Journal of Social and Economic Statistics

Vol. 2, No.1, Summer 2013

A COMPARISON OF BINARY LOGIT AND PROBIT MODELS WITH A SIMULATION STUDY[*]

Selen CAKMAKYAPAN[a] , Atilla GOKTAS[b]

Abstract

Logit and probit models which widely used are members of the family of generalized linear models. Particularly, when the dependent variable is binary, both models may be used for the estimation of the functional relationship between dependent and independent variables. Since those models are utilized for the same purposes, the question of which model performs better comes to the mind. For this intention, a Monte Carlo simulation was carried out to compare both the binary probit and logit models under different conditions. In data generation stage, by employing latent variable approach, different sample sizes, different cut points, and different correlations between dependent variable and independent variables were taken into account. To make a comparison between logit and probit models, residuals, deviations and different Pseudo-R squares which are used for qualitative data analysis, were calculated and the results were interpreted.

Keywords: Binary Logit Model, Binary Probit Model, Latent Variable, Monte Carlo Simulation, Pseudo R-Square

JEL Classification: C15, C53, C63

Authors' Affiliation

[a] – Hacettepe University, Department of Statistic, selencakmakyapan@hacettepe.edu.tr

[b] – Mugla Sıtkı Kocman University, Department of Statistic, gatilla@mu.edu.tr

[*] ***An earlier version of this paper was presented at *The 6th International Conference on Applied Statistics*, November 2012, Bucharest.

1. Introduction

Generalized linear models extend classical linear models, and all probability models to be discussed can be subsumed under generalized linear models. Therefore, specific treatments of the models belonging to this family can also be generalized and regarded as common to all models of the category (Liao, 1994).

Logit and probit models are the most commonly used members of the family of generalized linear models. As the simplest logit and probit model, response variable in binary logit and probit models have only two categories. The occurrence and nonoccurrence of these events are the categories in the dependent variables.

Binary logit and probit models assume an underlying dependent variable defined as Y^* which can be presented as a functional relationship in Eq. 1.

$$Y^* = \sum_{k=1}^{K} \beta_k x_k + \varepsilon \qquad (1)$$

In practice, Y^* is unobserved or called a latent variable ranging from $-\infty$ to ∞ that generates the observed Y is binary dependent variable.

Both of these models may be used to analyze same data sets for the same purpose. For this reason, the question of which model performs better may come to the mind. In this study, a comparison of binary probit and logit models via a simulation study was performed under different sample sizes, different correlations between dependent and independent variables and different cut points for converting the latent variable to be binary.

In this paper we present the very general information about the latent variable approach, binary logit model, binary probit model, goodness of fit measure for binary logit and probit model and the stages and results of simulation study.

2. A Latent Variable Model For Binary Variables

Suppose that there is an unobserved or latent variable Y^* ranging from $-\infty$ to ∞ that generates the observed Y. In the generation processes of Y, the larger values of Y^* are classified as $Y = 1$, while those with smaller values of Y^* are observed as $Y = 0$. The latent variable is assumed to be linearly dependent to the observed X 's throughout the structural model. Y^* is linked to the observed binary variable Y with the measurement equation as below:

$$Y = \begin{cases} 1, & Y_i^* > \tau \\ 0, & Y_i^* \leq \tau \end{cases} \quad (2)$$

where τ is the threshold or cut point. If $Y_i^* > \tau$ then $Y = 0$. If Y^* is less than the threshold, τ then $Y = 1$.

The link between the latent Y^* and the observed Y is illustrated in Figure 1 for the model $Y^* = \sum_{k=1}^{3} \beta_k x_k + \varepsilon$. In this figure, the vertical axis represents Y^*, with the threshold τ indicated by a horizontal dashed line. The distribution of Y^* is shown by the bell-shaped curves which could be treated perhaps a third dimension of the figure. When Y^* is larger than τ, indicated by the shaded region, $Y = 1$ is observed (Long, 1997).

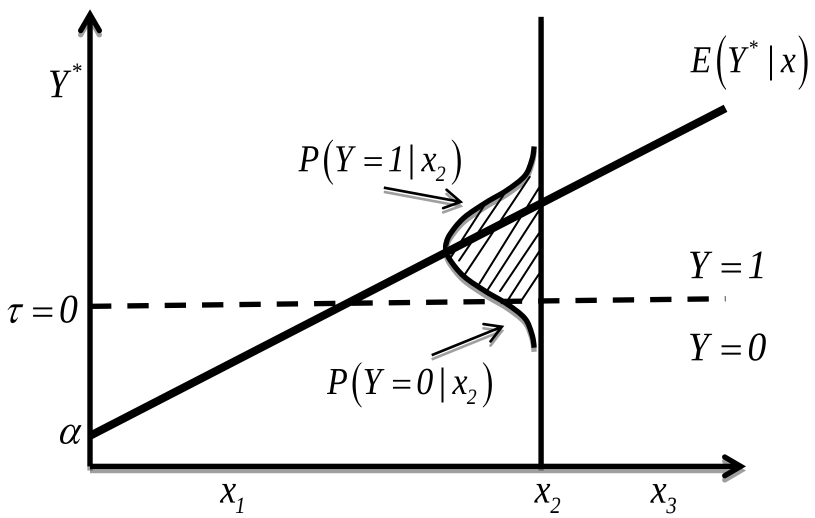

Figure 1. The Distribution of Y^* given x and probability of observed values in the binary response model

3. Binary Logit and Probit Model

When the latent variable is unobserved or the dependent variable is binary, the model cannot be estimated using ordinary least squares. Instead, maximum likelihood estimation is used, which requires assumptions about the distribution of the errors. Most often, the choice is between normal errors which result in the probit model, and logistic errors which result in the logit model (Long, 1997).

A logit model that takes a binary outcome variable is specified as follows:

$$log\left[\frac{P(Y=1)}{1-P(Y=1)}\right]=\sum_{k=1}^{K}\beta_k x_k \qquad (3)$$

A specification of the logit model of event probability, L, representing the logistic distribution:

The probit model represents another type of widely used statistical model for fitting data with binomial distributions. Φ represents the standard normal cumulative distribution and probit models are specified as follows:

$$P(Y=1)=L\left(\sum_{k=1}^{K}\beta_k x_k\right)=\frac{e^{\sum_{k=1}^{K}\beta_k x_k}}{1+\sum_{k=1}^{K}\beta_k x_k} \qquad (4)$$

$$P(Y=1)=\Phi\left(\sum_{k=1}^{K}\beta_k x_k\right)=\int_{-\infty}^{\sum_{k=1}^{K}\beta_k x_k} exp(-u^2/2)/\sqrt{2\pi}du \qquad (5)$$

Since Y^* is unobserved, the variance of the errors cannot be estimated. In the probit model, it is assumed that $Var(\varepsilon|x)=1$ and in the logit model that $Var(\varepsilon|x)=\pi^2/3\approx 3.29$. For detailed information, see Aldrich and Nelson (1984), Liao (1994), Maddala (1983), Long (1997), Greene (1990).

4. Goodness of Fit Measures

Analogous to the residual sum of squares in linear regression, the goodness-of-fit of a generalized linear model can be measured by the scaled deviance

$$D(Y;\hat{\mu}) = 2[\,l(Y;Y) - l(\hat{\mu},Y)] \qquad (6)$$

where $l(Y;Y)$ is the maximum likelihood achievable for an exact fit in which the fitted values are equal to the observed values, and $l(\hat{\mu},Y)$ is the log-likelihood function calculated at the estimated parameters β. The deviance function is very useful for comparing two models when one model has parameters that are a subset of the second model. The deviance is additive for such nested models if maximum likelihood estimates are used (McCullagh-Nelder, 1989).

Consider two nested models with the second having some covariates omitted and denote the maximum likelihood estimates in the two models by $\hat{\mu}_1$ and $\hat{\mu}_2$, respectively. Then the deviance difference $D(Y;\hat{\mu}_1) - D(Y;\hat{\mu}_2)$ is identical to the likelihood-ratio statistic and has an approximate χ^2 distribution with degrees of freedom equal to the difference between the numbers of parameters in the two models. For probability distributions in the exponential family the χ^2 approximation is usually quite accurate for differences of deviance even though it may be inaccurate for the deviances themselves (McCullagh-Nelder 1989).

The Pearson residuals are elements of the Pearson chi-square, that can be used to detect ill-fitted factor/covariate patterns. For a binomial distribution with mi trials in the *i*th observation, it is defined as Eq. 7. Large values of r_i suggest a failure of the model to fit a given observation.

$$r_i = \frac{Y_i - m_i\hat{\mu}_i}{\sqrt{m_i\hat{\mu}_i(1-\hat{\mu}_i)}} \qquad (7)$$

The Akaike information criterion is a way of selecting a model from a set of models. It can be said to describe the tradeoff between bias and variance in model construction, or loosely speaking between accuracy and complexity of the model and it is defined as follows:

$$AIC = 2k - 2\ln(L) \tag{8}$$

where k is the number of parameters and L is the maximized value of the likelihood function for the estimated model.

A large of different pseudo- R^2 measure for binary dependent variable models are surveyed. Pseudo- R^2 's are used to describe how well a model fits a set of data. Measures include those based solely on the maximized likelihoods with and without the restriction that slope coefficients are zero, those which require further calculation based on parameter estimates of the coefficients and variances and those that are based solely on whether the qualitative predictions of the model are correct or not. The theme of the survey is that while there is no obvious criterion for choosing which Pseudo- R^2 to use (Veall-Zimmermann, 1996). Also, different pseudo- R^2 have quite different values for the same model and most of them are not seen as a number between 0 and 1.

Some of Pseudo- R^2 are presented in Table.1 where LRT ($LRT = 2(l_m - l_0)$) is likelihood ratio statistics and l_m is the log-likelihood value of the model, l_0 is the log-likelihood value if the non-intercept coefficients are restricted to zero. LRT^* ($LRT^* = 2(l_{max} - l_0)$) and l_{max} represents maximum likelihood ratio and maximum log-likelihood value, respectively. H_i is the value of the cumulative distribution function for observation i . For detailed information, see Veall-Zimmermann (1996), Cameron-Windmeijer (1997), Cox-Wermuth (1992), Hagle- Mitchell (1992), Tardiff (1976).

Table 1. Pseudo- R^2 's (Veall-Zimmermann, (1996))

$R^2_{AN} = LRT / (LRT + N)$	Aldrich and Nelson (1984)
$R^2_{VZ} \dfrac{LRT}{(LRT + N)} / \dfrac{LRT^*}{(LRT^* + N)} = \dfrac{LRT}{(LRT + N)} / \dfrac{-2l_0}{N - 2l_0}$	Veall and Zimmerman (1990, 1992)
$R^2_{MF} = \dfrac{LRT}{LRT^*} / \dfrac{(l_m - l_0)}{(l_{max} - l_0)} = 1 - \dfrac{l_m}{l_0}$	McFadden (1973)
$R^2_M = 1 - exp(-LRT / N)$	Maddala (1983)
$R^2_{CU} = \dfrac{1 - exp(-LRT / N)}{1 - exp(-LRT^* / N)}$	Cragg and Uhler (1970)
$R^2_{MZ} = \dfrac{\sum_{i=1}^{N} (\hat{Y}_i^* - \bar{Y}^*)^2}{\sum_{i=1}^{N} (\hat{Y}_i^* - \bar{Y}^*)^2 + N\hat{\sigma}^2}$	McKelvey and Zavoina (1975)
$R^2_C = \dfrac{\left[cov(Y , H) \right]^2}{var(Y).var(H)} = \dfrac{var(H)}{var(Y)}$	Neter and Maynes (1970), Morrison (1972), Goldberger (1973) and Efron (1978)
$R^2_L = 1 - \dfrac{\sum_{i=1}^{N} (Y_i - H_i)^2}{\sum_{i=1}^{N} (Y_i - \bar{Y})^2}$	Lave (1970)

5. Simulation Study

The main purpose of this study is to determine whether there exists a priority or a difference between binary logit and probit models in fitting under certain conditions that are different sample sizes, different correlations between variables and different cut points for latent dependent variable.

Latent variable used in this study is treated to be continuous and affected by three independent variables coming from multivariate standard normal distribution so their means are zero and the variances are one of course.

To generate data from multivariate standard normal distribution, three different variance-covariance matrices were considered. These matrices were determined arbitrarily that they were positive definitive and correlations between independent variable and dependent variables were zero. Thus, multicolinearity has been avoided. Special covariance values were selected to create different correlation between dependent and independent variables. Covariances between variables are identical to their correlations because the variables have been generated from multivariate standard normal distributions. The variance-covariance matrices were named, “high”, “low” and “no”, respectively.

When linear regression model has been fitted to the data set of each generated from multivariate standard normal distribution with the following covariance matrices , $R^2_{OLS} \approx 0.9$, $R^2_{OLS} \approx 0.30$ and $R^2_{OLS} \approx 0.04$ are obtained approximately.

$$\Sigma_{high} = \begin{bmatrix} 1 & 0.4 & 0.5 & -0.7 \\ 0.4 & 1 & 0 & 0 \\ 0.5 & 0 & 1 & 0 \\ -0.7 & 0 & 0 & 1 \end{bmatrix}$$

$$\Sigma_{low} = \begin{bmatrix} 1 & 0.4 & 0.2 & -0.3 \\ 0.4 & 1 & 0 & 0 \\ 0.2 & 0 & 1 & 0 \\ -0.3 & 0 & 0 & 1 \end{bmatrix}$$

$$\Sigma_{no} = \begin{bmatrix} 1 & 0.01 & 0.1 & -0.1 \\ 0.01 & 1 & 0 & 0 \\ 0.1 & 0 & 1 & 0 \\ -0.1 & 0 & 0 & 1 \end{bmatrix}$$

In order to examine the effect of sample size in model selection, 5 different sample sizes were considered: 1000, 500, 200, 100 and 40. For each of the matrices and sample sizes, data generation was repeated 1000 times which was found to be sufficient.

After data generation, the latent dependent variable transformed to a binary case for two different cut points: 0 and 0.53. A cut point is z score in standard normal distribution table corresponds to event probability.

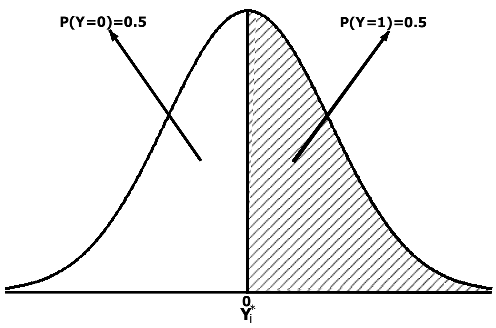

Figure 2. The cut points for $\boldsymbol{P(Y=1)=0.5}$

Dependent variable gets value : $Y_i = \begin{cases} 1, & Y_i^* > 0 \\ 0, & Y_i^* \leq 0 \end{cases}$ for $P(Y=1)=0.5$.

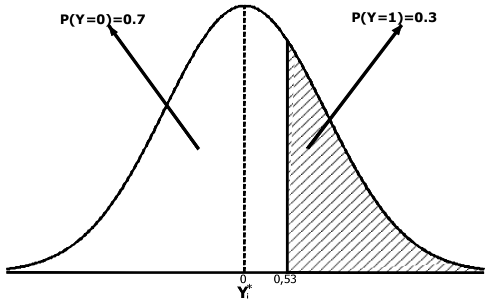

Figure 3. The cut points for $\boldsymbol{P(Y=1)=0.7}$

Dependent variable gets value : $Y_i = \begin{cases} 1, & Y_i^* > 0.53 \\ 0, & Y_i^* \leq 0.53 \end{cases}$ for $P(Y=1)=0.7$

In this study, 30 different data generation were performed and generated a total of 30000 data. In Table 2, data generation was summarized.

Table 2. Data generation and classification

Data				
Cut Point = 0			Cut Point = 0.53	
Sample Sizes	Repeat number	Variance-Covarince Matrices	Sample Sizes	Repeat number
40, 100, 200, 500, 1000	1000	High	40, 100, 200, 500, 1000	1000
		Low		
		No		

In the next step, parameter and probability estimations were obtained using both binary logit and probit models. And then, goodness of fit measures and pseudo-R^2's and their means on 1000 replication were calculated.

Student-t test was used to check whether there is a statisticaly significant difference between logit and probit model in terms of goodness of fit measures under different condition. Also, R^2_{OLS}'s were calculated from linear regression for the latent dependent variable and independent variables.

6. Simulation Results

Table 3 and 4 present only measure means and the representation of bold face of those measures are statistically significant differences between binary logit and binary probit model. Since the rest of the other measure (deviance, AIC, R^2_{AN} etc.) means are not significantly different in terms of logit and probit model, thus we have excluded them from the result tables. In the tables; *L* denotes logit model, *P* denotes probit model and N denotes sample size. For example, according to Pearson residuals in table 3, logit model is better than the probit one in “high” and “low” cases, for 500 and 1000 sample sizes. This is because measure mean values from the logit model are significantly smaller than the values from the probit. In “no” case, no matter what the condition is both models fit the data set identically so there is no priority.

When dependent and independent are uncorrelated, used models are expected to give inaccurate results so goodness of fit measure values for the model should be bad. In no case, this is true. For example, in table 4, according to R^2_{MZ}, logit model is better for any simple sizes. Since there is not much difference between table 3 and table 4 in interpretation thus cut points do not influence model selection.

Table 3. Statistically significant measure values for Cut Points = 0

Σ	N	R^2_{OLS}	L_Residuals	P_Residuals	L_ R^2_{MZ}	P_ R^2_{MZ}	L_ R^2_C	P_ R^2_C
High	40	0.90060	17.859	**16.291**	0.83929	0.84792	0.75049	0.75162
	100	0.90236	52.457	**48.910**	0.82240	0.82912	0.74232	0.74096
	200	0.90015	115.820	114.240	0.80991	**0.81586**	0.72830	0.72526
	500	0.90015	**304.670**	320.960	0.80261	**0.80784**	**0.72185**	0.71777
	1000	0.89957	**619.170**	676.760	0.80070	**0.80559**	**0.71945**	0.71496
Low	40	0.33250	37.820	**37.411**	0.28766	**0.31855**	0.25201	0.25391
	100	0.30984	98.487	98.364	0.22613	**0.25561**	0.21314	0.21301
	200	0.29908	198.490	198.840	0.21017	**0.23981**	0.20099	0.20040
	500	0.29534	**497.600**	499.480	0.19939	**0.22893**	0.19469	0.19390
	1000	0.29210	**995.080**	999.030	0.19449	**0.22412**	0.19064	0.18993
No	40	0.10026	39.647	39.548	**0.11358**	0.13907	0.09735	0.09805
	100	0.05203	99.953	99.934	**0.04820**	0.05942	0.04614	0.04620
	200	0.03703	199.980	199.970	**0.02732**	0.03452	0.02939	0.02940
	500	0.02595	499.990	499.990	**0.01561**	0.01983	0.01877	0.01877
	1000	0.02327	1000.000	1000.000	**0.01191**	0.01519	0.01601	0.01600

Table 4. Statistically significant measure values for Cut Points = 0.53

Σ	N	R^2_{OLS}	L_Residuals	P_Residuals	L_ R^2_{MZ}	P_ R^2_{MZ}	L_ R^2_C	P_ R^2_C
High	40	0.89818	15.702	**14.327**	0.79292	0.80330	0.73750	0.73976
	100	0.90133	46.683	**42.726**	0.76663	0.77289	0.73717	0.73474
	200	0.90052	102.250	**99.450**	0.73845	0.74438	0.72256	0.71899
	500	0.90017	**271.120**	283.320	0.72231	**0.72746**	**0.71210**	0.70751
	1000	0.90021	**547.280**	598.820	0.71983	**0.72489**	**0.71126**	0.70652
Low	40	0.34761	37.587	**36.173**	0.34260	**0.36754**	0.26645	0.26569
	100	0.30630	96.354	96.099	0.27748	**0.30570**	0.20472	0.20350
	200	0.29774	**195.440**	196.610	0.26171	**0.29004**	0.18944	0.18775
	500	0.29219	**492.040**	497.420	0.25091	**0.27930**	0.18037	0.17841
	1000	0.29200	**985.540**	998.240	0.24942	**0.27773**	0.17976	0.17770
No	40	0.10852	**39.019**	38.703	**0.24484**	0.27546	0.11509	0.11579
	100	0.05048	99.919	99.837	**0.21004**	0.24381	0.04330	0.04316
	200	0.03469	199.910	199.900	**0.19544**	0.23044	0.02653	0.02647
	500	0.02763	499.920	499.950	**0.18994**	0.22560	0.01958	0.01953
	1000	0.09297	**39.709**	39.239	**0.25284**	0.28356	0.09082	0.09057

6. Conclusion

In this study, different Pseudo-R^2 had quite different values for the same model and also there is no obvious criterion for choosing which Pseudo-R^2 to use. This situation reduced the credibility of these measures. So, deviance or pearson residuals were considered more appropriate for comparison binary logit and probit model. While according to model's deviance there is no difference between the models in all conditions, according to model's residulas the models fit differently in high and low cases and also for sample sizes. In high and low case, logit model is better for large sample sizes (500 and 1000) and probit model is better in small sizes (40,100,200). The sample size is effective to prefer which model is better. We can say that different correlations and cut points did not affect goodness of fit measures.

In order to avoid hardship caused by pseudo R square's instability, pearson residuals were considered for goodness of fit. According to residuals, sample sizes were effective in model selection. When differences were statistically significant for small sample sizes, probit model's residuals were lower so it was better model. Logit model was better model for large sample sizes.

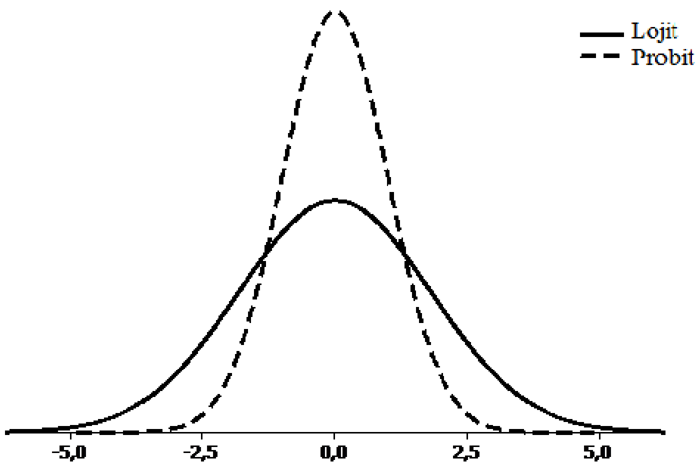

Figure 4. Distribution Curves for Logit and Probit model

Because of variance of probit model is one and variance of logit model is $\pi^2 / 3$ logit model has more flat distribution. Although the both models on the same axis as shown in Figure.4, logit model has heavier tails due to greater spread of the distribution curve. These situation causes that logit model is better than probit model in larger sample size. This is because when the sample size increases, probability of observes in tail increases too. This is the reason why logit model is better than probit model for large sample sizes (see also Amemiya (1981), Maddala (1983)). The result of study were obtained in this direction also.

References

Aldrich, J.H., Nelson, F.D. (1984). *Linear Probability, Logit, and Probit Models*. London: Sage Publications.

Amemiya, T. (1981). Qualitative Response Models: A Survey. *Journal of Economic Literature*, 19, pp. 1483-1536.

Cameron, A.C. Windmeijer, A.G. (1997). An R-squared Measure of Goodness of Fit for Some Common Nonlinear Regression Models. *Journal of Econometrics*, 77, pp. 329-342.

Hagle, T.M., Mitchell II, G.E. (1992). Goodness-of-Fit Measures for Probit and Logit. *Amercian Journal of Political Science*, 36, pp. 762-784.

Liao, T.M. (1994). *Interpreting Probability Models Logit, Probit, and Other Generalized Linear Models*. London: Sage Publications.

Long, J.S. (1997). *Regression Models for Categorical and Limited Dependent Variables*. California: Sage Publications.

Maddala, G.S. (1983). *Limited-Dependent and Qualitatives Variables in Econometrics*. Cambrige: Cambrige University Press.

McCullagh, P., Nelder, J.A. (1989). *Generalized Linear Models. 2nd ed. London*: Chapman and Hall.

Tardiff, T.J. (1976). A Note on Goodness-of-Fit Statistics For Probit and Logit Models. *Transportation*, 5, pp. 377-388.

Veall, M.R., Zimmermann, K.F. (1994). Evaluating Pseudo-R2's for Binary Probit Models. *Quality & Quality*, 28, 151-164.

Veall, M.R., Zimmermann, K.F. (1996). Pseudo-R2 Measures for Some Common Limited Dependent Variable Models. *Sunderforschungsbereich*, 386, pp.1-34

Windmeijer, F.A.G. (1995). Goodness of Fit Measures in Binary Choice Models. *Econometric Review*, 14, pp. 101-116.

Appendix

Appendix involve all results for all goodness of fit measures obtained from the simulation study

Table A1. Results for Cut points=0

Σ	*N*	*L_Deviance*	*P_Deviance*	*L_Residuals*	*P_Residuals*	*L_Akaike*	*P_Akaike*
High	40	15.611	15.456	17.859	**16.291**	0.54028	0.53639
	100	41.142	40.855	52.457	**48.910**	0.47142	0.46855
	200	87.223	86.792	115.820	114.240	0.46612	0.46396
	500	224.010	223.160	**304.670**	320.960	0.46003	0.45832
	1000	452.240	450.700	**619.170**	676.760	0.45824	0.45670
Low	40	43.135	43.055	37.820	**37.411**	1.22840	1.22640
	100	114.180	114.120	98.487	98.364	1.20180	1.20120
	200	232.270	232.220	198.490	198.840	1.19140	1.19110
	500	586.050	585.970	**497.600**	499.480	1.18410	1.18390
	1000	1177.800	1177.700	**995.080**	999.030	1.18380	1.18370
No	40	50.239	50.219	39.647	39.548	1.40600	1.40550
	100	132.950	132.950	99.953	99.934	1.38950	1.38950
	200	270.300	270.330	199.980	199.970	1.38160	1.38160
	500	682.620	682.620	499.990	499.990	1.37720	1.37720
	1000	1369.100	1369.100	1000.000	1000.000	1.37510	1.37510

Table A1. Continue

Σ	*N*	R^2_{OLS}	*L_* R^2_{AL}	*P_* R^2_{AL}	*L_adj.* R^2_{AL}	*P_adj.* R^2_{AL}	*L_* R^2_{VZ}	*P_* R^2_{VZ}
High	40	0.90060	0.49026	0.49129	0.84325	0.84502	0.85038	0.85217
	100	0.90236	0.48965	0.49041	0.84220	0.84351	0.84543	0.84674
	200	0.90015	0.48521	0.48578	0.83456	0.83554	0.83648	0.83748
	500	0.90015	0.48325	0.48371	0.83119	0.83198	0.83237	0.83317
	1000	0.89957	0.48255	0.48297	0.82999	0.83071	0.83089	0.83160
Low	40	0.33250	0.21124	0.21231	0.36333	0.36517	0.36630	0.36817
	100	0.30984	0.18649	0.18686	0.32076	0.32140	0.32192	0.32255
	200	0.29908	0.17865	0.17884	0.30728	0.30760	0.30797	0.30829
	500	0.29534	0.17438	0.17449	0.29993	0.30012	0.30035	0.30053
	1000	0.29210	0.17144	0.17156	0.29488	0.29508	0.29520	0.29540
No	40	0.10026	0.08925	0.08961	0.15351	0.15413	0.15527	0.15590
	100	0.05203	0.04445	0.04448	0.07645	0.07651	0.07674	0.07680
	200	0.03703	0.02874	0.02875	0.04943	0.04945	0.04954	0.04956
	500	0.02595	0.01852	0.01852	0.03186	0.03186	0.03190	0.03191
	1000	0.02327	0.01585	0.01585	0.02726	0.02726	0.02729	0.02729

Table A1. continue

Σ	*N*	$L\_R^2_L$	$P\_R^2_L$	$L\_R^2_{MF}$	$P\_R^2_{MF}$	$L\_R^2_M$	$P\_R^2_M$
High	40	0.74221	0.73958	0.71356	0.71642	0.61790	0.61941
	100	0.73561	0.73402	0.70107	0.70315	0.61694	0.61805
	200	0.72246	0.72164	0.68427	0.68583	0.61039	0.61124
	500	0.71632	0.71618	0.67633	0.67757	0.60750	0.60817
	1000	0.71420	0.71424	0.67354	0.67465	0.60646	0.60707
Low	40	0.24684	0.24546	0.20855	0.21003	0.23916	0.24053
	100	0.21130	0.21068	0.17075	0.17120	0.20662	0.20707
	200	0.19990	0.19957	0.15930	0.15951	0.19628	0.19650
	500	0.19384	0.19374	0.15334	0.15346	0.19074	0.19087
	1000	0.18973	0.18972	0.14977	0.14989	0.18709	0.18723
No	40	0.09598	0.09540	0.07632	0.07670	0.09527	0.09568
	100	0.04603	0.04593	0.03457	0.03460	0.04592	0.04596
	200	0.02935	0.02933	0.02170	0.02171	0.02934	0.02935
	500	0.01876	0.01875	0.01372	0.01372	0.01875	0.01876
	1000	0.01601	0.01601	0.01167	0.01167	0.01600	0.01600

Table A1. Continue

Σ	*N*	$L\_R^2_{MZ}$	$P\_R^2_{MZ}$	$L\_R^2_C$	$P\_R^2_C$	$L\_R^2_{CU}$	$P\_R^2_{CU}$
High	40	0.83929	0.84792	0.75049	0.75162	0.83088	0.83291
	100	0.82240	0.82912	0.74232	0.74096	0.82536	0.82685
	200	0.80991	**0.81586**	0.72830	0.72526	0.81522	0.81635
	500	0.80261	**0.80784**	**0.72185**	0.71777	0.81056	0.81147
	1000	0.80070	**0.80559**	**0.71945**	0.71496	0.80889	0.80970
Low	40	0.28766	**0.31855**	0.25201	0.25391	0.32149	0.32334
	100	0.22613	**0.25561**	0.21314	0.21301	0.27634	0.27695
	200	0.21017	**0.23981**	0.20099	0.20040	0.26213	0.26243
	500	0.19939	**0.22893**	0.19469	0.19390	0.25448	0.25465
	1000	0.19449	**0.22412**	0.19064	0.18993	0.24954	0.24973
No	40	**0.11358**	0.13907	0.09735	0.09805	0.12856	0.12913
	100	**0.04820**	0.05942	0.04614	0.04620	0.06143	0.06149
	200	**0.02732**	0.03452	0.02939	0.02940	0.03918	0.03919
	500	**0.01561**	0.01983	0.01877	0.01877	0.02502	0.02503
	1000	**0.01191**	0.01519	0.01601	0.01600	0.02135	0.02134

Table A2. Results for Cut points=0.53

Σ	***N***	***L_Deviance***	***P_Deviance***	***L_Residuals***	***P_Residuals***	***L_Akaike***	***P_Akaike***
High	40	13.867	13.706	15.702	**14.327**	0.49668	0.49265
	100	35.098	34.867	46.683	**42.726**	0.41098	0.40867
	200	75.057	74.662	102.250	**99.450**	0.40528	0.40331
	500	194.770	193.950	**271.120**	283.320	0.40154	0.39989
	1000	391.920	390.330	**547.280**	598.820	0.39792	0.39633
Low	40	36.784	36.725	37.587	**36.173**	1.06960	1.06810
	100	99.573	99.467	96.354	96.099	1.05570	1.05470
	200	203.350	203.220	**195.440**	196.610	1.04670	1.04610
	500	514.590	514.370	**492.040**	497.420	1.04120	1.04070
	1000	1031.500	1031.200	**985.540**	998.240	1.03750	1.03720
No	40	44.10100	44.062	**39.019**	38.703	1.25250	1.25150
	100	116.18000	116.170	99.919	99.837	1.22180	1.22170
	200	237.62000	237.610	199.910	199.900	1.21810	1.21810
	500	599.21000	599.200	499.920	499.950	1.21040	1.21040
	1000	43.88700	43.871	**39.709**	39.239	1.24720	1.24680

Table A2. Continue

Σ	***N***	R^2_{OLS}	L_ R^2_{AL}	P_ R^2_{AL}	L_adj. R^2_{AL}	P_adj. R^2_{AL}	L_ R^2_{VZ}	P_ R^2_{VZ}
High	40	0.89818	0.45880	0.46000	0.83502	0.83720	0.84005	0.84226
	100	0.90133	0.45888	0.45957	0.83516	0.83642	0.84014	0.84140
	200	0.90052	0.45566	0.45625	0.82930	0.83038	0.83080	0.83188
	500	0.90017	0.45202	0.45252	0.82268	0.82359	0.82394	0.82486
	1000	0.90021	0.45233	0.45281	0.82324	0.82411	0.82360	0.82447
Low	40	0.34761	0.21104	0.21188	0.38409	0.38562	0.38802	0.38957
	100	0.30630	0.17265	0.17334	0.31422	0.31548	0.31551	0.31677
	200	0.29774	0.16280	0.16324	0.29630	0.29710	0.29681	0.29761
	500	0.29219	0.15662	0.15692	0.28505	0.28559	0.28542	0.28597
	1000	0.29200	0.15642	0.15667	0.28468	0.28514	0.28489	0.28535
No	40	0.10852	0.10127	0.10196	0.18431	0.18557	0.18573	0.18704
	100	0.05048	0.04087	0.04096	0.07439	0.07454	0.07497	0.07512
	200	0.03469	0.02563	0.02566	0.04664	0.04671	0.04675	0.04682
	500	0.02763	0.01914	0.01915	0.03484	0.03486	0.03485	0.03487
	1000	0.09297	0.08058	0.08091	0.14666	0.14726	0.14967	0.15029

Table A2. continue

Σ	*N*	$L\_R^2_L$	$P\_R^2_L$	$L\_R^2_{MF}$	$P\_R^2_{MF}$	$L\_R^2_M$	$P\_R^2_M$
High	40	0.72689	0.72429	0.71192	0.71527	0.57215	0.57388
	100	0.73122	0.72888	0.70879	0.71070	0.57196	0.57296
	200	0.71638	0.71533	0.69121	0.69284	0.56715	0.56802
	500	0.70597	0.70574	0.67952	0.68087	0.56176	0.56249
	1000	0.70502	0.70512	0.67832	0.67963	0.56220	0.56290
Low	40	0.26329	0.25918	0.23695	0.23819	0.23866	0.23971
	100	0.20108	0.20003	0.17673	0.17762	0.18993	0.19076
	200	0.18666	0.18620	0.16216	0.16269	0.17758	0.17810
	500	0.17802	0.17790	0.15349	0.15384	0.16981	0.17017
	1000	0.17758	0.17759	0.15274	0.15303	0.16944	0.16973
No	40	0.11232	0.11065	0.09852	0.09939	0.10855	0.10936
	100	0.04300	0.04256	0.03646	0.03653	0.04220	0.04229
	200	0.02636	0.02625	0.02200	0.02203	0.02613	0.02617
	500	0.01953	0.01949	0.01612	0.01613	0.01939	0.01940
	1000	0.09008	0.08785	0.07932	0.07968	0.08577	0.08614

Table A2. Continue

Σ	*N*	$L\_R^2_{MZ}$	$P\_R^2_{MZ}$	$L\_R^2_C$	$P\_R^2_C$	$L\_R^2_{CU}$	$P\_R^2_{CU}$
High	40	0.79292	0.80330	0.73750	0.73976	0.81756	0.82006
	100	0.76663	0.77289	0.73717	0.73474	0.81725	0.81869
	200	0.73845	0.74438	0.72256	0.71899	0.80657	0.80780
	500	0.72231	**0.72746**	**0.71210**	0.70751	0.79865	0.79969
	1000	0.71983	**0.72489**	**0.71126**	0.70652	0.79825	0.79925
Low	40	0.34260	**0.36754**	0.26645	0.26569	0.34280	0.34432
	100	0.27748	**0.30570**	0.20472	0.20350	0.27082	0.27201
	200	0.26171	**0.29004**	0.18944	0.18775	0.25252	0.25327
	500	0.25091	**0.27930**	0.18037	0.17841	0.24135	0.24185
	1000	0.24942	**0.27773**	0.17976	0.17770	0.24067	0.24108
No	40	**0.24484**	0.27546	0.11509	0.11579	0.15555	0.15675
	100	**0.21004**	0.24381	0.04330	0.04316	0.06044	0.06056
	200	**0.19544**	0.23044	0.02653	0.02647	0.03720	0.03725
	500	**0.18994**	0.22560	0.01958	0.01953	0.02753	0.02754
	1000	**0.25284**	0.28356	0.09082	0.09057	0.12473	0.12527