The Interpretation of Coefficients in Models with Qualitative Dependent Variables*

Marc J. LeClere
College of Business and Management, University of Maryland, College Park, MD 20742

ABSTRACT

A substantial body of empirical accounting, finance, management, and marketing research utilizes single equation models with discrete dependent variables. Generally, the interpretation of the coefficients of the exogenous variables is limited to the sign and relative magnitude. This paper presents three methods of interpreting the coefficients in these models. The first method interprets the coefficients as marginal probabilities and the second method interprets the coefficients as elasticities of probability. The third method utilizes sensitivity analysis and examines the effect of hypothetical changes in exogenous variables on the probability of choice. This paper applies these methods to a published research study.

Subject Areas: Accounting and Statistical Techniques.

INTRODUCTION

During the last ten years, single equation models with qualitative or discrete dependent variables have become increasingly popular in accounting, finance, management, and marketing research. In general, the estimation of these models involves either probit or logistic regression.[1] Despite their popularity, researchers in these disciplines generally neglect the interpretive potential of the coefficients in these non-linear regression methods. The common interpretation examines the sign and relative magnitude because the coefficients do not have the same degree of interpretability as the coefficients in ordinary least squares (OLS) regression. In the linear probability model, coefficients represent the marginal change in probability associated with a unit change in the exogenous variables. However, in the probit model, since

$$P_i = F(Z_i) = F(\alpha + X\beta) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{Z_i} e^{\frac{-u^2}{2}} du = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\alpha + X\beta} e^{\frac{-u^2}{2}} du \tag{1}$$

where u is a random variable distributed as $N(0,1)$, and

[1]A review of the last ten years of the *Academy of Management Journal, Administrative Science Quarterly, Journal of Marketing Research, Journal of Marketing, Journal of Finance, The Accounting Review, Journal of Accounting Research*, and the *Journal of Accounting and Economics*, the last eight years of the *Journal of Management*, and the last six years of *Marketing Science* reveals twenty-two research studies that employed probit and seventy-two research studies that employed logistic regression. These research studies employed probit and logistic regression in such diverse areas as accounting choice [12] [18] [24], bankruptcy prediction [7] [8] [25] [27], capital structure [3] [22] [32], corporate governance [9] [13] [23], marketing channels [4] [10], and consumer behavior [19] [20]. Amemiya [2] provided a general review of the use of qualitative response models in the economics literature.

*I wish to acknowledge the helpful comments received on earlier drafts of this paper from Felicia B. LeClere. Special thanks are due to Harry H. Kelejian for helpful conversations.

$$Z_i = F^{-1}(P_i) = \alpha + X\beta, \tag{2}$$

the coefficients represent the change in standard units of the normally distributed variable per unit change in the exogenous variables. For logistic regression, since

$$P_i = F(Z_i) = F(\alpha + X\beta) = \frac{1}{1 + e^{-Z_i}} = \frac{1}{1 + e^{-(\alpha + X\beta)}} \tag{3}$$

and

$$Z_i = \log\left(\frac{P_i}{1 - P_i}\right) = \alpha + X\beta, \tag{4}$$

the coefficients represent the change in the log of the odds ratio per unit change in the exogenous variables [1] [15] [17] [21] [29]. In probit or logistic regression, since Z_i (the observed probit or logit) is a monotonically increasing function of P_i, examining the sign of a coefficient indicates the direction of the effect of a change in an independent variable. The relative magnitude of a coefficient indicates the relative influence that a variable has on the probability of choice.

Although sign and magnitude can be informative, there are three alternative methods to interpret the coefficients of probit or logistic regression models.[2] The first method provides an interpretation of the coefficients. It examines the relationship between a change in the probability and a change in an exogenous variable by taking the partial derivative of P_i with respect to the exogenous variable. The partial derivative (with respect to an exogenous variable X_k) for the probit model is

$$\frac{\partial P_i}{\partial X_k} = \frac{1}{\sqrt{2\pi}}\, e^{-\frac{(\alpha + X\beta)^2}{2}}\, \beta_k \tag{5}$$

and

$$\frac{\partial P_i}{\partial X_k} = \frac{e^{-(\alpha + X\beta)}}{(1 + e^{-(\alpha + X\beta)})^2}\, \beta_k \tag{6}$$

for the logistic regression model.

A common and inaccurate interpretation of the partial derivative evaluated at the mean value of the independent variables is as a marginal probability with respect to a unit change in each independent variable. The inaccuracy arises because the magnitude of a partial derivative is a function of the coefficient of the exogenous variables and is not invariant to scale. As a result, the partial derivative has no bound, can be greater than 1, and can generate probabilities outside of the plausible range [28]. The partial derivative misstates the marginal probability and the effect

[2]The ninety-four articles mentioned in footnote 1 provide evidence on the infrequency of alternative interpretation of coefficients. Four research studies provided partial derivatives of the regression models [6] [30] [31], one provided elasticities of probability [14], and four provided limited sensitivity analysis [5] [10] [11] [26].

of a unit change on the dependent variable because both functions are non-linear. The consequence of non-linearity is that each change in the logit or probit does not give the same change in the predicted probability. The probability changes at different rates for different ranges of probits and logits because the value of the derivative depends on where the evaluation occurs in terms of P_i. The benefit of the partial derivative is it may approximate the true marginal probability [28].

A second method of interpretation resolves these difficulties by scaling the partial derivatives and constructing an elasticity of probability. Elasticities are independent of the units which measure the dependent and exogenous variables. Standard economic interpretation of the elasticity of probability is that it indicates the responsiveness of changes in the probability to very small changes in the exogenous variables. A high elasticity (>1) implies that the probability is very responsive to changes in the exogenous variables as there is a greater than proportionate change in the probability relative to the exogenous variable. A low elasticity (<1) implies that the probability is very unresponsive to changes in the exogenous variables as there is a less than proportionate change in the probability relative to the exogenous variable. For the probit model the elasticity of probability is

$$\eta = \frac{X_k}{P_i}\frac{\partial P_i}{\partial X_k} = X_k\left(\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\alpha + X\beta} e^{\frac{-u^2}{2}}\,du\right)^{-1}\left(\frac{1}{\sqrt{2\pi}}\,e^{-\frac{(\alpha + X\beta)^2}{2}}\beta_k\right), \tag{7}$$

and for the logistic regression model

$$\eta = \frac{X_k}{P_i}\frac{\partial P_i}{\partial X_k} = X_k\frac{e^{-(\alpha + X\beta)}}{(1 + e^{-(\alpha + X\beta)})}\beta_k. \tag{8}$$

The unitlessness of the elasticity of probability facilitates direct comparisons among exogenous variables and provides an estimate of the relative effect of the exogenous variables on the probability of choice.

A third method of interpretation is sensitivity analysis. It involves evaluating the coefficients to determine how they affect changes in the probability of choice with changing levels of the exogenous variables. There are two ways to perform this interpretation. The first evaluates the respective functions at the mean of the explanatory variables and uses the cumulative normal density function table (for probit) or evaluates the log of the odds ratio (for logistic regression) to derive the overall probability of choice. Substituting new, hypothetical values in the predicted equations, keeping one or several variables at their mean values, and computing the new overall probability of choice, estimates the effect on the probability that Y_i=1 of a change in one or several of the significant explanatory variables. The second constructs several hypothetical cases by assigning meaningful values to each of the explanatory variables. These values may or may not be the mean values of the exogenous variables. Computing the difference between observed probabilities for two hypothetical cases estimates the effect of a change in one or several explanatory variables on the probability that Y_i=1 for a specific hypothetical case. The benefit provided by sensitivity analysis concerns evaluating the effect on probability of choice of an exogenous variable with a dichotomous measurement. Sensitivity analysis demonstrates the effect of its presence or absence at various levels of the other exogenous variables. In comparison, the elasticity of probability shows the responsiveness of the probability of choice to very small changes in the exogenous variable. Yet categorical variables change in units of one (1 to 0 or 0 to 1) rather than in very small amounts. A very small change in a categorical variable is not mathematically strictly permissible nor does it correspond to any social phenomena [28].

APPLICATION

This section applies the three interpretation techniques discussed above to the coefficients in a logistic regression model contained in a published study by Hill and Ingram (HI) [16]. They attempted to explain why savings and loan associations choose a Regulatory Accounting Principle (RAP) over a Generally Accepted Accounting Principle (GAAP) and argued that firms will be more likely to adopt the RAP the greater (1) the interaction between potential tax benefits and closeness to the minimum net worth requirement, and (2) the interaction between potential tax benefits and prior experience of selling loans in the secondary market. Savings and loan associations will also be likely to adopt a RAP if not regulated by the SEC and if their audit firm supports the choice of a RAP. Control variables are size, measured as total assets, and leverage. Table 1 presents the coefficients, means of the variables, partial derivatives, and elasticities.[3]

Two points illustrate the difficulty of interpreting the partial derivatives as marginal probabilities. First, since the function is non-linear, it is incorrect to state that a unit change in tax-net worth results in a 8900.82 percent increase in the probability of choice. Because the partial derivative has no bound, it generates a marginal probability outside the plausible range for a one unit increase. The actual marginal effect on the probability of choice for very small increases (dx) in tax-net worth is an increase of 89.0082dx. Second, the variables have different scales and the marginal probabilities are not independent of scale. Tax-net worth ([taxes/assets] × [.03/(net worth/total deposits)]), market-net worth (extent × [.03/(net worth/total deposits)]), and leverage (net worth/total liabilities) are ratios of different magnitudes and size (total assets) is a continuous variable. (The "extent" measure in the market-net worth variable is a dichotomous variable that measures prior experience in the secondary loan market.) Direct comparisons of the marginal probabilities are of limited use because of the difference in scales. For example, tax-net worth has a mean of .0049 while market-net worth has a mean of .3246. The realization that a one unit increase is a two hundred and four-fold increase in the mean of tax-net worth tempers the magnitude of the partial derivative of tax-net worth.

These two difficulties suggest examining the elasticities of probabilities. Elasticities facilitate direct comparisons because they are independent of scale. The elasticities are .8550, .3913, -.1819, and .0. The exogenous variables have low elasticities (all elasticities have absolute values less than one) and the probability of choice is not very responsive to changes in the exogenous variables. However, there are substantial differences among the elasticities. As an example, a very small

[3] The coefficients are from the logistic regression model presented in Table 4 of [16]. Table 2 of [16] contains the means of the variables. The dependent variable is 1 for firms that used a RAP and 0 for firms that used GAAP. Table 1 omits the partial derivative and elasticity of probability of the function with respect to SEC regulation and auditor because both variables are categorical variables.

Table 1: Logistic regression of choice between RAP and GAAP.

	β_K	$\overline{X}_k$	$\frac{\partial P_i}{\partial X_k}$	η
Intercept	-3.503	—	—	—
Tax-Net Worth	356.180	.0049	89.0082	.8550
Market-Net Worth	2.461	.3246	.6150	.3913
SEC Regulation	-3.059	.1062	—	—
Auditor	1.973	.8584	—	—
Size	-.192	1.9343	-.0480	-.1819
Leverage	.0001	19.8663	.0000	.0000

change in tax-net worth and size will result in a less than proportionate increase in the probability of the firm electing a RAP but the increase in probability is equal to .8550 of the change in tax-net worth and .1819 of the change in size. Choice of a RAP or GAAP is more elastic with respect to tax-net worth than it is to market-net worth, size, or leverage.

Table 2 contains the results of sensitivity analysis. It presents the estimated probabilities of a hypothetical firm electing a RAP. The analysis examines the estimated probabilities at four levels of tax-net worth (mean value and 10 percent, 20 percent, and 30 percent increases) and the four possible combinations of SEC and auditor support while holding market-net worth, size, and leverage at the mean values. This estimates the effect of SEC regulation and auditor support, at increasing levels of tax-net worth, on the probability of electing a RAP. The results show that when the firm is SEC regulated and does not have auditor support, the probability of electing a RAP is very small. Auditor support increases the probabilities of adoption, despite SEC regulation. The probabilities increase again when there is no auditor support but the firm is not subject to SEC regulation. The probabilities increase dramatically for a firm that is not SEC regulated but has auditor support. The effect of an increase in tax-net worth on the probability of choice is dependent on the level of the other variables. For a regulated firm with no auditor support, the effect of tax-net worth on the probability of electing a RAP is very small (an .8 percent increase from the mean level to the mean plus 30 percent level), but for a non-regulated firm (with or without auditor support), the effect is substantial (increases of 10.2 percent and 10.9 percent, respectively, from the mean level to

Table 2: Estimated probabilities of electing RAP versus GAAP for a hypothetical firm.

	Level of Tax-Net Worth			
Characteristics	Mean	+10%	+20%	+30%
SEC and NO AUD	.012	.014	.017	.020
SEC and AUD	.082	.096	.113	.132
NO SEC and NO AUD	.209	.240	.274	.311
NO SEC and AUD	.655	.694	.731	.764

the mean plus 30 percent level). These relationships are not apparent from the sign and relative magnitude of the estimated coefficients.

CONCLUSIONS

This paper has demonstrated and applied three methods of interpreting coefficients in probit and logistic regression models. The first and second method interpret the coefficients as marginal probabilities and elasticities of probability. These methods quantify the responsiveness of probability of choice to changes in the exogenous variables. The paper discussed difficulties in interpreting the partial derivative as a marginal probability. The third method uses sensitivity analysis to analyze the effect of changes in the exogenous variables on the probability of choice. This method allows control over units of change, comparisons across meaningful levels of variables, and demonstrates changes in the size of the probability of choice within a limited range. Because a large amount of accounting, finance, management, and marketing research uses probit and logistic regression, these interpretation methods warrant consideration in the presentation of the empirical results of such research. [Received: December 10, 1990. Accepted: September 23, 1991.]

REFERENCES

[1] Aldrich, J. H., & Nelson, F. D. *Linear probability, logit, and probit models.* Beverly Hills: Sage Publications, 1984.

[2] Amemiya, T. Qualitative response models: A survey. *Journal of Economic Literature,* 1981, *19*(4), 1483-1536.

[3] Amihud, A., Lev, B., & Travlos, N. G. Corporate control and the choice of investment financing: The case of corporate acquisitions. *Journal of Finance,* 1990, *45*(2), 603-616.

[4] Anderson, E., & Coughlan, A. T. International market entry and expansion via independent or integrated channels of distribution. *Journal of Marketing,* 1987, *51*(1), 71-82.

[5] Barnett, W. P., & Carroll, G. R. Competition and mutualism among early telephone companies. *Administrative Science Quarterly,* 1987, *32*(3), 400-421.

[6] Cameron, T. A., & James, M. D. Estimating willingness to pay from survey data: An alternative pre-test-market evaluation procedure. *Journal of Marketing Research,* 1987, *24*(4), 389-395.

[7] Casey, C., & Bartczak, N. Using operating cash flow data to predict financial distress: Some extensions. *Journal of Accounting Research,* 1985, *23*(1), 384-401.

[8] Casey, C., McGee, V. E., & Stickney, C. P. Discriminating between reorganized and liquidated firms in bankruptcy. *The Accounting Review,* 1986, *61*(2), 249-262.

[9] Cochran, P. P., Wood, R. A., & Jones, T. B. The composition of board of directors and incidence of golden parachutes. *Academy of Management Journal,* 1985, 28(3), 664-671.

[10] Coughlan, A. T. Competition and cooperation in marketing channel choice: Theory and application. *Marketing Science,* 1985, *4*(2), 110-129.

[11] Coughlan, A. T., & Schmidt, R. M. Executive compensation, management turnover, and firm performance. *Journal of Accounting and Economics,* 1985, *7*(1-3), 43-66.

[12] Dhaliwal, D. S., Salamon, G., & Smith, E. D. The effect of owner versus management control on the choice of accounting methods. *Journal of Accounting and Economics,* 1982, *4*(1), 41-53.

[13] Friedman, S. D., & Singh, S. CEO succession and stockholder reaction: The influence of organizational context and event content. *Academy of Management Journal,* 1989, *32*(4), 718-744.

[14] Gautschi, D. A. Specification of patronage models for retail center choice. *Journal of Marketing Research,* 1981, *18*(2), 162-175.

[15] Hanushek, E. A., & Jackson, J. E. *Statistical methods for social scientists.* Orlando, FL: Academic Press, 1977.

[16] Hill, J. W., & Ingram, R. W. Selection of GAAP or RAP in the savings and loan industry. *The Accounting Review,* 1989, *64*(4), 667-679.

[17] Jackson, B. B. *Multivariate data analysis.* Homewood, IL: Richard D. Irwin, 1983.

[18] Johnson, W. B., & Ramanan, R. Discretionary accounting changes from 'successful efforts' to 'full cost' methods. *The Accounting Review,* 1988, *63*(1), 96-110.

[19] Jones, J. M., & Zufryden, F. S. An approach for assessing demographic and price influences on brand purchase behavior. *Journal of Marketing*, 1982, *46*(1), 36-46.
[20] Jones, J. M., & Zufryden, F. S. Adding explanatory variables to a consumer purchase behavior model: An exploratory study. *Journal of Marketing Research*, 1980, *17*(3), 323-334.
[21] Maddala, G. S. *Limited-dependent and qualitative variables in econometrics*. New York: Cambridge University Press, 1983.
[22] Marsh, P. The choice between equity and debt: An empirical study. *Journal of Finance*, 1982, *37*(1), 121-144.
[23] Mizruchi, M. S., & Stearns, L. B. A longitudinal study of the formation of interlocking directorates. *Administrative Science Quarterly*, 1988, *33*(2), 194-210.
[24] Niehaus, G. R. Ownership structure and inventory method choice. *The Accounting Review*, 1989, *64*(2), 269-284.
[25] Ohlson, J. A. Financial ratios and the probabilistic prediction of bankruptcy. *Journal of Accounting Research*, 1980, *18*(1), 109-131.
[26] Palmer, D., Friedland, R., Jenning, P. D., & Powers, M. E. The economics and politics of structure: The multidivisional form and the large U. S. corporation. *Administrative Science Quarterly*, 1987, *32*(1), 25-48.
[27] Pastena, V., & Ruland, W. The merger/bankruptcy alternative. *The Accounting Review*, 1986, *61*(2), 288-301.
[28] Petersen, T. A comment on presenting results from logit and probit models. *American Sociological Review*, 1985, *50*(1), 130-131.
[29] Pindyck, R. S., & Rubinfeld, D. L. *Econometric models and economic forecasts*. New York: McGraw-Hill, 1981.
[30] Pisano, G. P. The R&D boundaries of the firm: An empirical analysis. *Administrative Science Quarterly*, 1990, *35*(1), 153-176.
[31] Romanelli, E. Environments and strategies of organization start-up: Effects on early survival. *Administrative Science Quarterly*, 1989, *34*(3), 369-387.
[32] Smith, R. L. The choice of issuance procedure and the cost of competitive and negotiated underwriting: An examination of the impact of rule 50. *Journal of Finance*, 1987, *42*(3), 703-720.

Marc J. LeClere is Assistant Professor of Accounting and Stoy Malone & Company Fellow at the University of Maryland. Dr. LeClere received his Ph.D. in business administration from The Pennsylvania State University. His current research interests include the choice of accounting methods by firms, corporate and individual tax policy, and research methods. His articles have appeared in such journals as *The Accounting Review* and *Advances in Taxation*.