Applied Microeconometrics (L8): Ordered and Multinomial Outcomes.

Nicholas Giannakopoulos

University of Patras
Department of Economics

ngias@upatras.gr

December 3, 2019

Overview

Ordered and Multinomial Response Outcomes

Ordered probit

Multinomial Logit

Modeling

- Ordered response models (dependent variable is NOT strictly continuous and NOT binary). Example:
  - rating of the corporate payment default risk, A (best) to D (worst)
  - rating of the sovereign default risk, A (best) to D (worst)
  - rating of self-assessed health, 0 (poor) to 3 (excellent health)
- Multinomial response models. Examples:
  - individual's employment status: unemployed (y=0); wage-employed (y=1); self-employed (y=2)
  - health care utilization: no use of health care (y=0); a GP visit but no use of hospital visits (y=1); a hospital visit, with or without a GP visit (y=2)
  - traveling from Patras to Athens: car (y=0); bus (y=1); train (y=2)

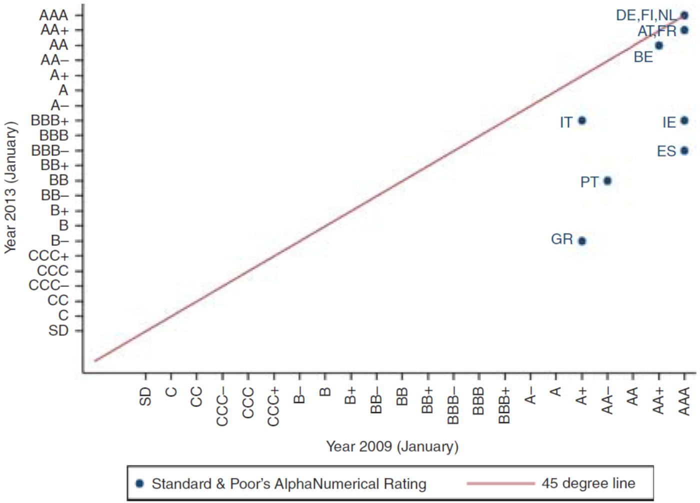
Example: Sovereign Rating

120 *G. Ferri et al. / Journal of Banking & Finance 25 (2001) 115–148*

Table 1
Moody's and S&P alphanumeric ratings' conversion into numeric values

Moody's	S&P	Numeric equivalent
Aaa	AAA	100
Aa1	AA+	95
Aa2	AA	90
Aa3	AA−	85
A1	A+	80
A2	A	75
A3	A−	70
Baa1	BBB+	65
Baa2	BBB	60
Baa3	BBB−	55
Ba1	BB+	50
Ba2	BB	45
Ba3	BB−	40
B1	B+	35
B2	B	30
B3	B−	25
Caa1	From CCC+ to CCC−	20
Caa2	CC	15
Caa3	C	10
Caa	D	5

Example: Sovereign Ratings (Demoussis, Drakos and Giannakopoulos, 2017)



Example: Self-assessed health in Greece

Heatlh status	Indicator	2009	2014
Very good	5	43,8	43,5
Good	4	31,5	25,4
Fair	3	16,5	15,2
Bad	2	5,6	4,2
Very bad	1	2,5	1,7
UNK	n.a	0,1	0,0
Total	100	100	100

Source: National Health Survey (EL.STAT)

Modeling Ordered Response Variables

- Two models
  - Ordered probit: Normal CDF $\Phi(\cdot)$
  - Ordered logit: Normal CDF $\Lambda(\cdot)$
- Ordered response y, $y = [0, 1, 2, ..., J]$
- Latent variable model:

$$\begin{aligned} y^\star =& \beta_1 x_1 + ... + \beta_k x_k + \epsilon \\ =& x'\beta + \epsilon \end{aligned} \tag{1}$$

Ordered probit

- We do not observe the latent variable, $y^\star$
- BUT we observe outcomes (choices)

$$
\begin{aligned}
y &= 0 \;\; \textit{if} \;\; y^\star \leq \alpha_1 \\
y &= 1 \;\; \textit{if} \;\; \alpha_1 < y^\star \leq \alpha_2 \\
y &= 2 \;\; \textit{if} \;\; \alpha_2 < y^\star \leq \alpha_3 \\
& \cdots \\
y &= J \;\; \textit{if} \;\; \alpha_J < y^\star
\end{aligned} \tag{2}
$$

Ordered probit

- Suppose y can take three values: 0, 1 or 2

$$\begin{aligned} y = 0 \;\; \textit{if} \;\; & \quad\quad\quad x'\beta + \epsilon \leq \alpha_1 \\ y = 1 \;\; \textit{if} \;\; & \alpha_1 < x'\beta + \epsilon \leq \alpha_2 \\ y = 2 \;\; \textit{if} \;\; & \alpha_2 < x'\beta + \epsilon \end{aligned} \tag{3}$$

Ordered probit

- Probabilities of observing $y = 0, 1, 2$

$$\begin{aligned} Pr(y=0) &= Pr(x'\beta + \epsilon \leq \alpha_1) \\ &= Pr(\epsilon \leq \alpha_1 - x'\beta) \\ &= \Phi(\alpha_1 - x'\beta) \\ &= 1 - \Phi(x'\beta - \alpha_1) \end{aligned} \tag{4}$$

$$\begin{aligned} Pr(y=2) &= Pr(x'\beta + \epsilon > \alpha_2) \\ &= Pr(\epsilon > \alpha_2 - x'\beta) \\ &= 1 - \Phi(\alpha_2 - x'\beta \\ &= \Phi(x'\beta - \alpha_2)) \end{aligned} \tag{5}$$

$$Pr(y=1) = \Phi(x'\beta - \alpha_1) - \Phi(x'\beta - \alpha_2) \tag{6}$$

Hint: $\Phi(\alpha) = 1 - \Phi(-\alpha)$

Ordered probit

- Interpretation: Partial effects on Probabilities of observing $y = 0, 1, 2$

$$\frac{\partial Pr(y=0|x)}{\partial x_k} = -\phi(x'\beta - \alpha_1)\beta_k \tag{7}$$

$$\frac{\partial Pr(y=1|x)}{\partial x_k} = -[\phi(x'\beta - \alpha_1) - \phi(x'\beta - \alpha_2)]\beta_k \tag{8}$$

$$\frac{\partial Pr(y=2|x)}{\partial x_k} = \phi(x'\beta - \alpha_2)\beta_k \tag{9}$$

Multinomial response: Multinomial logit

- Unordered Outcomes
- Employment status: self-employed (SE), wage-earner (WE) or unemployed (UE)
- logit model (0/1) is extended to more than two outcomes
- pep var: $J = 3$, $SE = 0, WE = 1, UE = 2$

$$\begin{aligned} y &= 0 \;\; \textit{if} \;\; \text{individual is UE} \\ y &= 1 \;\; \textit{if} \;\; \text{individual is WE} \\ y &= 2 \;\; \textit{if} \;\; \text{individual is SE} \end{aligned} \tag{10}$$

- conditional probability ($j = 0, 1, 2$)

$$Pr(y_1 = j|x_i) \tag{11}$$

Multinomial logit

- Probabilities in logit form

$$Pr(y=1|x_i) = \frac{exp(x_i\beta_1)}{1+exp(x_i\beta_1)+exp(x_i\beta_2)}$$

$$Pr(y=2|x_i) = \frac{exp(x_i\beta_2)}{1+exp(x_i\beta_1)+exp(x_i\beta_2)} \tag{12}$$

$$Pr(y=0|x_i) = 1 - Pr(y_i=1|x_i) - Pr(y_i=2|x_i)$$

$$= \frac{1}{1+exp(x_i\beta_1)+exp(x_i\beta_2)}$$

Multinomial logit

- Marginal effects

$$\frac{\partial Pr(y=1|x_i)}{\partial x_{ik}} = Pr(y_i=1|x_i)\Big[\beta_{1k} - \frac{\beta_{1k}exp(x_i\beta_1)+\beta_{2k}exp(x_i\beta_2)}{1+exp(x_i\beta_1)+exp(x_i\beta_2)}\Big]$$

$$\frac{\partial Pr(y=2|x_i)}{\partial x_{ik}} = Pr(y_i=2|x_i)\Big[\beta_{2k} - \frac{\beta_{1k}exp(x_i\beta_1)+\beta_{2k}exp(x_i\beta_2)}{1+exp(x_i\beta_1)+exp(x_i\beta_2)}\Big]$$

$$\frac{\partial Pr(y=0|x_i)}{\partial x_{ik}} = Pr(y_i=0|x_i)\Big[- \frac{\beta_{1k}exp(x_i\beta_1)+\beta_{2k}exp(x_i\beta_2)}{1+exp(x_i\beta_1)+exp(x_i\beta_2)}\Big] \tag{13}$$

Multinomial logit(Odds Ratio)

- Interpreting Odds Ratios
  - Odds ratios in logistic regression can be interpreted as the effect of a one unit of change in X in the predicted odds ratio with the other variables in the model held constant.

$$\frac{\text{odds (if the corresponding variable is incremented by 1)}}{\text{odds (if variable not increamented)}}$$

$$\frac{P(\text{event}|x+1)/(1-P(\text{event}|x+1)}{P(\text{event}|x)/P(\text{event}|x)}$$

Multinomial logit(Odds Ratio)

- Important property of odds ratios: they are constant
  - say you estimate the logistic regression model
  - $-13.70837 + .1685x_1 + .0039x_2$
  - The effect of the odds of a 1-unit increase in x_1 is $exp(.1685) = 1.18$
  - The odds increase by 18%
  - Incrementing x_1 increases the odds by 18% regardless of the value of x_2 (0, 1000, etc.)