Ordinal Outcomes

Objectives

Derive the ORM as a nonlinear probability model and as a latent variable model

Apply methods of interpretation from the BRM to the ORM

Discuss parallel regression assumption and its consequences

Ordinal Variables

Categories can be ranked

Distances between categories are unknown

Examples from your own field

Consequence of treating an ordinal dependent variable as though it were linear

True Versus Assumed Level of Measurement

		True Level		
		Nominal	**Ordinal**	**Interval**
Assumed Level	**N**	Ok	*Inefficient*	*Inefficient*
	O	*Biased*	Ok	*Inefficient*
	I	*Biased*	*Biased*	Ok

Nonlinear Probability Model

Write model in terms of cumulative probabilities

$$C_{i,j} = \Pr(y_i \le j) = \sum_{k=1}^{j} \Pr(y_i = j), \quad j = 1, \dots, J$$

as a function of a vector of independent variables

$$C_{i,j} = F(\alpha_j + \mathbf{x}_i \beta), \quad j = 1, \dots, J-1$$

Note: $(J-1)\alpha_j$ parameters

CDF

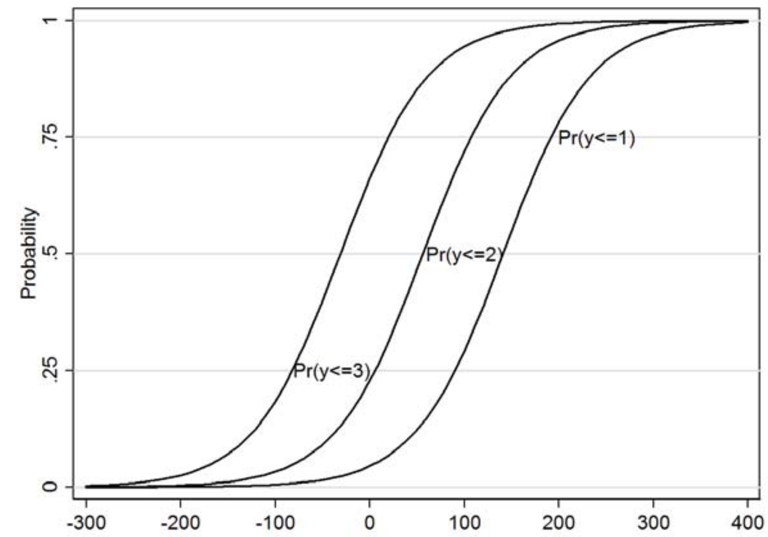
probit:

$$\Phi(\mathbf{x}\beta) = \int_{-\infty}^{\mathbf{x}\beta} \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{t^2}{2}\right) dt$$

logit:

$$\Lambda(\mathbf{x}\beta) = \frac{\exp(\mathbf{x}\beta)}{1+\exp(\mathbf{x}\beta)}$$

Cumulative Probabilities for a Four-Category Response



Conditional probabilities of ordered outcomes

$$\Pr(y_i = j \mid \mathbf{x}_i) = \begin{cases} F(\alpha_1 + \mathbf{x}_i\beta) & j=1 \\ F(\alpha_j + \mathbf{x}_i\beta) - F(\alpha_{j-1} + \mathbf{x}_i\beta) & 1<j\leq J-1 \\ 1-F(\alpha_{J-1} + \mathbf{x}_i\beta) & j=J \end{cases}$$

ORM models cumulative probabilities

Ordinal Logit Model:

$$C_{i,j} = \Pr(y_i \leq j \mid \mathbf{x}_i) = \frac{\exp(\alpha_j + \mathbf{x}_i\beta)}{1+\exp(\alpha_j + \mathbf{x}_i\beta)}$$

Ordinal Probit Model:

$$C_{i,j} = \Pr(y_i \leq j \mid \mathbf{x}_i) = \Phi(\alpha_j + \mathbf{x}_i\beta)$$

Alternative Parameterization

Instead of defining

$C_{i,j} = \Pr(y_i \le j)$

We could have defined

$C_{i,j} = \Pr(y_i > j)$

Since the distribution is symmetric, this is equivalent to

$1 - \Pr(y_i \le j)$

Question (for you)

How would this impact the coefficients from the standard parameterization?

A Latent Variable Model

Structural Model

Matrix form

$y_i^* = \mathbf{x}_i \boldsymbol{\beta} + \varepsilon_i$

For a single independent variable

$y_i^* = \alpha + \beta x_i + \varepsilon_i$

For two

$y_i^* = \beta_0 + \beta_1 \text{Age} + \beta_2 \text{Educ} + \varepsilon_i$

Measurement Model

Consider the dependent variable:

> “A working mother can establish just as warm and secure of a relationship with her child as a mother who does not work.”

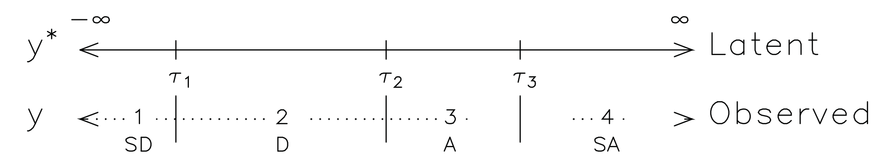
With response categories:

$$y = \begin{cases} 1 \Rightarrow \text{SD-Strongly Disagree} \\ 2 \Rightarrow \text{D-Disagree} \\ 3 \Rightarrow \text{A-Agree} \\ 4 \Rightarrow \text{SA-Strongly Agree} \end{cases}$$

The observed y is related to y* according to the measurement model:

$$y = \begin{cases} 1 \Rightarrow \text{SD-Strongly Disagree} & \text{if } \tau_0 = -\infty \le y^* < \tau_1 \\ 2 \Rightarrow \text{D-Disagree} & \text{if } \tau_1 \le y^* < \tau_2 \\ 3 \Rightarrow \text{A-Agree} & \text{if } \tau_2 \le y^* < \tau_3 \\ 4 \Rightarrow \text{SA-Strongly Agree} & \text{if } \tau_3 \le y^* < \tau_4 = \infty \end{cases}$$

Graphically:

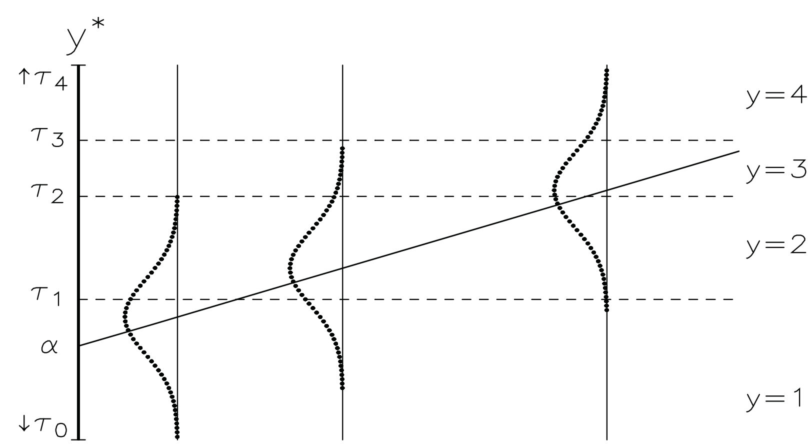


General measurement model:

$y_i = m \quad \text{if } \tau_{m-1} \le y_i^* < \tau_m \quad \text{for } m = 1 \text{ to } J$

Deriving $\Pr(y = 1 | \mathbf{x})$

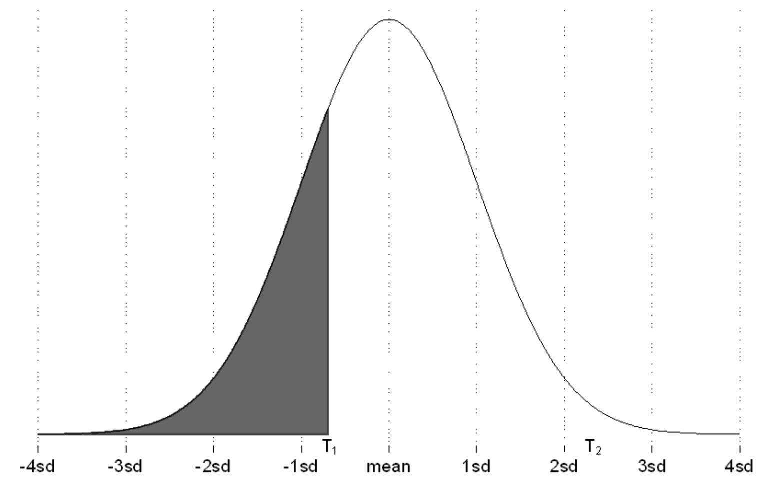
$\Pr(y_i = 1 | \mathbf{x}_i) = \Pr(\tau_0 \le y_i^* < \tau_1 | \mathbf{x}_i)$



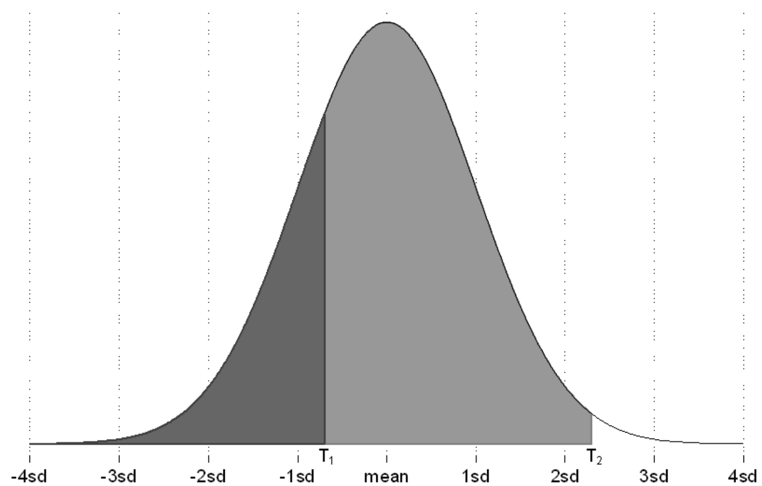
Deriving $\Pr(y=m|\mathbf{x})$

The $\Pr(y=m)$ is the *area of the error distribution* between τ_m and τ_{m-1}

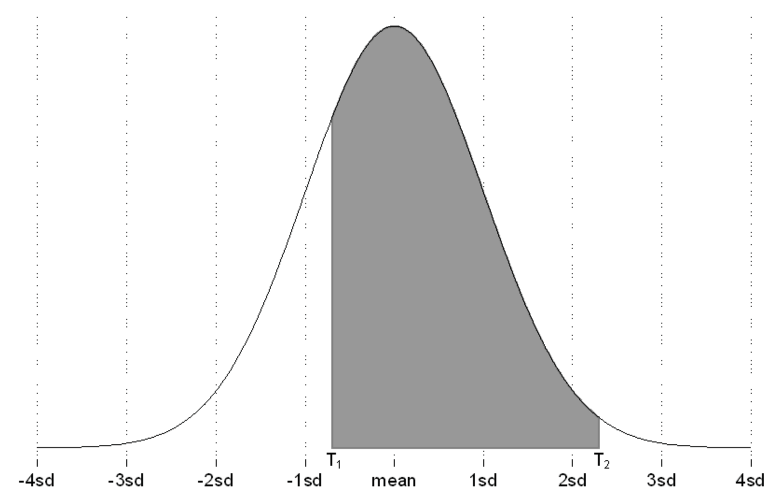
To compute the area less than τ_1, we use the function $\Phi(\tau_1)$ or $\Lambda(\tau_1)$



For the area less than τ_2, we use the function $\Phi(\tau_2)$ or $\Lambda(\tau_2)$



For the area between τ_1 and τ_2, we use $\Phi(\tau_1)-\Phi(\tau_2)$ or $\Lambda(\tau_1)-\Lambda(\tau_2)$



For example

$y = 2$ when y^* is between τ_1 and τ_2:

$$\begin{aligned}\Pr(y_i = 2 \mid \mathbf{x}_i) &= \Pr(\tau_1 \le y_i^* < \tau_2 \mid \mathbf{x}_i) \\ &= \Pr(\tau_1 \le \mathbf{x}_i\beta + \varepsilon_i < \tau_2 \mid \mathbf{x}_i) \\ &= \Pr(\tau_1 - \mathbf{x}_i\beta \le \varepsilon_i < \tau_2 - \mathbf{x}_i\beta \mid \mathbf{x}_i) \\ &= \Pr(\varepsilon_i < \tau_2 - \mathbf{x}_i\beta \mid \mathbf{x}_i) - \Pr(\varepsilon_i \le \tau_1 - \mathbf{x}_i\beta \mid \mathbf{x}_i) \\ &= F(\tau_2 - \mathbf{x}_i\beta) - F(\tau_1 - \mathbf{x}_i\beta)\end{aligned}$$

A general formula

$$\Pr(y_i = m \mid \mathbf{x}_i) = F(\tau_m - \mathbf{x}_i\beta) - F(\tau_{m-1} - \mathbf{x}_i\beta)$$

For example

Four outcome categories:

$$\begin{aligned}\Pr(y_i = 1 \mid x_i) &= \Phi(\tau_1 - \alpha - \beta x_i) - \Phi(\tau_0) \\ &= \Phi(\tau_1 - \alpha - \beta x_i) - \Phi(-\infty) \\ &= \Phi(\tau_1 - \alpha - \beta x_i) - 0\end{aligned}$$

$$\Pr(y_i = 2 \mid x_i) = \Phi(\tau_2 - \alpha - \beta x_i) - \Phi(\tau_1 - \alpha - \beta x_i)$$

$$\Pr(y_i = 3 \mid x_i) = \Phi(\tau_3 - \alpha - \beta x_i) - \Phi(\tau_2 - \alpha - \beta x_i)$$

$$\begin{aligned}\Pr(y_i = 4 \mid x_i) &= \Phi(\tau_4) - \Phi(\tau_3 - \alpha - \beta x_i) \\ &= \Phi(\infty) - \Phi(\tau_3 - \alpha - \beta x_i) \\ &= 1 - \Phi(\tau_3 - \alpha - \beta x_i)\end{aligned}$$

A Question(for you)

Can you replicate these predicted probabilities?

Given,

$\alpha = -.50$, $\beta = .052$, $\tau_1 = .75$, $\tau_2 = 3.5$, and $\tau_3 = 5.0$

	Predicted probabilities		
	$x = 15$	$x = 40$	$x = 80$
$\Pr(y = 1 \mid x)$	0.68	0.20	0.00
$\Pr(y = 2 \mid x)$	0.32	0.77	0.44
$\Pr(y = 3 \mid x)$	0.00	0.03	0.47
$\Pr(y = 4 \mid x)$	0.00	0.00	0.09

Identification

Since y^* is latent, its mean and variance cannot be estimated

The variance is identified by assuming the $Var(\varepsilon \mid \mathbf{x})$

Distributional Assumptions

Ordered probit: $\varepsilon \sim N(0,1)$

Ordered logit: $\varepsilon \sim \lambda(0,\pi^2/3)$

- cdf for normal with mean $\mu = 0$ and $\sigma^2 = 1$

$$\Phi(\varepsilon) = \int_{-\infty}^{\varepsilon} \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{t^2}{2}\right) dt$$

- cdf for logistic with mean $\mu = 0$ and $\sigma^2 = \pi^2/3$

$$\Lambda(\varepsilon) = \frac{\exp(\varepsilon)}{1+\exp(\varepsilon)}$$

Question (for you)

While the above assumptions identity the variance, the mean of y* is still unidentified

What are the consequences of this?

The mean of y* is unidentified (an illustration)

Start with the formula for the predicted probability:

$$\Pr(y = m \mid x) = F(\tau_m - \alpha - \beta x) - F(\tau_{m-1} - \alpha - \beta x)$$

Now add zero $0 = \delta - \delta$ and rearrange:

$$\Pr(y = m = \mid x) = F([\delta - \delta] + \tau_m - \alpha - \beta x) - F([\delta - \delta] + \tau_{m-1} - \alpha - \beta x)$$

$$= F([\tau_m - \delta] - [\alpha - \delta] - \beta x) - F([\tau_{m-1} - \delta] - [\alpha - \delta] - \beta x)$$

Relabel:

$$\Pr(y = m \mid x) = F(\tau_m^* - \alpha^* - \beta x) - F(\tau_{m-1}^* - \alpha^* - \beta x)$$

Question (for you)

We now have,

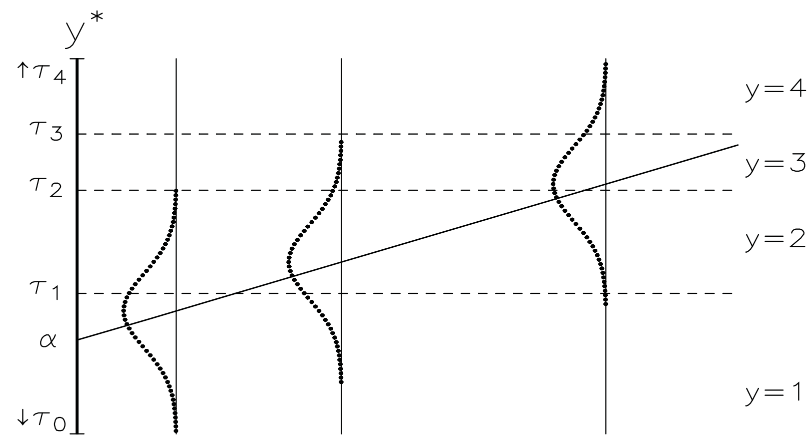
- a true model with parameters τ_m, α, and β
- and an "imposter" with τ_m^*, α^*, and β

Will both give identical predictions?

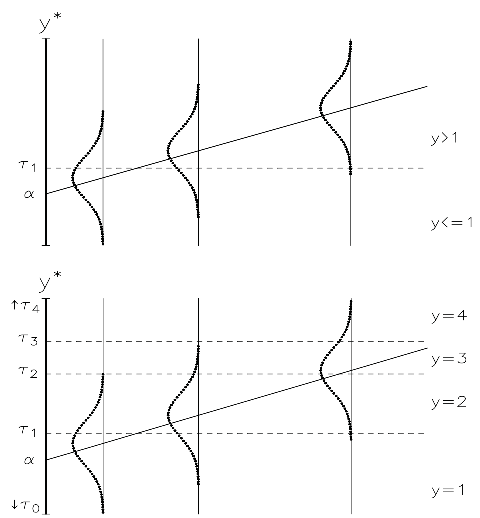
What assumption will identify the model?

The Structural Model

$y_i^* = \mathbf{x}_i\beta + \varepsilon_i$



Adding or deleting "cutpoints"



Estimation

The probability of a given outcome is:

$$\Pr(y_i = m \mid \mathbf{x}_i, \beta, \tau) = F(\tau_m - \mathbf{x}_i\beta) - F(\tau_{m-1} - \mathbf{x}_i\beta)$$

The probability of whatever is observed for the *i*th case is:

$$p_i = \begin{cases} \Pr(y_i = 1 \mid \mathbf{x}_i, \beta, \tau) & \text{if } y = 1 \\ \vdots & \vdots \\ \Pr(y_i = m \mid \mathbf{x}_i, \beta, \tau) & \text{if } y = m \\ \vdots & \vdots \\ \Pr(y_i = J \mid \mathbf{x}_i, \beta, \tau) & \text{if } y = J \end{cases}$$

If the observations are independent, then

$$L(\beta,\tau \mid \mathbf{y},\mathbf{X}) = \prod_{i=1}^{N} p_i = \prod_{j=1}^{J} \prod_{y_i=j} \left[F\left(\tau_j - \mathbf{x}_i\beta\right) - F\left(\tau_{j-1} - \mathbf{x}_i\beta\right) \right]$$

Software Issues

Parameterizations of the Model

You must know which parameterization is being used

Methods of Numerical Maximization

Different methods of numerical maximization produce different test statistics

Failure to Converge

The ORM takes longer to converge than most models; 5-10 iterations is typical

If the number of cases in a response category is small, the model may fail to converge

When this occurs, you can merge the smallest category into an adjacent category

Question (for you)

Will combining adjacent categories result in

- biased parameter estimates?
- a loss of efficiency?

Data

1977 and 1989 General Social Survey

The sample consists of 2,293 randomly sampled non-institutionalized adults

Outcome

Respondents asked to evaluate the statement:

"A working mother can establish just as warm and secure a relationship with her child as a mother who does not work."

Outcome categories were:

1=Strongly Disagree with the statement

2=Disagree with the statement

3=Agree with the statement

4=Strongly Agree with the statement

Descriptive Information

```
. spex ordwarm3

. use "http://www.indiana.edu/~jslsoc/stata/spex_data/ordwarm3.dta", clear

. codebook warm yr89 male age ed prst, compact

Variable     Obs Unique      Mean  Min  Max  Label
-------------------------------------------------------------------------------
warm        2293      4  2.607501    1    4  Working mom can have warm relations w
                                                 child?
yr89        2293      2  .3986044    0    1  Survey year: 1=1989 0=1977
male        2293      2  .4648932    0    1  Gender: 1=male 0=female
age         2293     72  44.93546   18   89  Age in years
ed          2293     21  12.21805    0   20  Years of education
prst        2293     58  39.58526   12   82  Occupational prestige
-------------------------------------------------------------------------------
```



```
. sum warm yr89 male age ed prst

    Variable |       Obs        Mean    Std. Dev.       Min        Max
-------------+--------------------------------------------------------
        warm |      2293    2.607501    .9282156          1          4
        yr89 |      2293    .3986044    .4897178          0          1
        male |      2293    .4648932    .4988748          0          1
         age |      2293    44.93546    16.77903         18         89
          ed |      2293    12.21805    3.160827          0         20
-------------+--------------------------------------------------------
        prst |      2293    39.58526    14.49226         12         82

. tab warm

Working mom |
   can have |
       warm |
relations w |
     child? |      Freq.     Percent        Cum.
------------+-----------------------------------
        1SD |        297       12.95       12.95
         2D |        723       31.53       44.48
         3A |        856       37.33       81.81
        4SA |        417       18.19      100.00
------------+-----------------------------------
      Total |      2,293      100.00
```

Descriptive Table

Name	Mean	Std Dev	Min	Max	Description
WARM			1.00	4.00	1=SD; 2=D; 3=A; 4=SA
YR89	0.40		0.00	1.00	Survey Year: 1=1989; 0=1977
MALE	0.47		0.00	1.00	1=male; 0=female
AGE	44.94	16.78	18.00	89.00	Age in years
ED	12.22	3.16	0.00	29.00	Years of education
PRST	39.59	14.49	12.00	82.00	Occupational prestige

Note: N=2,293. WARM has categories: 1=SD; 2=D; 3=A; 4=SA with marginal percentages 13, 32, 37, and 18, respectively.

Estimating the Ordinal Logit Model

```
. ologit warm yr89 male age ed prst

Iteration 0:   log likelihood = -2995.7704
Iteration 1:   log likelihood = -2851.8105
Iteration 2:   log likelihood = -2850.3921
Iteration 3:   log likelihood = -2850.3904
Iteration 4:   log likelihood = -2850.3904

Ordered logistic regression                       Number of obs   =       2293
                                                  LR chi2(5)      =     290.76
                                                  Prob > chi2     =     0.0000
Log likelihood = -2850.3904                       Pseudo R2       =     0.0485

------------------------------------------------------------------------------
        warm |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
        yr89 |   .5311259    .079811     6.65   0.000     .3746993    .6875525
        male |  -.7406042   .0784181    -9.44   0.000    -.8943008   -.5869076
         age |  -.0222286   .0024621    -9.03   0.000    -.0270543    -.017403
          ed |   .0630341   .0159337     3.96   0.000     .0318046    .0942636
        prst |   .0054823    .003288     1.67   0.095    -.0009622    .0119267
-------------+----------------------------------------------------------------
       /cut1 |  -2.217185   .2266492                     -2.661409   -1.772961
       /cut2 |  -.3868419   .2213616                     -.8207027     .047019
       /cut3 |    1.49904   .2230236                      1.061922    1.936159
------------------------------------------------------------------------------
```


Estimating the Ordinal Probit Model

```
. oprobit warm yr89 male age ed prst

Iteration 0:   log likelihood = -2995.7704
Iteration 1:   log likelihood = -2854.0632
Iteration 2:   log likelihood = -2853.9365
Iteration 3:   log likelihood = -2853.9365

Ordered probit regression                         Number of obs   =       2293
                                                  LR chi2(5)      =     283.67
                                                  Prob > chi2     =     0.0000
Log likelihood = -2853.9365                       Pseudo R2       =     0.0473

------------------------------------------------------------------------------
        warm |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
        yr89 |   .3242996   .0468088     6.93   0.000     .2325561    .4160431
        male |  -.4214407   .0455103    -9.26   0.000    -.5106392   -.3322422
         age |  -.0125318    .001439    -8.71   0.000    -.0153522   -.0097113
          ed |   .0366577   .0093007     3.94   0.000     .0184286    .0548867
        prst |   .0029157   .0019209     1.52   0.129    -.0008493    .0066806
-------------+----------------------------------------------------------------
       /cut1 |  -1.280546   .1310568                     -1.537413   -1.023679
       /cut2 |  -.2150214   .1293864                     -.4686142    .0385713
       /cut3 |    .910862    .129861                      .6563392    1.165385
------------------------------------------------------------------------------
```


Comparing Models

```
. quietly ologit warm yr89 male age ed prst
. eststo ologit
. quietly oprobit warm yr89 male age ed prst
. eststo oprobit
```

```
. esttab ologit oprobit  , ///
>         mtitle(OLogit OProbit) b(%9.2f) wide nostar noparen nogaps nonumbers

----------------------------------------------------------------
                   OLogit                    OProbit
----------------------------------------------------------------
warm
yr89                 0.53         6.65         0.32         6.93
male                -0.74        -9.44        -0.42        -9.26
age                 -0.02        -9.03        -0.01        -8.71
ed                   0.06         3.96         0.04         3.94
prst                 0.01         1.67         0.00         1.52
----------------------------------------------------------------
cut1
_cons               -2.22        -9.78        -1.28        -9.77
----------------------------------------------------------------
cut2
_cons               -0.39        -1.75        -0.22        -1.66
----------------------------------------------------------------
cut3
_cons                1.50         6.72         0.91         7.01
----------------------------------------------------------------
N                    2293                      2293
----------------------------------------------------------------
t statistics in second column
```


The OLM in Terms of Odds Ratios

The cumulative probability that the outcome is less than or equal to m is:

$$\Pr(y \le m \mid \mathbf{x}) = \sum_{j=1}^{m} \Pr(y = j \mid \mathbf{x}) = F(\tau_m - \mathbf{x}\beta) \quad \text{for } m = 1, J-1$$

The odds that an outcome is less than or equal to m is:

$$\Omega_m(\mathbf{x}) = \frac{\Pr(y \le m \mid \mathbf{x})}{\Pr(y > m \mid \mathbf{x})} = \exp(\tau_m - \mathbf{x}\beta)$$

To determine the effect of a unit change in x_k:

$$\frac{\Omega_m(\mathbf{x}, x_k + 1)}{\Omega_m(\mathbf{x}, x_k)} = \frac{\exp(\tau_m - \mathbf{x}\beta, x_k + 1)}{\exp(\tau_m - \mathbf{x}\beta, x_k)} = \exp(-\beta_k)$$

Questions (for you)

In the BLM the factor change equals $\exp(\beta_k)$

In the OLM the factor change equals $\exp(-\beta_k)$

Why?

Why do you think this model is also known as the *proportional odds model*?

How do both impact interpretation?

For a unit increase in x_k, the odds of lower outcomes compared to higher outcomes change by the factor $\exp(-\beta_k)$, holding all other variables constant

For a unit increase in x_k, the odds of higher outcomes compared to lower outcomes change by the factor $\exp(\beta_k)$, holding all other variables constant

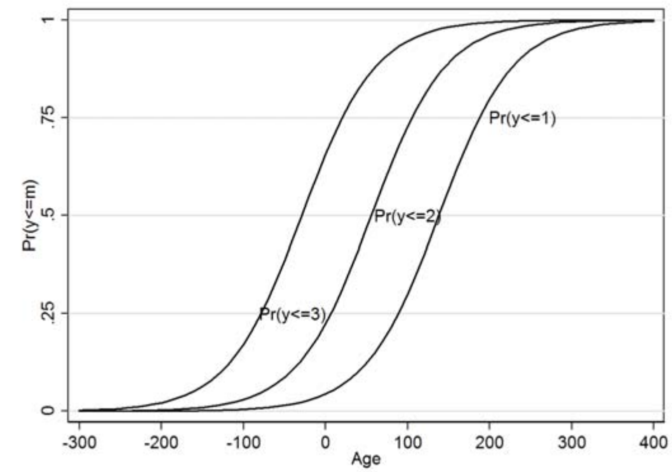
The Proportional Odds Assumption

or more generally, the Parallel Regression Assumption

The equality of odds ratios for any comparison of higher versus lower outcomes

The cumulative probability curve are parallel

Graphically



An Informal Test of the Parallel Regression Assumption

To illustrate the test, recode the dependent variable into J-1 dummy variables

Estimate 3 binary regressions:

$\Pr(y \le 1 \mid \mathbf{x}) = F(\mathbf{x}\beta_1)$

$\Pr(y \le 2 \mid \mathbf{x}) = F(\mathbf{x}\beta_2)$

$\Pr(y \le 3 \mid \mathbf{x}) = F(\mathbf{x}\beta_3)$

If the assumption of parallel regressions is true, then

$\beta_1 = \beta_2 = \beta_3$


```
. gen warm1 = warm <= 1 if warm<.
. label variable warmlt2 "1=SD; 0=D,A,SA"

. gen warm2 = warm <= 2 if warm<.
. label variable warmlt3 "1=SD,D; 0=A,SA"

. gen warm3 = warm <= 3 if warm<.
. label variable warmlt4 "1=SD,D,A; 0=SA"

. qui logit warm1 yr89 male age ed prst, nolog
. eststo warm1

. qui logit warm2 yr89 male age ed prst, nolog
. eststo warm2

. qui logit warm3 yr89 male age ed prst, nolog
. eststo warm3
```

```
. esttab warm*, b(%9.2f) nonumber not mtitles(warm1 warm2 warm3)

--------------------------------------------------------------
                  warm1           warm2           warm3
--------------------------------------------------------------
main
yr89             -0.98***        -0.57***        -0.33**
male              0.31*           0.70***         1.09***
age               0.02***         0.03***         0.02***
ed               -0.10***        -0.05**         -0.06*
prst              0.00           -0.01*          -0.00
_cons            -1.52***        -0.53*           1.29***
--------------------------------------------------------------
N                 2293            2293            2293
--------------------------------------------------------------
* p<0.05, ** p<0.01, *** p<0.001
```


Interpretation

Consider:

Factor Change in Odds (for ordinal logit model only)

Predicted Probabilities

- for Observed Data
- for Ideal Types
- at Selected Values
- over a Range of a Variable

Discrete Change in Predicted Probability

Partial or Marginal Change in Predicted Probability

Computing Factor Change in Odds

```
. ologit warm yr89 male age ed prst
. listcoef, help

ologit (N=2293): Factor Change in Odds

  Odds of: >m vs <=m

----------------------------------------------------------------------
        warm |      b         z     P>|z|    e^b    e^bStdX      SDofX
-------------+--------------------------------------------------------
        yr89 |   0.53113    6.655   0.000   1.7008   1.2971     0.4897
        male |  -0.74060   -9.444   0.000   0.4768   0.6911     0.4989
         age |  -0.02223   -9.028   0.000   0.9780   0.6887    16.7790
          ed |   0.06303    3.956   0.000   1.0651   1.2205     3.1608
        prst |   0.00548    1.667   0.095   1.0055   1.0827    14.4923
----------------------------------------------------------------------
       b = raw coefficient
       z = z-score for test of b=0
   P>|z| = p-value for z-test
     e^b = exp(b) = factor change in odds for unit increase in X
 e^bStdX = exp(b*SD of X) = change in odds for SD increase in X
   SDofX = standard deviation of X
```


Interpreting Factor Change

```
ologit (N=2293): Factor Change in Odds

  Odds of: >m vs <=m

----------------------------------------------------------------------
        warm |      b         z     P>|z|    e^b    e^bStdX      SDofX
-------------+--------------------------------------------------------
        yr89 |   0.53113    6.655   0.000   1.7008   1.2971     0.4897
        male |  -0.74060   -9.444   0.000   0.4768   0.6911     0.4989
         age |  -0.02223   -9.028   0.000   0.9780   0.6887    16.7790
          ed |   0.06303    3.956   0.000   1.0651   1.2205     3.1608
        prst |   0.00548    1.667   0.095   1.0055   1.0827    14.4923
----------------------------------------------------------------------
```

A standard deviation increase in years of education (about 3 years) increases the odds of being more supportive of working women by a factor of 1.22 (z=3.96, p<.001), holding all other variables constant.

Computing Percent Change

```
. listcoef, help percent

ologit (N=2293): Percentage Change in Odds

  Odds of: >m vs <=m

----------------------------------------------------------------------
        warm |       b          z      P>|z|       %      %StdX      SDofX
-------------+--------------------------------------------------------
        yr89 |   0.53113    6.655   0.000     70.1      29.7     0.4897
        male |  -0.74060   -9.444   0.000    -52.3     -30.9     0.4989
         age |  -0.02223   -9.028   0.000     -2.2     -31.1    16.7790
          ed |   0.06303    3.956   0.000      6.5      22.0     3.1608
        prst |   0.00548    1.667   0.095      0.5       8.3    14.4923
----------------------------------------------------------------------
       b = raw coefficient
       z = z-score for test of b=0
   P>|z| = p-value for z-test
       % = percent change in odds for unit increase in X
   %StdX = percent change in odds for SD increase in X
   SDofX = standard deviation of X
```


Interpreting Percent Change

```
ologit (N=2293): Percentage Change in Odds

  Odds of: >m vs <=m

----------------------------------------------------------------------
        warm |       b          z      P>|z|       %      %StdX      SDofX
-------------+--------------------------------------------------------
        yr89 |   0.53113    6.655   0.000     70.1      29.7     0.4897
        male |  -0.74060   -9.444   0.000    -52.3     -30.9     0.4989
         age |  -0.02223   -9.028   0.000     -2.2     -31.1    16.7790
          ed |   0.06303    3.956   0.000      6.5      22.0     3.1608
        prst |   0.00548    1.667   0.095      0.5       8.3    14.4923
----------------------------------------------------------------------
```

The odds of being more supportive of working women are 52% less for men (z=-9.44, p<.011), holding all other variables constant.

Approaches to interpretation using Pr(y=1)

Predicted probabilities for observed data `[predict]`

Predicted probabilities for ideal types `[prvalue]`

Tables of predicted probabilities `[prtab]`

Plotting predicted probabilities `[prgen]`

Discrete change in predicted probabilities `[prchange & prvalue]`

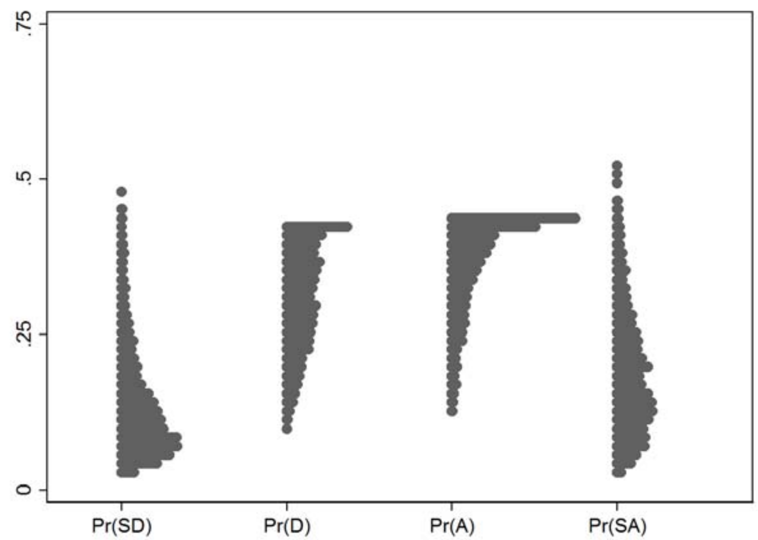
Marginal change `[prchange]`

Predicted Probabilities

The predicted probability that $y = m$ given $\mathbf{x}$ is:

$$\widehat{\Pr}\left(y = m \mid \mathbf{x}\right) = F\left(\hat{\tau}_m - \mathbf{x}\hat{\boldsymbol{\beta}}\right) - F\left(\hat{\tau}_{m-1} - \mathbf{x}\hat{\boldsymbol{\beta}}\right)$$

Dot plot of Predicted Probabilities for Observed Data



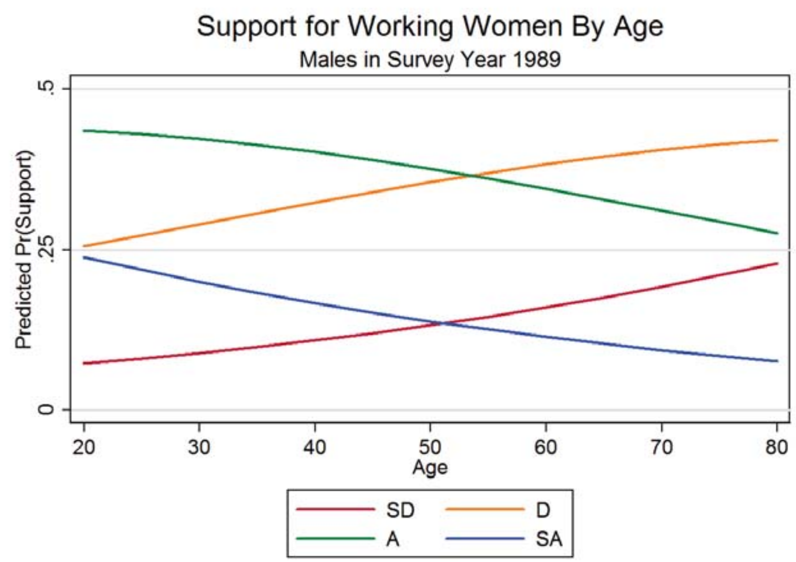

```
. ologit warm yr89 male age ed prst
. predict pry1 pry2 pry3 pry4

. label var pry1 "Pr(SD)"
. label var pry2 "Pr(D)"
. label var pry3 "Pr(A)"
. label var pry4 "Pr(SA)"

. dotplot pry1 pry2 pry3 pry4, ylabel(0(.25).75)
. graph export orm-02-dotplot.png , width(1200) replace
```


Table of Predicted Probabilities

1977	SD	D	A	SA
Men	0.19	0.40	0.32	0.10
Women	0.10	0.31	0.41	0.18
Men-Women	0.09	0.09	-0.09	-0.08
1989	SD	D	A	SA
Men	0.12	0.34	0.39	0.15
Women	0.06	0.23	0.44	0.27
Men-Women	0.06	0.11	-0.05	-0.12
Change from 1977 to 1989				
Men	-0.07	-0.06	0.07	0.05
Women	-0.04	-0.08	0.03	0.09


```
. prtab yr89 male, rest(mean) novarlbl

ologit: Predicted probabilities for warm

Predicted probability of outcome 1 (1SD)

----------------------------
          |        male
     yr89 | 0Female    1Male
----------+-----------------
   0_1977 |  0.0993   0.1877
   1_1989 |  0.0608   0.1196
----------------------------

Predicted probability of outcome 2 (2D)

----------------------------
          |        male
     yr89 | 0Female    1Male
----------+-----------------
   0_1977 |  0.3080   0.4026
   1_1989 |  0.2269   0.3391
----------------------------

::output deleted::

         yr89        male         age          ed        prst
x=  .39860445   .46489315   44.935456   12.218055   39.585259
```

Plotting Predicted Probabilities




```
. prgen age, x(male=1 yr89=1) from(20) to(80) generate(m89) gap(5) rest(mean)

ologit: Predicted values as age varies from 20 to 80.

         yr89       male        age         ed       prst
x=          1          1  44.935456  12.218055  39.585259

. label var m89p1 "SD"
. label var m89p2 "D"
. label var m89p3 "A"
. label var m89p4 "SA"

. codebook m89p* , compact

Variable   Obs Unique      Mean       Min       Max  Label
-------------------------------------------------------------------------------
m89p1       13     13  .1392124   .072406  .2285378  SD
m89p2       13     13  .3480605  .2549954  .4202596  D
m89p3       13     13  .3674588  .2753192  .4350054  A
m89p4       13     13  .1452682  .0758835  .2375932  SA
-------------------------------------------------------------------------------
```

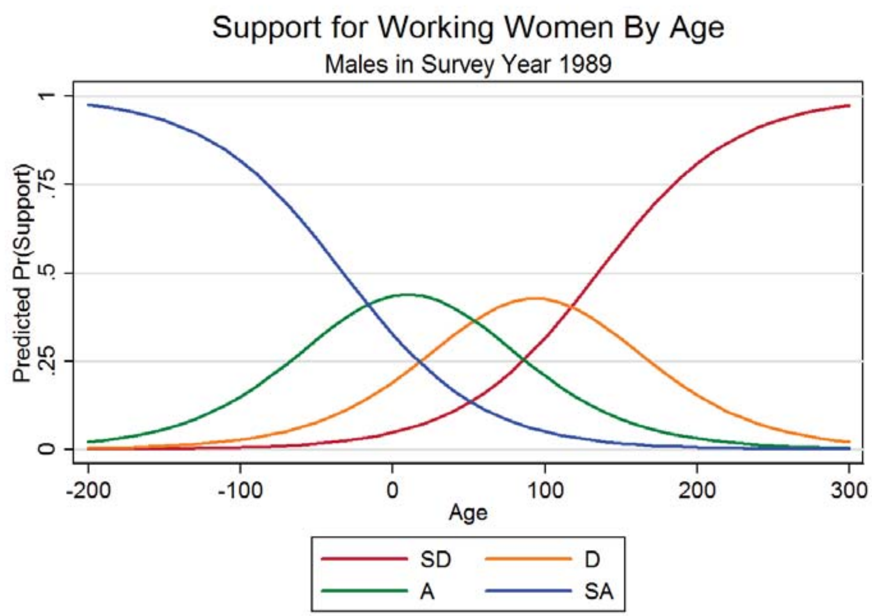


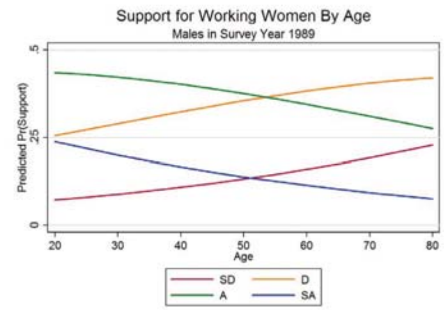
```
. graph twoway connected m89p1 m89p2 m89p3 m89p4 m89x ///
>         , title("Support for Working Women By Age") ///
>         subtitle(Males in Survey Year 1989) ///
>         xtitle("Age") ///
>         ytitle("Predicted Pr(Support)") ///
>         xlabel(20(10)80) ylabel(0(.25).50, grid) ///
>         msymbol(O D S T) ///
>         saving(orm-m89.gph,replace)

. graph export orm-04-predprob-agew89.png , width(1200) replace
```

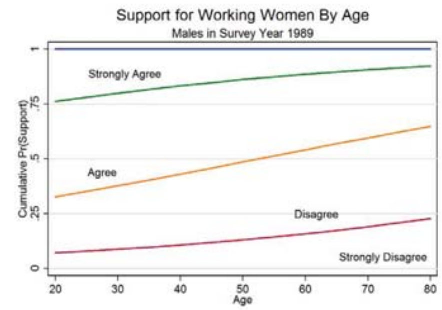

Question (for you)

Why do we have crossing lines?

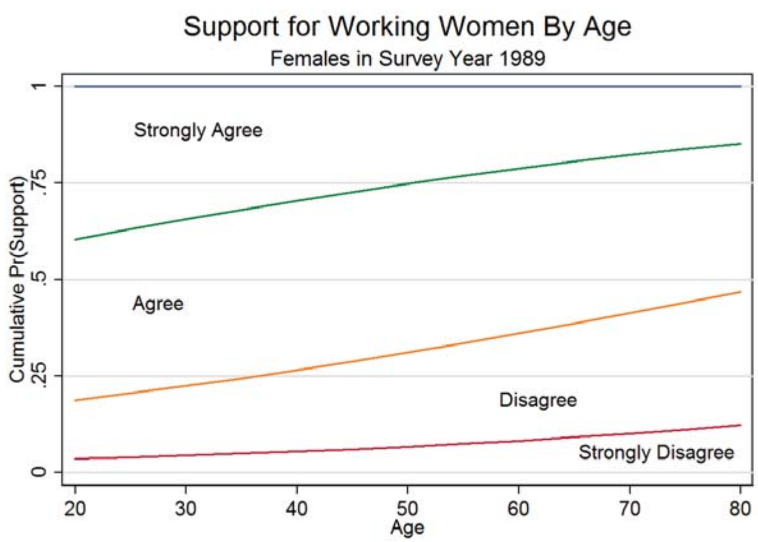
What happened to the parallel curves?





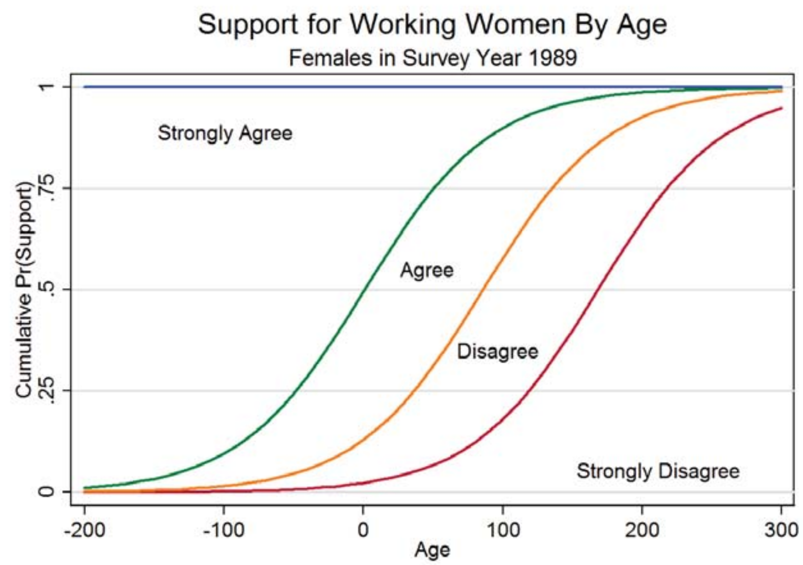


Plotting Predicted Cumulative Probabilities

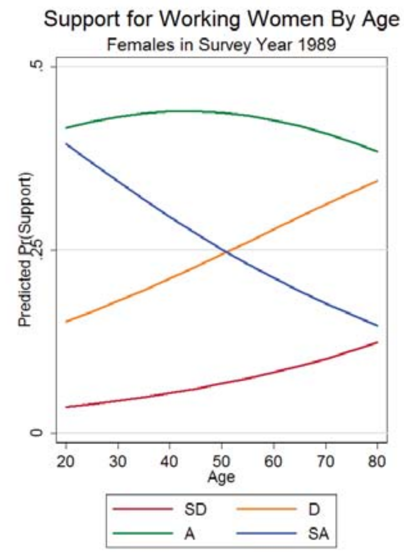


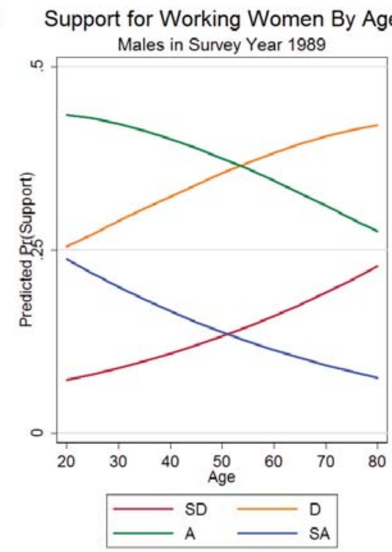

```
. label var w89s1 "SD"
. label var w89s2 "SD or D"
. label var w89s3 "SD, D or A"

. codebook w89s* , compact

Variable   Obs Unique      Mean        Min        Max  Label
--------------------------------------------------------------------------------
w89s1       13     13  .0723478  .0358844  .1237714  SD
w89s2       13     13  .3180948  .1883808  .4683307  SD or D
w89s3       13     13  .7409536  .6047547  .8530888  SD, D or A
w89s4       13      2         1  .9999999         1  pr(y<=4)
--------------------------------------------------------------------------------

. graph twoway connected w89s1 w89s2 w89s3 w89s4 w89x ///
>         , title("Support for Working Women By Age") ///
>         subtitle(Females in Survey Year 1989) ///
>         xtitle("Age") ///
>         ytitle("Cumulative Pr(Support)") ///
>         xlabel(20(10)80) ylabel(0(.25)1, grid) ///
>         msymbol(none none none none) ///
>         text(.89 25 "Strongly Agree",place(e)) ///
>         text(.44 25 "Agree",place(e)) ///
>         text(.19 58 "Disagree",place(e)) ///
>         text(.055 65 "Strongly Disagree",place(e)) ///
>         , legend(off)

. graph export orm-06-cumulativeprob-agew89.png , width(1200) replace
```

Parallel Lines



Comparing Men and Women





Discrete change in predicted probabilities

Discrete change is defined as:

$$\frac{\Delta \Pr(y=m \mid \mathbf{x})}{\Delta x_k} = \Pr(y=m \mid \mathbf{x}, x_k = x_E) - \Pr(y=m \mid \mathbf{x}, x_k = x_S)$$

Interpretation

When x_k changes from x_S to x_E, the predicted probability of outcome m changes by $\frac{\Delta \Pr(y=m \mid \mathbf{x})}{\Delta x_k}$, holding all other variables at $\mathbf{x}$

Question (for you)

What impacts the discrete change?

Computing Discrete Change Coefficients

```
. prchange, rest(mean)

ologit: Changes in Probabilities for warm

yr89
        Avg|Chg|         1SD          2D          3A         4SA
0->1   .06459879  -.05080348  -.07839411   .05431128   .07488631

male
        Avg|Chg|         1SD          2D          3A         4SA
0->1   .09050236   .07567564   .10532907  -.08177939  -.09922536

age
            Avg|Chg|         1SD          2D          3A         4SA
Min->Max   .18753199   .18630873   .18875524   -.1824809  -.19258308
   -+1/2   .00273808   .00220662   .00326955  -.00245824   -.0030179
  -+sd/2   .04581587   .03711103   .05452073  -.04094052  -.05069122
MargEfct   .00273811    .0022066   .00326962  -.00245833  -.00301789

ed
            Avg|Chg|         1SD          2D          3A         4SA
Min->Max   .15220216  -.14217279   -.1622315   .14657542   .15782891
   -+1/2   .00776388  -.00625769  -.00927007    .0069696   .00855814
  -+sd/2   .02452281  -.01979138  -.02925426   .02198741   .02705821
MargEfct   .00776449  -.00625728  -.00927171   .00697111   .00855787
```



```
prst
            Avg|Chg|         1SD          2D          3A         4SA
Min->Max   .04688536  -.03700081  -.05676991   .04017711   .05359362
   -+1/2    .0006753  -.00054421  -.00080639    .0006063   .00074431
  -+sd/2   .00978541  -.00788773  -.01168308   .00878367   .01078716
MargEfct    .0006753  -.00054421  -.00080639    .0006063    .0007443

               1SD          2D          3A         4SA
Pr(y|x)  .11175791   .32789621   .39833155   .16201432

           yr89      male       age        ed      prst
   x=   .398604   .464893   44.9355   12.2181   39.5853
sd_x=   .489718   .498875    16.779   3.16083   14.4923
```


Average Change

Average absolute discrete change

$$\overline{\Delta} = \frac{1}{J}\sum_{j=1}^{J}\left|\frac{\Delta \Pr\left(y = j \mid \overline{\mathbf{x}}\right)}{\Delta x_k}\right|$$

Interpretation (you try)

Adding CI to the DC

1977	SD	D	A	SA
Men	0.19	0.40	0.31	0.10
Women	0.10	0.31	0.41	0.18
Men-Women	0.09*	0.09*	-0.10*	-0.09*
1989	**SD**	**D**	**A**	**SA**
Men	0.12	0.34	0.39	0.15
Women	0.06	0.23	0.44	0.27
Men-Women	0.06*	0.11*	-0.05*	-0.12*
Change from 1977 to 1989				
Men	-0.07*	-0.06*	0.08*	0.06*
Women	-0.04*	-0.08*	0.03*	0.09*

* significant at the .05-level


```
. //gender difference for 77
. quietly prvalue, x(yr89=0 male=0) rest(mean) save
. prvalue, x(yr89=0 male=1) rest(mean) dif

ologit: Change in Predictions for warm

Confidence intervals by delta method

                 Current     Saved    Change    95% CI for Change
    Pr(y=1SD|x):  0.1877    0.0993    0.0885  [ 0.0687,   0.1082]
    Pr(y=2D|x):   0.4026    0.3080    0.0946  [ 0.0734,   0.1158]
    Pr(y=3A|x):   0.3144    0.4119   -0.0975  [-0.1192,  -0.0758]
    Pr(y=4SA|x):  0.0952    0.1808   -0.0856  [-0.1045,  -0.0667]

              yr89       male        age         ed       prst
Current=         0          1  44.935456  12.218055  39.585259
  Saved=         0          0  44.935456  12.218055  39.585259
   Diff=         0          1          0          0          0
```



```
. //gender difference for 89
. quietly prvalue, x(yr89=1 male=0) rest(mean) save
. prvalue, x(yr89=1 male=1) rest(mean) dif

ologit: Change in Predictions for warm

Confidence intervals by delta method

                 Current     Saved    Change    95% CI for Change
    Pr(y=1SD|x):  0.1196    0.0608    0.0588  [ 0.0444,   0.0731]
    Pr(y=2D|x):   0.3391    0.2269    0.1121  [ 0.0885,   0.1358]
    Pr(y=3A|x):   0.3895    0.4393   -0.0498  [-0.0675,  -0.0320]
    Pr(y=4SA|x):  0.1518    0.2730   -0.1211  [-0.1468,  -0.0954]

              yr89       male        age         ed       prst
Current=         1          1  44.935456  12.218055  39.585259
  Saved=         1          0  44.935456  12.218055  39.585259
   Diff=         0          1          0          0          0
```



```
. //year difference for males
. quietly prvalue, x(yr89=0 male=1) rest(mean) save
. prvalue, x(yr89=1 male=1) rest(mean) dif

ologit: Change in Predictions for warm

Confidence intervals by delta method

                 Current     Saved    Change    95% CI for Change
    Pr(y=1SD|x):  0.1196    0.1877   -0.0681  [-0.0881,  -0.0481]
    Pr(y=2D|x):   0.3391    0.4026   -0.0636  [-0.0844,  -0.0427]
    Pr(y=3A|x):   0.3895    0.3144    0.0751  [ 0.0530,   0.0971]
    Pr(y=4SA|x):  0.1518    0.0952    0.0566  [ 0.0384,   0.0748]

              yr89       male        age         ed       prst
Current=         1          1  44.935456  12.218055  39.585259
  Saved=         0          1  44.935456  12.218055  39.585259
   Diff=         1          0          0          0          0
```

```
. //year difference for females
. quietly prvalue, x(yr89=0 male=0) rest(mean) save
. prvalue, x(yr89=1 male=0) rest(mean) dif

ologit: Change in Predictions for warm

Confidence intervals by delta method

                     Current     Saved    Change   95% CI for Change
     Pr(y=1SD|x):     0.0608    0.0993   -0.0384  [-0.0502,  -0.0266]
     Pr(y=2D|x):      0.2269    0.3080   -0.0811  [-0.1052,  -0.0571]
     Pr(y=3A|x):      0.4393    0.4119    0.0274  [ 0.0157,   0.0391]
     Pr(y=4SA|x):     0.2730    0.1808    0.0921  [ 0.0641,   0.1202]

               yr89       male        age         ed       prst
Current=          1          0  44.935456  12.218055  39.585259
  Saved=          0          0  44.935456  12.218055  39.585259
   Diff=          1          0          0          0          0
```


Interpretation (you try)

Testing parallel regression assumption

```
. brant, detail

Estimated coefficients from j-1 binary regressions

              y>1         y>2         y>3
 yr89    .9787113   .57319466   .33257816
 male  -.31344341  -.69708427  -1.0924897
  age  -.01719524  -.02577173  -.01910713
   ed   .09796804   .05016246   .05555507
 prst  -.00255688   .00899022   .00487506
_cons   1.5202206   .52724985  -1.2923037

Brant Test of Parallel Regression Assumption

    Variable |      chi2   p>chi2    df
-------------+--------------------------
         All |     47.67    0.000    10
-------------+--------------------------
        yr89 |     13.11    0.001     2
        male |     22.38    0.000     2
         age |      7.06    0.029     2
          ed |      3.75    0.154     2
        prst |      4.89    0.087     2
----------------------------------------

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


When the assumption of parallel regressions is rejected, alternative models for nominal outcomes should be considered

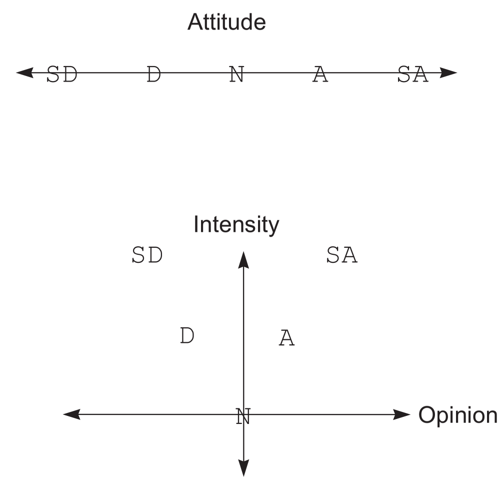
When is a Variable Ordinal?

Simply because the values of a variable *can* be ordered does not imply that the variable *should* be analyzed as ordinal

A variable can be ordered for one purpose, but be unordered or ordered differently for another purpose

For example, occupational groupings can be ordered both by the status of the occupation and by the income of the occupation (Miller and Volker, 1985)

Likert scales might reflect two dimensions

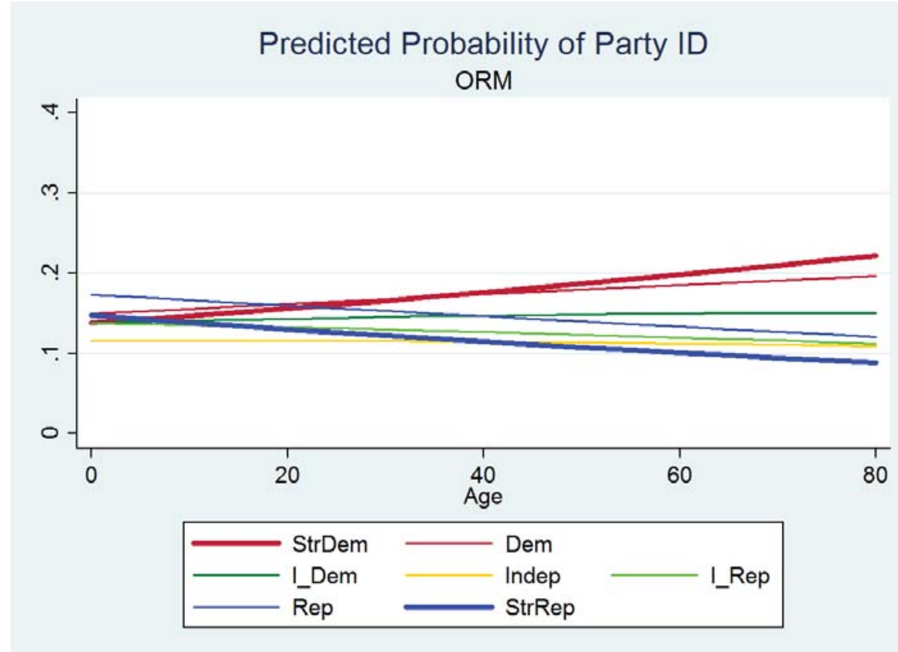


When the proper ordering is ambiguous, models for nominal variables should be considered

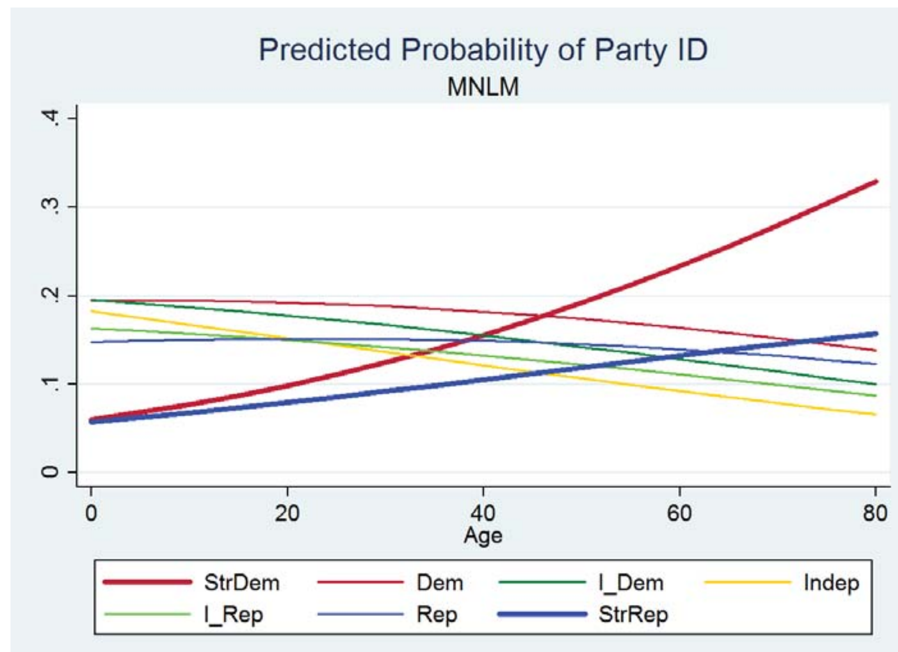
Data

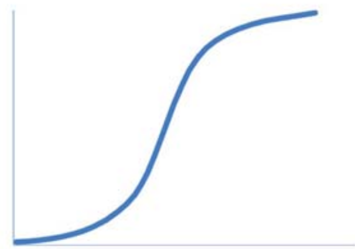
```
. codebook , compact

Variable    Obs Unique      Mean  Min     Max  Label
-------------------------------------------------------------------------------
partyid    2250      7  3.709333    1       7  Party ID
age        2250     75  45.17689   17      91  Age
-------------------------------------------------------------------------------

. tab partyid , m

   Party ID |      Freq.     Percent        Cum.
------------+-----------------------------------
   1_StrDem |        408       18.13       18.13
    2_WkDem |        394       17.51       35.64
    3_I/Dem |        331       14.71       50.36
    4_Indep |        257       11.42       61.78
    5_I/Rep |        283       12.58       74.36
    6_WkRep |        325       14.44       88.80
   7_StrRep |        252       11.20      100.00
------------+-----------------------------------
      Total |      2,250      100.00
```

```
. ologit partyid age

Ordered logistic regression                       Number of obs   =       2250
                                                  LR chi2(1)      =      10.97
                                                  Prob > chi2     =     0.0009
Log likelihood = -4336.8443                       Pseudo R2       =     0.0013

------------------------------------------------------------------------------
     partyid |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
         age |  -.0072398   .0021887    -3.31   0.001    -.0115295   -.0029501
-------------+----------------------------------------------------------------
       /cut1 |   -1.83178   .1127654                     -2.052796   -1.610764
       /cut2 |  -.9081746   .1057363                     -1.115414   -.7009352
       /cut3 |     -.3002   .1039622                     -.5039622   -.0964378
       /cut4 |   .1665945   .1040661                     -.0373713    .3705603
       /cut5 |   .7505707   .1061486                      .5425233    .9586182
       /cut6 |   1.754251   .1161458                       1.52661    1.981893
------------------------------------------------------------------------------

. listcoef

ologit (N=2250): Factor Change in Odds

  Odds of: >m vs <=m

----------------------------------------------------------------------
     partyid |      b         z     P>|z|    e^b    e^bStdX      SDofX
-------------+--------------------------------------------------------
         age | -0.00724   -3.308   0.001   0.9928   0.8822     17.3155
----------------------------------------------------------------------
```


```
. brant

Brant Test of Parallel Regression Assumption

    Variable |      chi2   p>chi2    df
-------------+--------------------------
         All |     89.12    0.000     5
-------------+--------------------------
         age |     89.12    0.000     5
----------------------------------------
```



End BRM