	0.18195**	.0615	0.16997***	.04959
Number of articles	0.06354***	.0117	0.07199***	.01079
Job prestige	-0.4460***	.1098	-0.42046***	.09007
δ	*-0.26084**	*.1116*	*-0.16262*	*.1505*
Articles x Female			*-0.03064*	*.0173*
Log likelihood	-836.28		-835.13	

$*p < .05$, $**p < .01$, $*** p < .001$

Reprinted from Allison (1999, p. 195)

Alternative (and broader) solution: Heterogeneous Choice Models

With heterogeneous choice (aka Location-Scale) models, the dependent variable can be ordinal or binary. For a binary dependent variable, the model (Keele & Park, 2006) can be written as

$$\Pr(y_i = 1) = g\left(\frac{x_i\beta}{\exp(z_i\gamma)}\right) = g\left(\frac{x_i\beta}{\exp(\ln(\sigma_i))}\right) = g\left(\frac{x_i\beta}{\sigma_i}\right)$$

In the above formula,

- g stands for the link function (in this case logit; probit is also commonly used, and other options are possible, such as the complementary log-log, log-log and cauchit).
- x is a vector of values for the ith observation. The x's are the explanatory variables and are said to be the determinants of the choice, or outcome.
- z is a vector of values for the ith observation. The z's define groups with different error variances in the underlying latent variable. The z's and x's need not include any of the same variables, although they can.
- β and γ are vectors of coefficients. They show how the x's affect the choice and the z's affect the variance (or more specifically, the log of σ).
- The numerator in the above formula is referred to as the choice equation, while the denominator is the variance equation. These are also referred to as the location and scale equations. Also, the choice equation includes a constant term but the variance equation does not.
- The conventional logit and probit models, which do not have variance equations, are special cases of the above, where $\sigma_i = 1$ for all cases.
- Allison's model is a special case of a heterogeneous choice model, where the dependent variable is a dichotomy and both the variance and choice equations include the same dichotomous grouping variable.

In Stata, heterogeneous choice models can be estimated via the user-written routine `oglm`.

```
. * oglm replication of Allison's Table 2:
. use "http://www.indiana.edu/~jslsoc/stata/spex_data/tenure01.dta", clear
(Gender differences in receipt of tenure (Scott Long 06Jul2006))
. keep if pdasample
(148 observations deleted)
. * Allison Table 2, Model 1

. oglm tenure female year yearsq select articles prestige, het(female) store(m1)

Heteroskedastic Ordered Logistic Regression       Number of obs   =       2797
                                                  LR chi2(7)      =     413.09
                                                  Prob > chi2     =     0.0000
Log likelihood = -836.28235                       Pseudo R2       =     0.1981

------------------------------------------------------------------------------
             |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
tenure       |
      female |  -.9391907   .3705243    -2.53   0.011    -1.665405   -.2129763
        year |   1.909544   .1996935     9.56   0.000     1.518152    2.300936
      yearsq |  -.1396868   .0169425    -8.24   0.000    -.1728935   -.1064801
      select |   .1819201   .0526572     3.45   0.001     .0787139    .2851264
    articles |   .0635345    .010219     6.22   0.000     .0435055    .0835635
    prestige |  -.4462073    .096904    -4.60   0.000    -.6361356   -.2562791
-------------+----------------------------------------------------------------
lnsigma      |
      female |   .3022305    .146178     2.07   0.039     .0157268    .5887341
-------------+----------------------------------------------------------------
       /cut1 |   7.490506   .6596628    11.36   0.000      6.19759    8.783421
------------------------------------------------------------------------------
. display "Allison's delta = " (1 - exp(.3022305)) / exp(.3022305)
Allison's delta = -.26083233
```

```
. * Allison Table 2, Model 2 with interaction added
. oglm tenure female year yearsq select articles prestige f_articles, het(female) store(m2)

Heteroskedastic Ordered Logistic Regression       Number of obs   =       2797
                                                  LR chi2(8)      =     415.39
                                                  Prob > chi2     =     0.0000
Log likelihood = -835.13347                       Pseudo R2       =     0.1992

------------------------------------------------------------------------------
             |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
tenure       |
      female |  -.3780597   .4500207    -0.84   0.401    -1.260084    .5039646
        year |   1.838257   .2029491     9.06   0.000     1.440484     2.23603
      yearsq |  -.1342828    .017024    -7.89   0.000    -.1676492   -.1009165
      select |   .1699659   .0516643     3.29   0.001     .0687057    .2712261
    articles |   .0719821   .0114106     6.31   0.000     .0496178    .0943464
    prestige |  -.4204742   .0961206    -4.37   0.000    -.6088671   -.2320813
  f_articles |  -.0304836   .0187427    -1.63   0.104    -.0672185    .0062514
-------------+----------------------------------------------------------------
lnsigma      |
      female |   .1774193   .1627087     1.09   0.276     -.141484    .4963226
-------------+----------------------------------------------------------------
       /cut1 |   7.365285   .6547121    11.25   0.000     6.082073    8.648497
------------------------------------------------------------------------------

. display "Allison's delta = " (1 - exp(.1774193)) / exp(.1774193)
Allison's delta = -.16257142
```

```
. * Test interaction term.  For the choice equation, LR tests are usually
. * preferable to Wald tests.  E.g. if you used male instead of female
. * in the above models the Wald tests would come out differently but the
. * lr tests would come out the same.  The choice coefficients are the coefficients
. * for a group that has values of 0 on all vars in the variance equation.
. lrtest m1 m2, stats

Likelihood-ratio test                                 LR chi2(1)  =      2.30
(Assumption: m1 nested in m2)                         Prob > chi2 =    0.1296

-----------------------------------------------------------------------------
       Model |    Obs    ll(null)   ll(model)     df          AIC         BIC
-------------+---------------------------------------------------------------
          m1 |   2797   -1042.828   -836.2824      8     1688.565    1736.055
          m2 |   2797   -1042.828   -835.1335      9     1688.267    1741.694
-----------------------------------------------------------------------------
               Note:  N=Obs used in calculating BIC; see [R] BIC note
```

Using Stepwise selection as a model building or diagnostic device

```
. sw, pe(.01) lr: oglm  tenure female year yearsq select  articles prestige,
eq2(female  year yearsq select  articles prestige ) flip store(m3)
LR test               begin with empty model
p = 0.0000 <  0.0100  adding  articles

Heteroskedastic Ordered Logistic Regression       Number of obs   =       2797
                                                  LR chi2(7)      =     428.03
                                                  Prob > chi2     =     0.0000
Log likelihood = -828.81224                       Pseudo R2       =     0.2052

------------------------------------------------------------------------------
             |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
tenure       |
      female |  -.4179259   .1742083    -2.40   0.016     -.759368   -.0764838
        year |   2.108752   .2486633     8.48   0.000     1.621381    2.596123
      yearsq |  -.1542213   .0208579    -7.39   0.000    -.1951019   -.1133406
      select |   .1744644   .0598623     2.91   0.004     .0571364    .2917924
    articles |   .0628407   .0157851     3.98   0.000     .0319026    .0937789
    prestige |  -.6118689   .1307262    -4.68   0.000    -.8680877   -.3556502
-------------+----------------------------------------------------------------
lnsigma      |
    articles |    .030149   .0091448     3.30   0.001     .0122256    .0480724
-------------+----------------------------------------------------------------
       /cut1 |   7.959556   .7637106    10.42   0.000      6.46271    9.456401
------------------------------------------------------------------------------

. * Another alternative.  General idea suggested by Maarten Buis.
. * articles is the problem, so find another way to deal with it.
. gen articles2 = articles^2
. oglm  tenure female year yearsq select   articles articles2 prestige, het(articles) store(m4)

Heteroskedastic Ordered Logistic Regression       Number of obs   =       2797
                                                  LR chi2(8)      =     439.77
                                                  Prob > chi2     =     0.0000
Log likelihood = -822.94311                       Pseudo R2       =     0.2109

------------------------------------------------------------------------------
             |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
tenure       |
```

```
      female |  -.3470778   .1470054    -2.36   0.018    -.6352031   -.0589526
        year |   1.764339   .2233366     7.90   0.000     1.326608    2.202071
      yearsq |  -.1282567   .0182644    -7.02   0.000    -.1640544   -.0924591
      select |   .1631087   .0503776     3.24   0.001     .0643704    .2618471
    articles |   .1481165   .0246791     6.00   0.000     .0997464    .1964866
   articles2 |   -.002716   .0008273    -3.28   0.001    -.0043374   -.0010945
    prestige |  -.4909742   .1124811    -4.36   0.000    -.7114332   -.2705152
-------------+----------------------------------------------------------------
lnsigma      |
    articles |   .0081942   .0095091     0.86   0.389    -.0104432    .0268316
-------------+----------------------------------------------------------------
       /cut1 |   7.375548   .6803437    10.84   0.000     6.042099    8.708997
------------------------------------------------------------------------------

. lrtest m3 m4, stats

Likelihood-ratio test                                  LR chi2(1)  =     11.74
(Assumption: m3 nested in m4)                          Prob > chi2 =    0.0006

-----------------------------------------------------------------------------
       Model |    Obs    ll(null)   ll(model)     df          AIC         BIC
-------------+---------------------------------------------------------------
          m3 |   2797   -1042.828   -828.8122      8     1673.624    1721.115
          m4 |   2797   -1042.828   -822.9431      9     1663.886    1717.313
-----------------------------------------------------------------------------
               Note:  N=Obs used in calculating BIC; see [R] BIC note
```

Problem with the Solution I: Model misspecification can have serious consequences

*Simulations where residual variances are equal across groups but the coefficients are not**					
αs: $\alpha_1^0 = \alpha_2^0 = 1$ $\alpha_2^1 = 2$ α_1^1 varies	Test of residual variances differ across groups, while αs are assumed to be the same		% of time LR test correctly rejects hyp of equal coefficients across groups	Effect of X2 allowed to differ across groups	
	Average estimated value of δ	% of times LR test falsely rejects hyp of equal residual variances		Average estimated value of δ	Average estimated value of X2 interaction term
$\alpha_1^1 = 0.50$	0.591	82.4%	99.9%	-0.491	3.346
$\alpha_1^1 = 1.00$	0.649	92.3%	90.7%	0.016	1.063
$\alpha_1^1 = 1.50$	0.802	98.4%	35.5%	0.522	0.359
$\alpha_1^1 = 2.0$	1.023	100.0%	5.1%	1.029	0.012
$\alpha_1^1 = 2.50$	1.303	100.0%	21.5%	1.539	-0.195
$\alpha_1^1 = 3.00$	1.631	100.0%	59.8%	2.054	-0.333

* By construction, in every simulation the true value of δ is 0, the hypothesis of equal residual variances is true, the hypothesis of equal coefficients is false, and the true value of the X2 interaction term is 1.

Problem with the Solution II: Radically different interpretations of the same results are possible.

Example: Hauser & Andrew's (Sociological Methodology 2006) Logistic Response Model with Partial Proportionality Constraints.

Hauser and Andrew replicated and extended Mare's analysis of school continuation. They argued that the relative effects of some (but not all) background variables are the same at each transition, and that multiplicative scalars express proportional change in the effect of those variables across successive transitions. Specifically, Hauser & Andrew estimate two new types of models.

logistic response model with proportionality constraints (LRPC):	*logistic response model with partial proportionality constraints* (LRPPC):
$\log_e\left(\frac{p_{ij}}{1-p_{ij}}\right) = \beta_{j0} + \lambda_j \sum_k \beta_k X_{ijk}$	$\log_e\left(\frac{p_{ij}}{1-p_{ij}}\right) = \beta_{j0} + \lambda_j \sum_{k=1}^{k'} \beta_k X_{ijk} + \sum_{k'+1}^{K} \beta_{jk} X_{ijk}$

Hauser & Andrew summarize their models in Table 5 of their paper:

TABLE 5
Fit of Selected Models of Educational Transitions: 1973 Occupational Changes in a Generation Survey

Model	Description	Log-Likelihood	DF for Model	Model Chi-square	Contrast	Contrast Chi-square	Contrast BIC	Pseudo R-squared
1	Fit the grand mean	−46830.8	0	—		—		0
2	An intercept for each transition	−38674.3	5	16313.0	2 vs. 1	16313.0	16256.0	0.17
3	An intercept for each transition and constant social background effects	−34333.3	13	24995.0	3 vs. 2	8682.0	8590.8	0.27
4	An intercept for each transition and proportional social background effects	−33529.7	19	26602.2	4 vs. 3	1607.3	1538.9	0.28
5	An intercept for each transition, constant effects of socioeconomic variables, interactions of BROKEN, FARM, and SOUTH with transition	−34112.0	28	25437.6	5 vs. 3	442.6	271.7	0.27
6	An intercept for each transition, proportional effects of socioeconomic variables, interactions of BROKEN, FARM, and SOUTH with transition	−33399.7	34	26862.1	6 vs. 5	1424.6	1356.2	0.29
7	Saturated model: Intercepts for each transition and interactions of all social background variables with transition	−33332.2	53	26997.2	7 vs. 6	135.1	−81.4	0.29

Here are `oglm`'s algebraically-equivalent models. Note that the fits are identical to those reported by Hauser and Andrew. Nonetheless, the interpretations are very different. Hauser and Andrew's models argue that there are real differences in effects across transitions, whereas the heterogeneous choice models imply that the apparent differences in effect are an artifact of differences in residual variability.

```
----------------------------------------------------------------------------------------------------
                    m1           m2           m3           m4           m5           m6           m7
----------------------------------------------------------------------------------------------------
N                88768        88768        88768        88768        88768        88768        88768
ll            -46830.8     -38674.3     -34333.3     -33529.7     -34112.0     -33399.7     -33332.2
df_m                 0            5           13           18           28           33           53
chi2          5.82e-11      16313.0      24995.0      26602.2      25437.6      26862.1      26997.2
r2_p          6.66e-16        0.174        0.267        0.284        0.272        0.287        0.288
----------------------------------------------------------------------------------------------------
```

Five of the Hauser & Andrew models can be estimated via conventional logistic regression. Model 4 (LRPC) and Model 6 (LRPPC) can be estimated via Stata code they present in their paper. Following is the `oglm` code for estimating models that are algebraically equivalent to m4 and m6. In both m4 and m6, dummy variables for transition are included in the variance equation. In m6, the non-ses variables are freed from constraints by including interaction terms for each non-ses variable with each transition.

```
*** Model 4: An intercept for each transition & proportional social background effects
* This is the first hetero choice model (equivalent to H & A's LRPC).
quietly oglm outcome trans2 trans3 trans4 trans5 trans6 dunc sibsttl9 ln_inc_trunc
edhifaom edhimoom broken farm16 south, het(trans2 trans3 trans4 trans5 trans6)
store(m4)

*** Model 6: An intercept for each transition, proportional effects of
* socioeconomic variables, interactions of broken, farm, and south with transition.
* This is the second hetero choice model (equivalent to H & A's LRPPC).
quietly oglm outcome trans2 trans3 trans4 trans5 trans6 broken farm16 south
trans2Xbroken trans2Xfarm16 trans2Xsouth trans3Xbroken trans3Xfarm16 trans3Xsouth
trans4Xbroken trans4Xfarm16 trans4Xsouth trans5Xbroken trans5Xfarm16 trans5Xsouth
trans6Xbroken trans6Xfarm16 trans6Xsouth dunc sibsttl9 ln_inc_trunc edhifaom edhimoom,
het(trans2 trans3 trans4 trans5 trans6) store(m6)
```

Problem 2: Parallel Lines/ Proportional odds assumption violated

Illustration of the problem: Working Mothers Example

```
. use "http://www.indiana.edu/~jslsoc/stata/spex_data/ordwarm2.dta"
(77 & 89 General Social Survey)
. * Parallel Lines/ Proportional Odds Model
. ologit  warm yr89 male white age ed prst, nolog

Ordered logistic regression                       Number of obs   =       2293
                                                  LR chi2(6)      =     301.72
                                                  Prob > chi2     =     0.0000
Log likelihood = -2844.9123                       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 |  -2.465362   .2389126                     -2.933622   -1.997102
       /cut2 |   -.630904   .2333155                     -1.088194    -.173614
       /cut3 |   1.261854   .2340179                      .8031873    1.720521
------------------------------------------------------------------------------
```

```
. est store ologit
.* Brant test shows assumptions are 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

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.
```

A non-parsimonious solution to the problem: Unconstrained Generalized Ordered Logit Model

Unconstrained Gologit Model. All betas are free to differ across levels of j.

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

```
. * Unconstrained gologit model - no vars required to meet parallel lines
. * Results are almost identical to running j-1 binary regressions,
. * like the Brant test reported.
. gologit2  warm yr89 male white age ed prst, npl lrf store(gologit)

Generalized Ordered Logit Estimates               Number of obs   =       2293
                                                  LR chi2(18)     =     350.92
                                                  Prob > chi2     =     0.0000
Log likelihood =  -2820.311                       Pseudo R2       =     0.0586

------------------------------------------------------------------------------
        warm |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
1SD          |
        yr89 |     .95575   .1547185     6.18   0.000     .6525074    1.258993
        male |  -.3009776   .1287712    -2.34   0.019    -.5533645   -.0485906
       white |  -.5287268   .2278446    -2.32   0.020    -.9752941   -.0821595
         age |  -.0163486   .0039508    -4.14   0.000    -.0240921   -.0086051
          ed |   .1032469   .0247377     4.17   0.000     .0547619     .151732
```

```
        prst |  -.0016912   .0055997    -0.30   0.763    -.0126665     .009284
       _cons |   1.856951   .3872576     4.80   0.000      1.09794    2.615962
-------------+----------------------------------------------------------------
2D           |
        yr89 |   .5363707   .0919074     5.84   0.000     .3562355     .716506
        male |   -.717995   .0894852    -8.02   0.000    -.8933827   -.5426072
       white |   -.349234   .1391882    -2.51   0.012    -.6220379     -.07643
         age |  -.0249764   .0028053    -8.90   0.000    -.0304747   -.0194782
          ed |   .0558691   .0183654     3.04   0.002     .0198737    .0918646
        prst |   .0098476   .0038216     2.58   0.010     .0023575    .0173377
       _cons |   .7198119    .265235     2.71   0.007     .1999609    1.239663
-------------+----------------------------------------------------------------
3A           |
        yr89 |   .3312184   .1127882     2.94   0.003     .1101577    .5522792
        male |  -1.085618   .1217755    -8.91   0.000    -1.324294   -.8469423
       white |  -.3775375   .1568429    -2.41   0.016     -.684944    -.070131
         age |  -.0186902   .0037291    -5.01   0.000     -.025999   -.0113814
          ed |   .0566852   .0251836     2.25   0.024     .0073263    .1060441
        prst |   .0049225   .0048543     1.01   0.311    -.0045918    .0144368
       _cons |  -1.002225   .3446354    -2.91   0.004    -1.677698   -.3267523
------------------------------------------------------------------------------
```

A More Parsimonious Solution: Partial Proportional Odds

Constrained Gologit Model – Partial Proportional Odds. Some betas differ across levels of j but others do not.

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

```
. * Partial proportional odds - relax the pl assumption when it is violated
. gologit2  warm yr89 male white age ed prst, auto lrf store(gologit2)

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

Step  1:  Constraints for parallel lines imposed for white (P Value = 0.7136)
Step  2:  Constraints for parallel lines imposed for ed (P Value = 0.1589)
Step  3:  Constraints for parallel lines imposed for prst (P Value = 0.2046)
Step  4:  Constraints for parallel lines imposed for age (P Value = 0.0743)
Step  5:  Constraints for parallel lines are not imposed for
          yr89 (P Value = 0.00093)
          male (P Value = 0.00002)

Wald test of parallel lines assumption for the final model:

 ( 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)  [1SD]white - [3A]white = 0
 ( 6)  [1SD]ed - [3A]ed = 0
 ( 7)  [1SD]prst - [3A]prst = 0
 ( 8)  [1SD]age - [3A]age = 0

           chi2(  8) =   12.80
         Prob > chi2 =    0.1190

An insignificant test statistic indicates that the final model
does not violate the proportional odds/ parallel lines assumption
```

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)

```
------------------------------------------------------------------------------

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
------------------------------------------------------------------------------
```

```
. * lrtests show that partial proportional odds is the most parsimonious model
. lrtest ologit gologit, force

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

. lrtest ologit gologit2, force

Likelihood-ratio test                                 LR chi2(4)  =     36.59
(Assumption: ologit nested in gologit2)               Prob > chi2 =    0.0000
```

```
. lrtest gologit gologit2, force

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

Concerns 1 & 2: Ordinality not required; predicted probabilities can go negative

```
. recode warm (1=3)(3=1), gen(xwarm)
(1153 differences between warm and xwarm)

. gologit2  xwarm yr89 male white age ed prst

Generalized Ordered Logit Estimates               Number of obs   =       2293
                                                  LR chi2(18)     =     351.13
                                                  Prob > chi2     =     0.0000
Log likelihood = -2820.2051                       Pseudo R2       =     0.0586

------------------------------------------------------------------------------
       xwarm |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
1            |
        yr89 |  -.3279931   .0895688    -3.66   0.000    -.5035447   -.1524415
        male |   .0985331   .0878095     1.12   0.262    -.0735703    .2706364
       white |   .0775744   .1325895     0.59   0.558    -.1822962    .3374451
         age |   .0148708   .0027818     5.35   0.000     .0094186    .0203229
          ed |  -.0341937   .0184685    -1.85   0.064    -.0703914    .0020039
        prst |  -.0050614   .0037438    -1.35   0.176    -.0123992    .0022764
       _cons |    .497562   .2618536     1.90   0.057    -.0156617    1.010786
-------------+----------------------------------------------------------------
2            |
        yr89 |  -.2527107   .0954108    -2.65   0.008    -.4397124   -.0657089
        male |  -.5372284   .0924572    -5.81   0.000    -.7184412   -.3560156
       white |  -.0387025   .1359797    -0.28   0.776    -.3052179    .2278128
         age |  -.0030129   .0028531    -1.06   0.291    -.0086049    .0025791
          ed |  -.0381041   .0188705    -2.02   0.043    -.0750895   -.0011186
        prst |   .0078674   .0038637     2.04   0.042     .0002948      .01544
       _cons |  -.1399591   .2710817    -0.52   0.606    -.6712695    .3913512
-------------+----------------------------------------------------------------
3            |
        yr89 |   .2502576   .1071648     2.34   0.020     .0402185    .4602966
        male |  -.9449406   .1143625    -8.26   0.000    -1.169087   -.7207942
       white |  -.4347512   .1472539    -2.95   0.003    -.7233635   -.1461389
         age |  -.0167564   .0033158    -5.05   0.000    -.0232554   -.0102575
          ed |   .0571525   .0230149     2.48   0.013     .0120442    .1022608
        prst |   .0061237   .0042714     1.43   0.152     -.002248    .0144954
       _cons |  -1.108264   .3067563    -3.61   0.000    -1.709495   -.5070325
------------------------------------------------------------------------------

WARNING! 133 in-sample cases have an outcome with a predicted probability that is
less than 0. See the gologit2 help section on Warning Messages for more information.
```

Concern 3: Interpreting Results (previous examples also apply here)

```
. * Another example - suggests gender may not have ordinal relationship
. * with health as it is coded

. webuse nhanes2f

. gologit2 health female, auto svy

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

Step  1:  Constraints for parallel lines are not imposed for
          female (P Value = 0.00150)

------------------------------------------------------------------------------

Generalized Ordered Logit Estimates

Number of strata   =        31                Number of obs      =      10335
Number of PSUs     =        62                Population size    =  116997257
                                              Design df          =         31
                                              F(   4,     28)    =       8.70
                                              Prob > F           =     0.0001

------------------------------------------------------------------------------
             |             Linearized
      health |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
poor         |
      female |   .1212723   .0975363     1.24   0.223    -.0776543    .3201988
       _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
------------------------------------------------------------------------------
```