

# Contents

	<i>List of tables</i>	<i>page</i> ix
	<i>List of figures</i>	xi
	<i>Preface</i>	xiii
1	Introduction: random utility and ordered choice models	1
2	Modeling binary choices	9
	2.1 Random utility formulation of a model for binary choice	10
	2.2 Probability models for binary choices	11
	2.3 Estimation and inference	16
	2.4 Covariance matrix estimation	26
	2.5 Application of the binary choice model to health satisfaction	28
	2.6 Partial effects in a binary choice model	30
	2.7 Hypothesis testing	39
	2.8 Goodness of fit measures	44
	2.9 Heteroscedasticity	54
	2.10 Panel data	57
	2.11 Parameter heterogeneity	75
	2.12 Endogeneity of a right-hand-side variable	80
	2.13 Bivariate binary choice models	83
	2.14 The multivariate probit and panel probit models	93
	2.15 Endogenous sampling and case control studies	96
3	A model for ordered choices	99
	3.1 A latent regression model for a continuous measure	99
	3.2 Ordered choice as an outcome of utility maximization	103
	3.3 An observed discrete outcome	105
	3.4 Probabilities and the log likelihood	108

	3.5	Log likelihood function	108
	3.6	Analysis of data on ordered choices	109
4		Antecedents and contemporary counterparts	111
	4.1	The origin of probit analysis: Bliss (1934a), Finney (1947a)	111
	4.2	Social science data and regression analysis for binary outcomes	116
	4.3	Analysis of binary choice	117
	4.4	Ordered outcomes: Aitchison and Silvey (1957), Snell (1964)	117
	4.5	Minimum chi squared estimation of an ordered response model: Gurland <i>et al.</i> (1960)	123
	4.6	Individual data and polychotomous outcomes: Walker and Duncan (1967)	125
	4.7	McKelvey and Zavoina (1975)	127
	4.8	Cumulative odds model	127
	4.9	Continuation ratio model	130
	4.10	The ordered regression model	130
	4.11	Other related models	132
	4.12	The latent continuous variable	134
5		Estimation, inference and analysis using the ordered choice model	136
	5.1	Application of the ordered choice model to self-assessed health status	136
	5.2	Distributional assumptions	138
	5.3	The estimated ordered probit (logit) model	138
	5.4	The estimated threshold parameters	140
	5.5	Interpretation of the model – partial effects and scaled coefficients	142
	5.6	Inference	150
	5.7	Prediction – computing probabilities	157
	5.8	Measuring fit	160
	5.9	Estimation issues	167
6		Specification issues and generalized models	181
	6.1	Functional form issues and the generalized ordered choice model (1)	181
	6.2	Model implications for partial effects	193
	6.3	Methodological issues	198
	6.4	Specification tests for ordered choice models	198

7	Accommodating individual heterogeneity	208
	7.1 Threshold models – the generalized ordered probit model (2)	209
	7.2 Nonlinear specifications – a hierarchical ordered probit (HOPIT) model	214
	7.3 Thresholds and heterogeneity – anchoring vignettes	219
	7.4 Heterogeneous scaling (heteroscedasticity) of random utility	232
	7.5 Individually heterogeneous marginal utilities	237
	Appendix: Equivalence of the vignette and HOPIT models	237
8	Parameter variation and a generalized model	239
	8.1 Random-parameters models	239
	8.2 Latent class and finite mixture modeling	247
	8.3 Generalized ordered choice model with random thresholds (3)	262
9	Ordered choice modeling with panel and time series data	268
	9.1 Ordered choice models with fixed effects	268
	9.2 Ordered choice models with random effects	275
	9.3 Testing for random or fixed effects: a variable addition test	278
	9.4 Extending parameter heterogeneity models to ordered choices	281
	9.5 Dynamic models	285
	9.6 Spatial autocorrelation	289
10	Bivariate and multivariate ordered choice models	290
	10.1 Multiple equations	290
	10.2 Bivariate ordered probit models	291
	10.3 Polychoric correlation	294
	10.4 Semi-ordered bivariate probit model	295
	10.5 Applications of the bivariate ordered probit model	295
	10.6 A panel data version of the bivariate ordered probit model	297
	10.7 Trivariate and multivariate ordered probit models	299
11	Two-part and sample selection models	302
	11.1 Inflation models	302
	11.2 Sample selection models	306
	11.3 An ordered probit model with endogenous treatment effects	319
12	Semiparametric and nonparametric estimators and analyses	320
	12.1 Heteroscedasticity	321
	12.2 A distribution free estimator with unknown heteroscedasticity	323



12.3	A semi-nonparametric approach	324
12.4	A partially linear model	327
12.5	Semiparametric analysis	327
12.6	A nonparametric duration model	329
	<i>References</i>	337
	<i>Index</i>	361

# Tables

2.1	Data used in binary choice application	<i>page</i> 29
2.2	Estimated probit and logit models	29
2.3	Alternative estimated standard errors for the probit model	30
2.4	Partial effects for probit and logit models at means of $x$	32
2.5	Marginal effects and average partial effects	37
2.6	Hypothesis tests	43
2.7	Homogeneity test	43
2.8	Fit measures for probit model	49
2.9a	Prediction success for probit model based on $\hat{y}_i = 1[\hat{F}_i > .5]$	50
2.9b	Predictions for probit model based on probabilities	50
2.10	Success measures for predictions by estimated probit model using $\hat{y}_i = 1[\hat{F}_i > .5]$	50
2.11	Heteroscedastic probit model	56
2.12	Cluster corrected covariance matrix (7,293 groups)	59
2.13	Fixed effects probit model	61
2.14	Estimated fixed effects logit models	65
2.15	Estimated random effects probit models	70
2.16a	Semiparametric random effects probit model	70
2.16b	Estimated parameters for four class latent class model	70
2.17	Random effects model with Mundlak correction	74
2.18	Estimated random parameter models	78
2.19	Estimated partial effects	78
2.20	Cross-tabulation of healthy and working	86
2.21	Estimated bivariate probit model	87
2.22	Estimated sample selection model	93
2.23	Estimated panel probit model	95
4.1	McCullagh application of an ordered outcome model	130
5.1	Estimated ordered choice models: probit and logit	139
5.2	Estimated partial effects for ordered choice models	144
5.3	Estimated expanded ordered probit model	147
5.4	Transformed latent regression coefficients	150

---

5.5	Estimated partial effects with asymptotic standard errors	158
5.6	Mean predicted probabilities by kids	159
5.7a	Predicted vs. actual outcomes for ordered probit model	165
5.7b	Predicted probabilities vs. actual outcomes for ordered probit model	165
5.8	Predicted vs. actual outcomes for automobile data	166
5.9	Grouped data for ordered choice modeling response frequency in a taste-testing experiment	168
5.10	Estimated ordered choice model based on grouped data	169
5.11	<i>Stata</i> and <i>NLOGIT</i> estimates of an ordered probit model	172
5.12	Software used for ordered choice modeling	179
6.1	Brant test for parameter homogeneity	186
6.2	Estimated ordered logit and generalized ordered logit (1)	191
6.3	Boes and Winkelmann estimated partial effects	194
7.1	Estimated generalized ordered probit models from Terza (1985)	211
7.2	Estimated hierarchical ordered probit models	217
7.3	Estimated partial effects for ordered probit models	218
7.4	Predicted outcomes from ordered probit models	219
7.5	Estimated heteroscedastic ordered probit model	235
7.6	Partial effects in heteroscedastic ordered probit model	236
8.1	Estimated random parameters ordered probit model	242
8.2	Implied estimates of parameter matrices	243
8.3	Estimated partial effects from random-parameters model	244
8.4	Estimated two-class latent class ordered probit models	256
8.5	Estimated partial effects from latent class models	257
8.6	Estimated generalized random-thresholds ordered logit model	266
9.1	Monte Carlo analysis of the bias of the MLE in fixed-effects discrete choice models (Means of empirical sampling distributions, $n = 1,000$ individuals, $R = 200$ replications)	270
9.2	Fixed-effects ordered choice models	279
9.3	Random effects ordered logit models – quadrature and simulation	280
9.4	Random effects ordered probit model with Mundlak correction	281
9.5	Random parameters ordered logit model	283
9.6	Latent-class ordered logit models	284
10.1	Applications of bivariate ordered probit since 2000	296
11.1	Estimated ordered probit sample selection model	310
12.1	Grouping of strike durations	334
12.2	Estimated logistic duration models for strike duration	334

# Figures

1.1	IMDb.com ratings ( <a href="http://www.imdb.com/title/tt0465234/ratings">www.imdb.com/title/tt0465234/ratings</a> )	<i>page</i> 3
2.1	Random utility basis for a binary outcome	11
2.2	Probability model for binary choice	14
2.3	Probit model for binary choice	15
2.4	Partial effects in a binary choice model	31
2.5	Fitted probabilities for a probit model	39
2.6	Prediction success for different prediction rules	51
2.7	ROC curve for estimated probit model	54
2.8	Distribution of conditional means of income parameter	80
3.1	Underlying probabilities for an ordered choice model	108
4.1	Insecticide experiment	112
4.2	Table of probits for values of $p_i$	114
4.3	Percentage errors in Pearson table of probability integrals	114
4.4	Implied spline regression in Bliss's probit model	115
5.1	Self-reported health satisfaction	137
5.2	Health satisfaction with combined categories	137
5.3	Estimated ordered probit model	140
5.4a	Sample proportions	141
5.4b	Implied partitioning of latent normal distribution	141
5.5	Partial effect in ordered probit model	145
5.6	Predicted probabilities for different ages	160
6.1	Estimated partial effects in Boes and Winkelmann (2006b) models	196
6.2	Estimated partial effects for linear and nonlinear index functions	196
7.1	Differential item functioning in ordered choices	220
7.2	KMST comparison of political efficacy	229
7.3	KMST estimated vignette model	230
8.1	Kernel density for estimate of the distribution of means of income coefficient	247



## List of figures

---

12.1	Table 1 from Stewart (2005)	325
12.2	Job satisfaction application, extended	326
12.3	Strike duration data	333
12.4	Estimated nonparametric hazard functions	335
12.5	Estimated hazard function from log-logistic parametric model	335