

The Proportional Odds Assumption in Ordered Logit/Probit Models

Afees Salisu

The proportional odds assumption

- One of the assumptions underlying ordered logistic (and ordered probit) regression is that the relationship between each pair of outcome groups is the same.

Odds assumption ...

- In other words, ordered logistic regression assumes that the coefficients that describe the relationship between, say, the lowest versus all higher categories of the response variable are the same as those that describe the relationship between the next lowest category and all higher categories, etc.

Odds assumption ...

- This is called the **proportional odds assumption** or the **parallel regression assumption**.
- In practice, violating this assumption may or may not alter your substantive conclusions. You need **to test whether this is the case**.

Odds assumption ...

- There are several tests for verifying this assumption.
- Among these tests are Wolfe Gould, Brant, Score, Likelihood ratio and Wald tests.
- The underlying null hypothesis is that the relationship is proportional; that is, parallel.

Odds assumption ...

- To test in Stata, use 'oparallel' command.
- It is a user written program.
- To download the command type "findit oparallel" in Stata.
- Once downloaded, you can type "oparallel" immediately after you estimate an ordered logit model ("ologit") to perform the test.

Odds assumption ...

- In the case of our example, the oparallel test is shown below

```
. oparallel
```

```
Tests of the parallel regression assumption
```

	Chi2	df	P>Chi2
Wolfe Gould	21.33	3	0.000
Brant	17.45	3	0.001
score	26.66	3	0.000
likelihood ratio	22.05	3	0.000
Wald	28.88	3	0.000

Odds assumption ...

- Interpretation:
- The relationship is not proportional across all the test statistics.

Dealing with violation ...

- Option 1: **Do nothing.** Use ordered logistic regression because the practical implications of violating this assumption are minimal.
- Option 2: **Use a multinomial logit model.** This frees you of the proportionality assumption, but it is less parsimonious and often dubious on substantive grounds.
- Option 3: **Dichotomize the outcome and use binary logistic regression.** This is common, but you lose information and it could alter your substantive conclusions.

Dealing with violation ...

- For instance, in the case of our example, one can merge categories 1 & 2 since their marginal effects are similar both in terms of sign and significance.
- **Stata command on how to dichotomize:**
 `recode healthstatus (1 2 = 0)(3 = 1),
 gen(health)`
- **Stata command for estimation:**
 `logit health age logincome numberdiseases`

Dealing with violation ...

- Option 4: Use a model that does not assume proportionality. Increasingly, this is common. Two user-submitted Stata commands fit these kinds of models:
 - “**gologit2**” – generalized ordered logit models (see Williams 2007, Stata Jn.).
 - “**oglm**” – heterogeneous choice models (see Williams 2010, Stata Jn.)

Dealing with violation ...

- Note that they are user-written programs. You have to install them before estimation.
- Stata estimation commands after installation:
oglm healthstatus age logincome
numberdiseases

gologit2 healthstatus age logincome
numberdiseases
- Note that the results obtained here are log odds

Dealing with violation ...

- For the odds ratio, use:

```
oglm healthstatus age logincome  
numberdiseases, or
```

```
gologit2 healthstatus1 age logincome  
numberdiseases, or
```

Dealing with violation ...

- For the marginal effects, use the Stata commands below immediately after estimating with `oglm / gologit2`;
 `margins, dydx(*) predict(outcome(1))
 atmeans`
 `margins, dydx(*) predict(outcome(2))
 atmeans`
 `margins, dydx(*) predict(outcome(3))
 atmeans`

Dealing with violation ...

- Like the *ologit* command, you can also perform relevant scenario analyses for the *oglm* and *gologit2* commands.
- For instance, you may be interested in evaluating how certain fixed values of age will affect the probability of reporting excellent health status

Dealing with violation ...

- . Stata commands:

```
margins, dydx(*) at(age=(25 30 35))  
predict(outcome(3)) atmeans
```

```
margins, dydx(*) at(age=25/30 )  
predict(outcome(3)) atmeans
```

- Please update your do-file appropriately.