

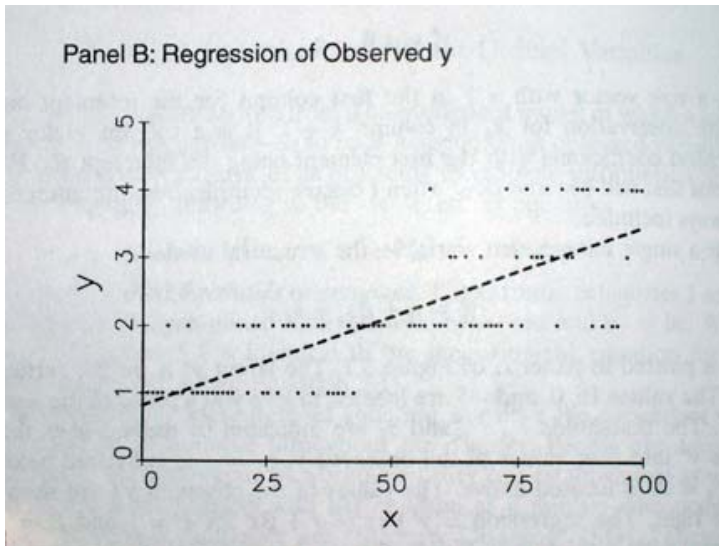
SC704: Topics in Multivariate Statistics
Instructor: Natalia Sarkisian

Ordered Logit

When the outcome variable is categorical but not binary - that is, either an ordinal variable or a nominal one with more than 2 categories—we can also use logit models, but need to modify them.

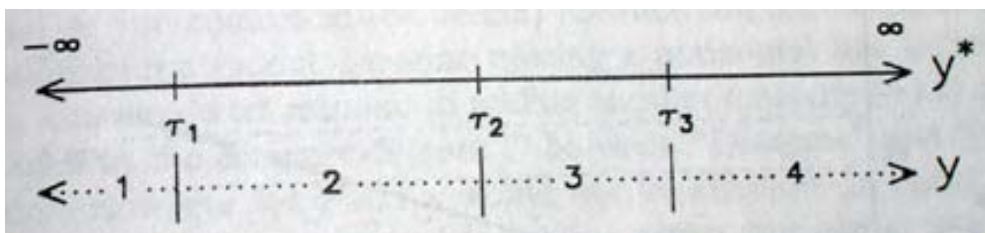
If your dependent variable has ordered categories (i.e. the order of categories is meaningful but the distances between them are arbitrary), you can use ordered logit. For some variables, the order is much clearer than for others, but always exercise caution and think whether this is the only order possible or whether another one might make sense as well.

It is inappropriate to use OLS for ordinal dependent variables - OLS assumes that the distances between categories are the same - e.g. the distance from "strongly agree" and "agree" equals to that from "agree" to "neither agree nor disagree", but in most cases we can't make that assumption. This is what OLS does if used with ordinal variables:

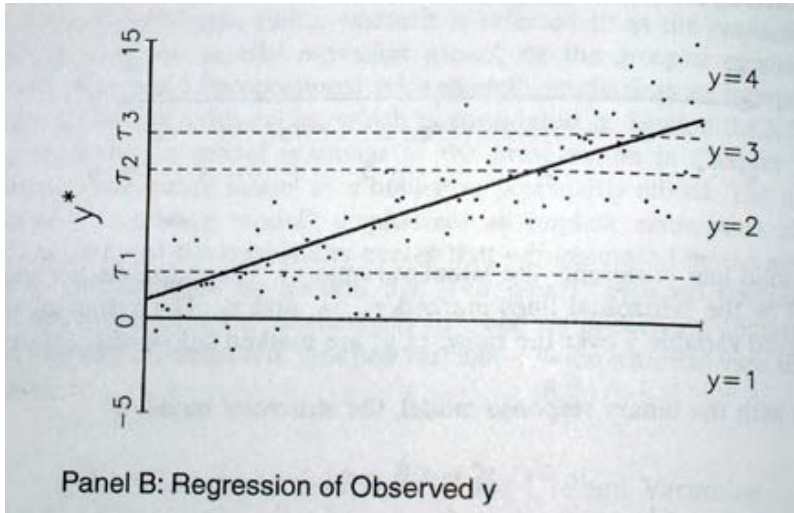


It is clear from this picture that if we changed intervals and decided that the distances are not all equal, that would change the slope. To avoid this problem, we can use ordered logit. It is based on the idea of a latent dependent variable, which we can only observe as a set of categories - but in fact, it is a continuous variable. E.g. even if we ask people's opinion on abortion in discrete categories, the most accurate representation of their views would be to position them somewhere on the continuum of support for abortion.

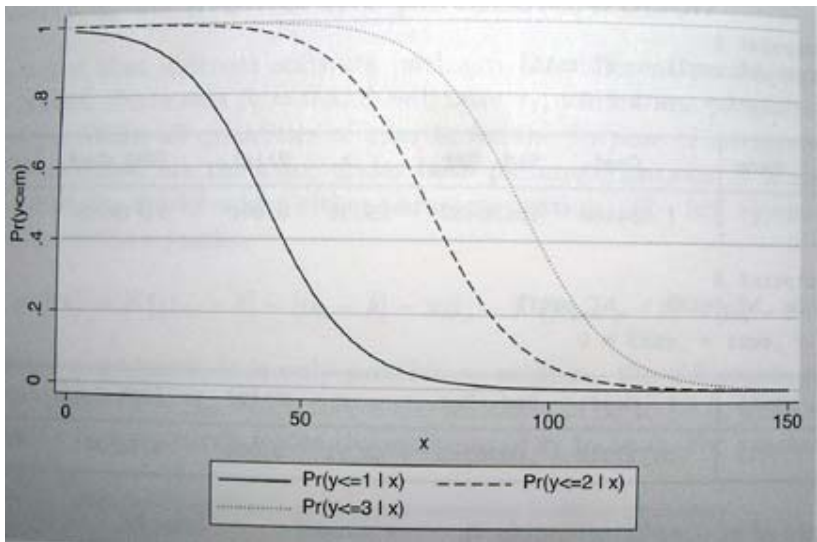
So we assume a latent dependent variable, and it is divided into intervals - those are categories we actually observe:



Then, our regression model of latent Y on X is assumed to look like this (you can see how the categories are mapped onto the latent variable - they are not equal).



This is one interpretation of ordered logit model. Another one is that it combines a set of binary logits by constraining them to be the same equation. We could estimate binary logit models for each category to predict probability of belonging to that group or any group below it. We could then require all of these logits to have the same slopes and we could estimate them simultaneously - the result is the ordered logit model. To understand why they have to be the same (this is called parallel slopes assumption), we can return to our latent Y model - the slope of the line is the same across all categories - for the entire span of the latent variable. That is how this assumption looks when we examine probabilities:



Now, let's run ordered logit model in Stata. I selected a variable that evaluates opinions on governmental spending on national defense:

```
. tab natarmsy
      national
defense --
version y |          Freq.      Percent      Cum.
-----+-----
too little |           477       35.39       35.39
about right |           591       43.84       79.23
too much   |           280       20.77      100.00
-----+-----
Total     |        1,348      100.00
```

```

. ologit natarmsy age sex childsex educ born
Iteration 0:   log likelihood = -1410.9409
Iteration 1:   log likelihood = -1391.9261
Iteration 2:   log likelihood = -1391.882
Iteration 3:   log likelihood = -1391.882
Ordered logit estimates
Log likelihood = -1391.882
Number of obs   =      1337
LR chi2(5)      =      38.12
Prob > chi2     =      0.0000
Pseudo R2      =      0.0135

```

	Coef.	Std. Err.	z	P> z	[95% Conf. Interval]	
natarmsy						
age	-.0111591	.0032866	-3.40	0.001	-.0176007	-.0047176
sex	.1686415	.1034603	1.63	0.103	-.034137	.37142
childsex	.0095746	.0347702	0.28	0.783	-.0585737	.0777229
educ	.0326995	.0174081	1.88	0.060	-.0014198	.0668188
born	.7142956	.1829363	3.90	0.000	.3557471	1.072844

_cut1	.3558196	.3847395	(Ancillary parameters)			
_cut2	2.341709	.3908546				

Measures of Fit:

Measures of fit for ordered logit models can be obtained using fitstat. Simulations indicate that McKelvey and Zavoina's R squared most closely approximate the R squared obtained by fitting OLS using the underlying latent variable, so this measure of R2 is the most appropriate. To choose the best-fitting model, we can do hypotheses tests using test and lrtest as well as use BIC comparisons.

Interpretation:

1. Coefficients and Odds Ratios

The output looks almost like the binary logit output - except for the cutoff values on the bottom - those are the values of latent Y which we used to create categories - those values used to cut up our imaginary Y (opposition to defense expenditures - larger number means more opposed) to get the observed three categories.

We focus our interpretation on coefficients - and we can interpret them the same way as we interpreted binary logit coefficients. So we can interpret the sign and the significance but not the size of the coefficients in the table above. We find that age decreases opposition to defense expenditures, and being foreign born increases such opposition. Education is only significant on .1 level and increases such opposition as well.

One type of interpretation of results that works exclusively for ordered logit (it doesn't exist for either binary or multinomial logit) is the interpretation of Y-standardized and fully standardized coefficients as the change (measured in standard deviations) in latent Y variable per unit of X or per standard deviation of X:

```

. listcoef, std
ologit (N=1337): Unstandardized and Standardized Estimates
Observed SD: .73511836
Latent SD: 1.8407959

```

	b	z	P> z	bStdX	bStdY	bStdXY	SDofX
natarmsy							
age	-0.01116	-3.395	0.001	-0.1941	-0.0061	-0.1055	17.3958
sex	0.16864	1.630	0.103	0.0840	0.0916	0.0456	0.4981
childsex	0.00957	0.275	0.783	0.0163	0.0052	0.0088	1.6975
educ	0.03270	1.878	0.060	0.0995	0.0178	0.0540	3.0423

```
born | 0.71430 3.905 0.000 0.1972 0.3880 0.1071 0.2760
```

 So one year increase in age decreases the latent Y (opposition to defense expenditures) by .006 standard deviations, and one standard deviation increase in age (which is 17.4 years) decreases the opposition to defense expenditures by .1055 standard deviations.

All other types of interpretation of results are very similar to binary logit. The only complication here is that we have multiple groups, so we will have to be careful about that. So for example we can obtain odds ratios:

```
. ologit natarmsy age sex child educ born, or
Ordered logit estimates
Number of obs = 1337
LR chi2(5) = 38.12
Prob > chi2 = 0.0000
Pseudo R2 = 0.0135
Log likelihood = -1391.882
```

```
-----
```

natarmsy	Odds Ratio	Std. Err.	z	P> z	[95% Conf. Interval]
age	.9889029	.0032501	-3.40	0.001	.9825533 .9952935
sex	1.183696	.1224655	1.63	0.103	.9664391 1.449792
child	1.009621	.0351047	0.28	0.783	.9431087 1.080823
educ	1.03324	.0179868	1.88	0.060	.9985812 1.069102
born	2.042747	.3736925	3.90	0.000	1.427247 2.923683

```
-----
```

We can also use listcoef with various options the same way as for binary. These are cumulative odds of belonging to a certain category or higher versus belonging to one of the lower categories. So we can say that the odds of thinking that we spend too much versus thinking that we spend about right or too little are 2 times higher for those who are foreign born. Similarly, the odds of thinking that we spend about right or too much versus that we spend too little are also twice as high for foreign born people as they are for American born.

To better understand what these are, let's calculate odds and odds ratios:

```
. tab natarmsy born
national | was r born in this
defense -- | country
version y | yes no | Total
-----+-----+-----
too little | 456 21 | 477
about right | 533 57 | 590
too much | 244 34 | 278
-----+-----+-----
Total | 1,233 112 | 1,345
. di (533+244)/456
1.7039474
*odds of saying about right or too much for native born (without any controls)
. di (57+34)/21
4.3333333
*odds for saying about right or too much for foreign born
*Odds ratio:
. di 4.3333333/1.7039474
2.5431145
Alternatively:
. di 244/(533+456)
.24671385
*odds of saying too much for native born
. di 34/(57+21)
.43589744
*odds of saying too much for foreign born
*Odds ratio:
. di .43589744/.24671385
1.7668138
```

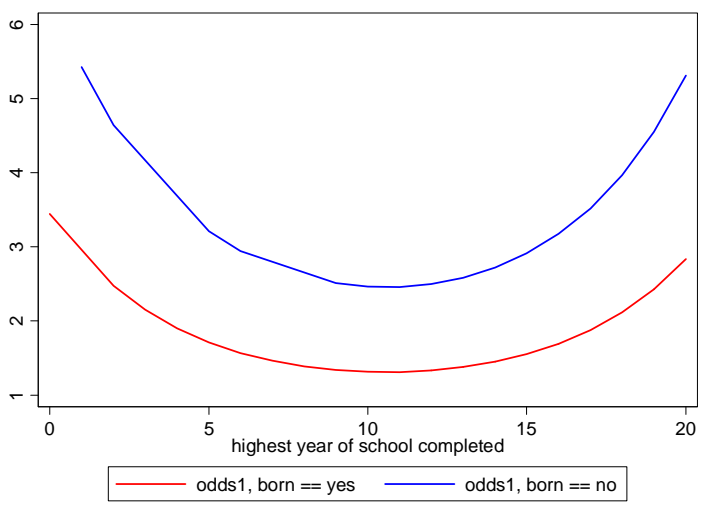
Note that the odds ratio for born in the ologit output is approximately in the middle between these two values: $(1.7668138 + 2.5431145)/2 = 2.1549642$. That's because ologit assumes that these two odds ratios are essentially the same and thus uses the average. That's the parallel slopes assumption in action. So we are assuming these two odds ratios are the same - if they differ significantly, the assumption is violated. We'll learn how to test that later.

For ologit, we can also use adjust command to explore odds (adjust does not allow us to obtain probabilities for ologit) - use exp option.

```
. sum educ
-----+-----
Variable |      Obs      Mean   Std. Dev.   Min   Max
-----+-----
educ     |    2753   13.36397   2.973924     0    20
. gen educm=educ-r(mean)
(12 missing values generated)
. gen educm2=educm^2
(12 missing values generated)
. ologit natarmsy age sex childs born educm educm2
Ordered logistic regression
Number of obs   =    1337
LR chi2(6)      =    47.41
Prob > chi2     =    0.0000
Pseudo R2      =    0.0168
Log likelihood = -1387.2381
```

natarmsy	Coef.	Std. Err.	z	P> z	[95% Conf. Interval]	
age	-.0124031	.0033202	-3.74	0.000	-.0189106	-.0058955
sex	.1785567	.1036235	1.72	0.085	-.0245415	.381655
childs	.003951	.0348873	0.11	0.910	-.0644269	.072329
born	.6292186	.185014	3.40	0.001	.2665979	.9918394
educm	.0485421	.0181214	2.68	0.007	.0130248	.0840593
educm2	.0086708	.0028482	3.04	0.002	.0030884	.0142531
/cut1	-.1498701	.3016316			-.7410572	.441317
/cut2	1.847077	.306773			1.245813	2.448341

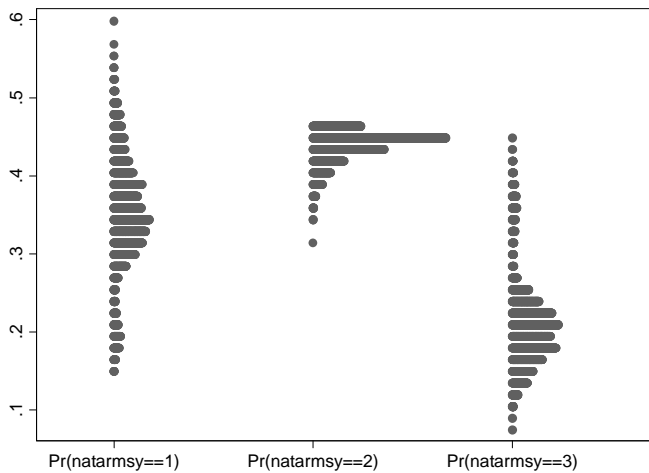
```
. qui adjust age sex childs if e(sample), gen(odds1) exp
. separate odds1, by(born)
storage display value
variable name type format label variable label
-----+-----
odds11 float %9.0g odds1, born == yes
odds12 float %9.0g odds1, born == no
. line odds11 odds12 educ, sort lcolor(red blue)
```



2. Predicted Probabilities.

Further, we can examine predicted probabilities the same way as for binary logit - but, now we will always have sets of predicted probabilities - reflecting the number of categories.

```
. qui ologit natarmsy age sex childs educ born
. predict p1 p2 p3
(option p assumed; predicted probabilities)
(26 missing values generated)
. dotplot p1 p2 p3
```



```
. prvalue
ologit: Predictions for natarmsy2
Pr(y=1|x):      0.3508
Pr(y=2|x):      0.4466
Pr(y=3|x):      0.2025
      age      sex      childs      educ      born
x=  46.367988  1.5459985  1.854899  13.352281  1.0830217
```

So for all average values, the probability of thinking that we spend too little is 35%, about right - 45%, and too much - 20%. That corresponds to the original distribution (see p.3).

Again, we can select specific values of independent variables to get meaningful results using prvalue. We can also get tables of predicted probabilities:

```
. prtab sex born
ologit: Predicted probabilities for natarmsy
Predicted probability of outcome 1
```

respondents sex	was r born in this country	
	yes	no
male	0.3860	0.2353
female	0.3469	0.2064

Predicted probability of outcome 2

respondents sex	was r born in this country	
	yes	no
male	0.4348	0.4562
female	0.4478	0.4481

Predicted probability of outcome 3

respondents sex	was r born in this country	
	yes	no
male	0.1792	0.3084
female	0.2054	0.3455

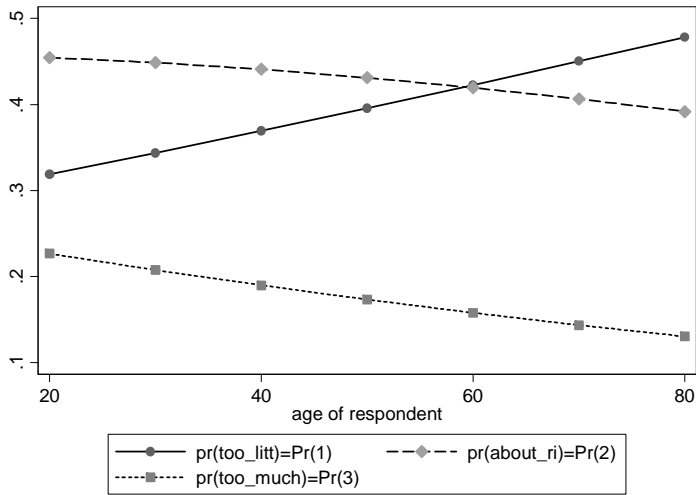
```
x=      age      sex      childs      educ      born
46.367988 1.5459985 1.854899 13.352281 1.0830217
```

And we can create graphs of predicted probabilities as well as cumulative predicted probabilities. Focusing on native born men:

```
. prgen age, from(20) to(80) gen(prage) x(born=1 sex=1) ncases(7)
ologit: Predicted values as age varies from 20 to 80.
```

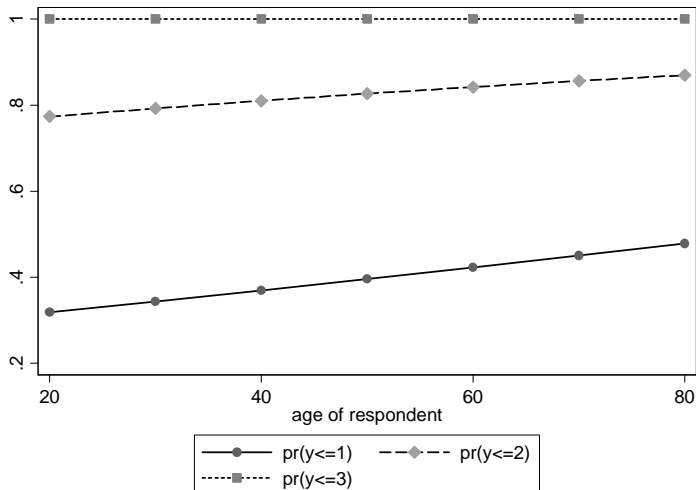
```
      age      sex      childs      educ      born
x= 46.367988      1 1.854899 13.352281      1
```

```
. graph twoway connected pragep1 pragep2 pragep3 pragep
```

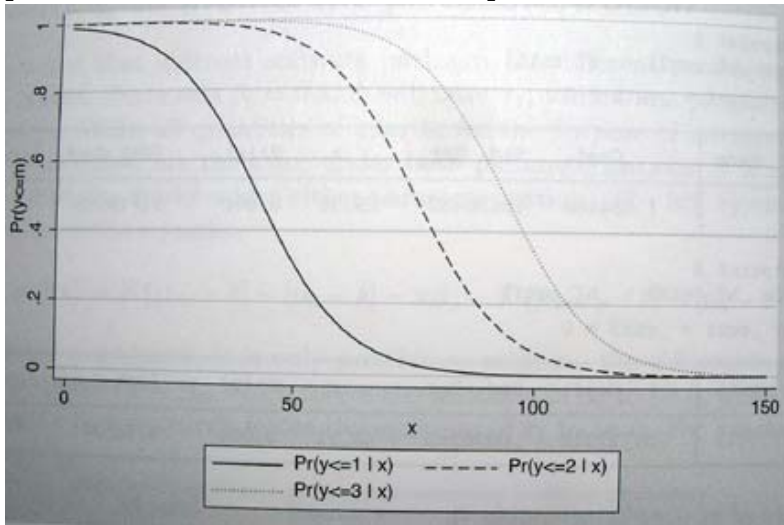


Note that the fact that slopes go in different directions is normal - as probability of being in one category increases, the probability of being in another category decreases. We can also graph cumulative probabilities - these should be parallel (reflects the assumption of parallel slopes):

```
. graph twoway connected prages1 prages2 prages3 pragec
```



In interpreting this graph, we focus on the distances between the lines rather than the lines themselves - your book shows how you can shade the areas to focus on areas rather than lines - see p.253-254). By the way, the lines don't look parallel because the curves are positioned differently along X axis:



3. Changes in Probabilities.

Similar to binary logit, we can examine discrete and marginal changes in probabilities using `prchange` and `mfx compute` commands. But here again, we will get changes for each outcome individually (it also calculates average change but that is difficult to interpret). So we don't interpret the first column.

E.g., examining changes for native born women:

```
. prchange, x(born=1 sex=2) rest(mean)
```

```
ologit: Changes in Probabilities for natarmsy
```

```
age
```

	Avg Chg	too_litt	about_ri	too_much
Min->Max	.12120405	.18180606	-.05837068	-.12343541
-+1/2	.00168547	.00252819	-.00070724	-.00182097
-+sd/2	.02930365	.04395548	-.01227728	-.0316782
MargEfct	.00168547	.00252821	-.00070725	-.00182096

```
sex
```

	Avg Chg	too_litt	about_ri	too_much
Min->Max	.02608323	-.03912485	.01296741	.02615744
-+1/2	.02546073	-.03819111	.01067123	.02751985
-+sd/2	.01268518	-.01902777	.00532132	.01370645
MargEfct	.02547157	-.03820735	.01068818	.02751917

```
childs
```

	Avg Chg	too_litt	about_ri	too_much
Min->Max	.01149452	-.01724178	.00459129	.0126505
-+1/2	.00144615	-.00216922	.00060683	.0015624
-+sd/2	.00245486	-.00368229	.00103009	.00265221
MargEfct	.00144614	-.00216921	.00060682	.00156239

```
educ
```

	Avg Chg	too_litt	about_ri	too_much
Min->Max	.10114521	-.15171781	.05172443	.09999338
-+1/2	.00493884	-.00740826	.0020723	.00533596
-+sd/2	.01502322	-.02253485	.00630137	.01623344
MargEfct	.00493892	-.00740839	.00207243	.00533595

```
born
```

```

      Avg|Chg|   too_litt   about_ri   too_much
Min->Max   .09369085  -.14053628   .00038877   .14014749
  -+1/2    .1070681  -.16060215   .04400143   .11660072
  +sd/2    .02976146 -.04464218   .01246849   .03217371
MargEfct   .10788701 -.16183052   .0452707    .11655982

```

```

      too_litt   about_ri   too_much
Pr(y|x)  .34689745  .44775236  .20535021

```

```

      age      sex   child   educ   born
x=    46.368    2    1.8549  13.3523  1
sd(x)= 17.3958  .498066  1.69753  3.04225  .276018

```

Diagnostics

1. Parallel slopes assumption

We discussed the parallel regression assumption (assumption that probability curves are parallel). Now we will learn to test it. This is crucial - if it does not hold, we should use other models (e.g. multinomial logit), but also there a few other alternatives - see p.184-187 in your textbook). The command we use here is part of Long and Freese's package we installed earlier.

```

. brant, detail
Estimated coefficients from j-1 binary regressions
      y>11      y>12
age  -.00944256  -.01591924
sex   .20596159   .1246965
child .01557951  -.00421577
educ  .02356209   .05218194
born  .96380662   .50986576
_cons -.64946649  -2.0783207

```

Brant Test of Parallel Regression Assumption

Variable	chi2	p>chi2	df
All	7.37	0.194	5
age	1.81	0.178	1
sex	0.32	0.573	1
child	0.15	0.694	1
educ	1.28	0.258	1
born	2.60	0.107	1

We interpret probability values - if the overall probability is less than chosen cutoff (e.g., .05), that we reject the assumption of parallel slopes and cannot use ordered logit model. We also get information on individual variables - that way we can see for which variables slopes are not parallel and consider respecifying the model in some fashion. In this case, none of the variables presents a problem. If the assumption is violated, we have three options - to use generalized ordered logit model, stereotype logit model, or multinomial logit model. Let's see an example where parallel slopes assumption is violated:

```

. codebook natfare

```

```

-----
natfare
welfare
-----

```

```

      type: numeric (byte)
      label: natfare

      range: [1,3]          units: 1
unique values: 3          missing .: 1451/2765
      tabulation: Freq.   Numeric  Label
                  279     1  too little
                  502     2  about right

```

```

533      3 too much
1451      .

```

```

. ologit natfare age sex child educ born
Ordered logistic regression
Number of obs = 1306
LR chi2(5) = 13.40
Prob > chi2 = 0.0199
Pseudo R2 = 0.0048
Log likelihood = -1379.0767

```

	Coef.	Std. Err.	z	P> z	[95% Conf. Interval]	
natfare						
age	-.0043618	.0033169	-1.32	0.189	-.0108629	.0021392
sex	-.0658052	.1042275	-0.63	0.528	-.2700873	.1384769
childs	-.054537	.0346097	-1.58	0.115	-.1223708	.0132967
educ	.0075963	.0181499	0.42	0.676	-.0279769	.0431695
born	-.4227496	.1761635	-2.40	0.016	-.7680237	-.0774754
/cut1	-2.094474	.3990914			-2.876679	-1.312269
/cut2	-.3762581	.3944567			-1.149379	.3968628

```

. brant, detail
Estimated coefficients from j-1 binary regressions
y>1      y>2
age      .0005797  -.00694478
sex     -.13511054  -.03743762
childs  -.07977849  -.03702665
educ    .05719413  -.02104598
born    -.50652198  -.33981992
_cons   1.4613488   .71516591

```

Brant Test of Parallel Regression Assumption

Variable	chi2	p>chi2	df
All	15.76	0.008	5
age	2.93	0.087	1
sex	0.50	0.478	1
childs	0.96	0.328	1
educ	11.03	0.001	1
born	0.62	0.432	1

A significant test statistic provides evidence that the parallel regression assumption has been violated. More specifically, the assumption is violated for education. Here we'll discuss generalized ordered logit as an alternative model; the other alternatives will be discussed later.

Generalized Ordered Logit

Let's estimate a generalized ordered logit model. We need a gologit2 command which is user-written. We find it by typing

```

. net search gologit2
. gologit2 natfare age sex child educ born
Generalized Ordered Logit Estimates
Number of obs = 1306
LR chi2(10) = 28.65
Prob > chi2 = 0.0014
Pseudo R2 = 0.0103
Log likelihood = -1371.4507

```

	Coef.	Std. Err.	z	P> z	[95% Conf. Interval]	
too_little						
age	.0004908	.0043505	0.11	0.910	-.008036	.0090176
sex	-.1339589	.1390404	-0.96	0.335	-.4064732	.1385553
childs	-.0849995	.0444353	-1.91	0.056	-.1720911	.0020921

educ		.0528277	.0230162	2.30	0.022	.0077168	.0979386
born		-.5331761	.2097069	-2.54	0.011	-.9441941	-.1221582
_cons		1.561235	.5049273	3.09	0.002	.571596	2.550874

about_right							
age		-.0069425	.0036582	-1.90	0.058	-.0141124	.0002275
sex		-.0413382	.1145167	-0.36	0.718	-.2657868	.1831103
childs		-.034073	.038231	-0.89	0.373	-.1090044	.0408584
educ		-.0203868	.0200331	-1.02	0.309	-.0596509	.0188774
born		-.3546384	.1975892	-1.79	0.073	-.7419061	.0326293
_cons		.7233665	.4381509	1.65	0.099	-.1353934	1.582126

This estimates the two models separately, the same way brant test did. We could do a similar test by comparing the two equations:

```
. test [too_little=about_right]

( 1) [too_little]age - [about_right]age = 0
( 2) [too_little]sex - [about_right]sex = 0
( 3) [too_little]childs - [about_right]childs = 0
( 4) [too_little]educ - [about_right]educ = 0
( 5) [too_little]born - [about_right]born = 0

      chi2( 5) =    16.34
      Prob > chi2 =    0.0059
```

Now, let's make use of some more advanced options:

```
. gologit2 natfare age sex childs educ born, autofit gamma
```

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

```
Step 1: sex meets the pl assumption (P Value = 0.5024)
Step 2: born meets the pl assumption (P Value = 0.3904)
Step 3: childs meets the pl assumption (P Value = 0.2221)
Step 4: age meets the pl assumption (P Value = 0.1549)
Step 5: The following variables do not meet the pl assumption:
      educ (P Value = 0.00082)
```

Wald test of parallel lines assumption for the final model:

```
( 1) [too little]sex - [about right]sex = 0
( 2) [too little]born - [about right]born = 0
( 3) [too little]childs - [about right]childs = 0
( 4) [too little]age - [about right]age = 0
      chi2( 4) =    4.63
      Prob > chi2 =    0.3272
```

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(sex born childs age)

```
-----
Generalized Ordered Logit Estimates          Number of obs   =    1306
                                              Wald chi2(6)     =    24.41
                                              Prob > chi2      =    0.0004
                                              Pseudo R2       =    0.0087
Log likelihood = -1373.774
( 1) [too little]sex - [about right]sex = 0
( 2) [too little]born - [about right]born = 0
( 3) [too little]childs - [about right]childs = 0
( 4) [too little]age - [about right]age = 0
```

		Coef.	Std. Err.	z	P> z	[95% Conf. Interval]
too little						
age		-.0043832	.0033113	-1.32	0.186	-.0108732 .0021068

sex	-.0702248	.1044244	-0.67	0.501	-.2748929	.1344433
childs	-.0514524	.0345432	-1.49	0.136	-.1191558	.016251
educ	.0533867	.022804	2.34	0.019	.0086916	.0980818
born	-.4252105	.1766018	-2.41	0.016	-.7713436	-.0790774
_cons	1.496639	.4354585	3.44	0.001	.643156	2.350122

about right						
age	-.0043832	.0033113	-1.32	0.186	-.0108732	.0021068
sex	-.0702248	.1044244	-0.67	0.501	-.2748929	.1344433
childs	-.0514524	.0345432	-1.49	0.136	-.1191558	.016251
educ	-.0206651	.0200009	-1.03	0.302	-.0598661	.0185359
born	-.4252105	.1766018	-2.41	0.016	-.7713436	-.0790774
_cons	.7614398	.4118299	1.85	0.064	-.0457319	1.568612

Alternative parameterization: Gammas are deviations from proportionality

natfare	Coef.	Std. Err.	z	P> z	[95% Conf. Interval]	

Beta						
age	-.0043832	.0033113	-1.32	0.186	-.0108732	.0021068
sex	-.0702248	.1044244	-0.67	0.501	-.2748929	.1344433
childs	-.0514524	.0345432	-1.49	0.136	-.1191558	.016251
educ	.0533867	.022804	2.34	0.019	.0086916	.0980818
born	-.4252105	.1766018	-2.41	0.016	-.7713436	-.0790774

Gamma_2						
educ	-.0740518	.022142	-3.34	0.001	-.1174493	-.0306543

Alpha						
_cons_1	1.496639	.4354585	3.44	0.001	.643156	2.350122
_cons_2	.7614398	.4118299	1.85	0.064	-.0457319	1.568612

I used autofit model to keep all the coefficients that are not significantly different constrained to be equal, and allow only unequal coefficients (here, coefficients for educ) to vary.

Gamma option allows the alternative parametrization which presents coefficients for the first model (y>1) and then presents any deviations from that model in other models as gamma coefficients. So here we can see that the only gamma is for education - the coefficient in y>2 model is -.074 smaller (and as we can see from the earlier output, the effect of education in that model is, in fact, not significant).

We can also use various post-estimation commands with gologit2, but in order to do that, we need to reestimate the model as gologit version 1 (they didn't adjust the postestimation commands to work with gologit2). We can do that using option v1.

```
. prvalue, x(educ=16) rest(mean)
prvalue does not work for the last type of model estimated.
. quietly gologit2 natfare age sex childs educ born, autofit gamma v1
. prvalue, x(educ=16) rest(mean)
gologit: Predictions for natfare
Confidence intervals by delta method
                                95% Conf. Interval
Pr(y=too_litt|x):    0.1848    [ 0.1568,    0.2129]
Pr(y=about_ri|x):   0.4225    [ 0.3866,    0.4584]
Pr(y=too_much|x):   0.3927    [ 0.3565,    0.4288]
      age          sex      childs      educ      born
x= 45.944104  1.5520674  1.754977      16  1.0964778
```

Now it works. Can also use prtab, prgen etc.

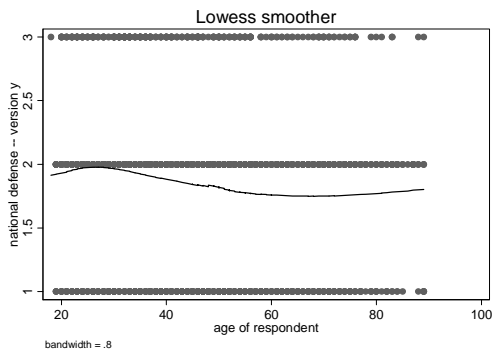
2. Multicollinearity.

As was the case for binary logit, we can test for multicollinearity by running OLS model instead of ordered logit and using vif.

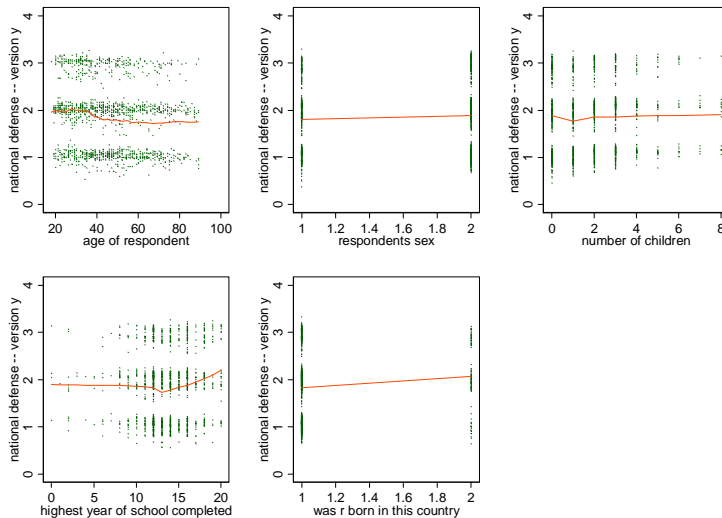
3. Linearity and Additivity

For additivity and the issue of interactions, the story is as complex as for binary logit and the same considerations apply. Rely on theory in selecting interactions, and use predicted probabilities and discrete changes to examine the results. Also see Chapter 9 -- 9.2, 9.3, 9.4 (pp.423-443) for more examples and tools to examine interactions in nonlinear models.

As for linearity, as always, we need to start our ordered logit analyses by conducting univariate and bivariate examination of the data. For bivariate examination, an ordered variable can be used in two ways - you can either use it as if it were continuous (especially if the number of categories is relatively high) or you can split it into dichotomies and use logistic-based tools. E.g.:

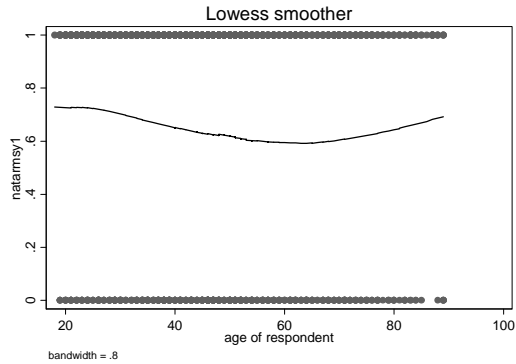


```
. mrunning natarmsy age sex childs born educ
```

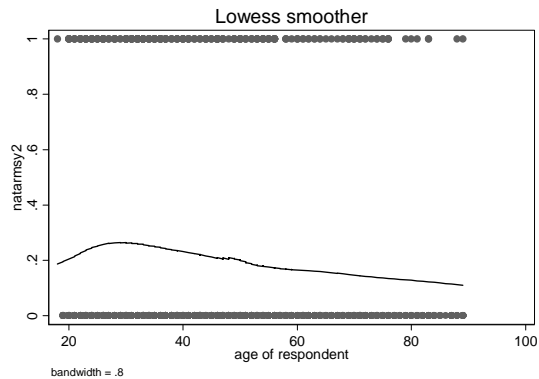


```
Or you can create dichotomies (note that these are cumulative dichotomies!):  
. gen natarmsy1=(natarmsy>1) if natarmsy~=.  
(1417 missing values generated)  
. gen natarmsy2=(natarmsy>2) if natarmsy~=.  
(1417 missing values generated)
```

```
And then we can use lowess, like in binary logit. E.g.:  
. lowess natarmsy1 age
```



```
. lowess natarmsy2 age
```



Looks like it's not quite linear and the shape of the relationship might differ for two equations - we could introduce age squared and then test parallel slopes.

```
. qui sum age
. gen agem=age-r(mean)
(14 missing values generated)
. gen agem2=agem^2
(14 missing values generated)

. gologit2 natarmsy agem agem2 sex childs educ born, autofit
```

```
-----
Testing parallel lines assumption using the .05 level of significance...
Step 1: Constraints for parallel lines imposed for agem (P Value = 0.7961)
Step 2: Constraints for parallel lines imposed for educ (P Value = 0.8065)
Step 3: Constraints for parallel lines imposed for sex (P Value = 0.7043)
Step 4: Constraints for parallel lines imposed for childs (P Value = 0.2648)
Step 5: Constraints for parallel lines imposed for born (P Value = 0.1295)
Step 6: Constraints for parallel lines are not imposed for
       agem2 (P Value = 0.00910)
```

```
Wald test of parallel lines assumption for the final model:
( 1) [too_little]agem - [about_right]agem = 0
( 2) [too_little]educ - [about_right]educ = 0
( 3) [too_little]sex - [about_right]sex = 0
( 4) [too_little]childs - [about_right]childs = 0
( 5) [too_little]born - [about_right]born = 0
      chi2( 5) =    3.83
      Prob > chi2 =    0.5744
```

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(agem educ sex childs born)
```

```
-----
Generalized Ordered Logit Estimates                    Number of obs   =      1337
                                                       Wald chi2(7)    =       48.66
                                                       Prob > chi2     =       0.0000
Log likelihood = -1385.356                            Pseudo R2      =       0.0181
( 1) [too_little]agem - [about_right]agem = 0
( 2) [too_little]educ - [about_right]educ = 0
( 3) [too_little]sex - [about_right]sex = 0
( 4) [too_little]childs - [about_right]childs = 0
( 5) [too_little]born - [about_right]born = 0
-----
```

	natarmsy	Coef.	Std. Err.	z	P> z	[95% Conf. Interval]	
too_little							
	agem	-.0152224	.0036664	-4.15	0.000	-.0224083	-.0080365
	agem2	.0005836	.0001825	3.20	0.001	.0002259	.0009413
	sex	.1554398	.1036979	1.50	0.134	-.0478043	.3586839
	childs	.0224146	.035386	0.63	0.526	-.0469407	.09177
	educ	.0418801	.0179471	2.33	0.020	.0067044	.0770557
	born	.7037078	.1833104	3.84	0.000	.3444259	1.06299
	_cons	-1.160804	.3766593	-3.08	0.002	-1.899042	-.4225649

about_right							
	agem	-.0152224	.0036664	-4.15	0.000	-.0224083	-.0080365
	agem2	-.0000121	.0002272	-0.05	0.958	-.0004574	.0004333
	sex	.1554398	.1036979	1.50	0.134	-.0478043	.3586839
	childs	.0224146	.035386	0.63	0.526	-.0469407	.09177
	educ	.0418801	.0179471	2.33	0.020	.0067044	.0770557
	born	.7037078	.1833104	3.84	0.000	.3444259	1.06299
	_cons	-2.980485	.3885241	-7.67	0.000	-3.741978	-2.218991

We can see that the square term of age is significant in one equation only. Turning to diagnosing linearity in multivariate context, here we need to estimate multiple binary models and do the diagnostics separately for them.

```
. boxtid logit natarmsy1 age sex childs educ born
```

```
-----
```

age		-.0104898	.0036217	-2.896	Nonlin. dev. 5.549	(P = 0.018)
	p1	-1.517852	1.362114	-1.114		

childs		.0340214	.0379598	0.896	Nonlin. dev. 0.156	(P = 0.693)
	p1	1.813904	3.115287	0.582		

educ		.0314742	.0198539	1.585	Nonlin. dev. 7.191	(P = 0.007)
	p1	7.9854	4.681117	1.706		

```
. boxtid logit natarmsy2 age sex childs educ born
```

```
-----
```

age		-.015219	.0044118	-3.450	Nonlin. dev. 0.247	(P = 0.619)
	p1	1.586871	1.546701	1.026		

childs		-.0103851	.0476276	-0.218	Nonlin. dev. 2.477	(P = 0.116)
	p1	-46.97068	.	.		

educ		.0466299	.0237333	1.965	Nonlin. dev. 5.772	(P = 0.016)
	p1	7.863133	4.870539	1.614		

```
-----
```

Here we also see a nonlinear relationship for age in the first but not the second model. Education appears nonlinear in both.

4. Outliers and Influential Observations

In order to do unusual data diagnostics for ordered logit, we should also rely on separate binary models we've used in previous steps. So we should obtain residuals and influence statistics from them (so all the same methods we discussed for binary logit apply here as well), e.g. getting standardized residuals:

```
. qui logit natarmsyl age sex child educ born
. predict resid1, rs
(1428 missing values generated)
. qui logit natarmsy2 age sex child educ born
. predict resid2, rs
(1428 missing values generated)
```

Note that the fact that you'll have to do a separate search for unusual data for each binary model may complicate things if they suggest that different observations are influential; you'll have to then test the potential effects of these influential observations on your ologit model (rather than just on individual binary logits).

5. Error term distribution

Like we did for binary logit, we can obtain robust standard errors for the ordered logit model in order to check whether our assumptions about error distribution hold (compare with the model on p.3):

```
. ologit natarmsy age sex child educ born, robust
Iteration 0:   log pseudolikelihood = -1410.9409
Iteration 1:   log pseudolikelihood = -1391.9261
Iteration 2:   log pseudolikelihood = -1391.882
Iteration 3:   log pseudolikelihood = -1391.882
Ordered logistic regression
Number of obs   =      1337
Wald chi2(5)    =      41.23
Prob > chi2     =      0.0000
Pseudo R2      =      0.0135
Log pseudolikelihood = -1391.882
```

	Coef.	Robust Std. Err.	z	P> z	[95% Conf. Interval]	
natarmsy						
age	-.0111591	.0032531	-3.43	0.001	-.0175351	-.0047831
sex	.1686415	.1039126	1.62	0.105	-.0350235	.3723065
child	.0095746	.0352056	0.27	0.786	-.0594272	.0785763
educ	.0326995	.0172806	1.89	0.058	-.0011699	.0665689
born	.7142956	.1695103	4.21	0.000	.3820615	1.04653
/cut1	.3558196	.3880631			-.4047701	1.116409
/cut2	2.341709	.3944431			1.568615	3.114803

Example of ordered logit:

Michelson, Melissa R. 2003. "The Corrosive Effect of Acculturation: How Mexican Americans Lose Political Trust." *Social Science Quarterly*, 82(4), pp.918-933.

Questions to answer about the article:

1. What are the dependent and the independent variables in this analysis?
2. What is reported in Table 2? How can we interpret these results? How do the authors discuss these results in the text?
3. What is reported in Table 3? How can we interpret these results?
4. In addition to what the authors chose to present, how else could they have presented their results?
5. What measures of model fit and model diagnostics are presented? What diagnostics and potential problems did the authors not address?