A Comparison of Logistic Regression Pseudo R^2 Indices

Thomas J. Smith
Northern Illinois University

Cornelius M. McKenna
Kishwaukee College

The present study examined seven distinct pseudo R^2 indices used in logistic regression, how values of these indices compared to values of ordinary least squares (OLS) R^2 obtained under similar conditions, and how values of these indices varied as a function of multicollinearity among predictors and base rate of the dependent variable. Monte Carlo simulation methods suggested that the Aldrich-Nelson pseudo R^2 index with Veall-Zimmermann correction resulted in values that most closely approximated the OLS R^2 values. Additionally, lower multicollinearity among predictors was associated with increased variability among the values resulting from the various indices. Changes in base rate had little effect on "corrected" versions of the indices, but did affect uncorrected versions of the indices.

Binary logistic regression is a frequently applied procedure used to predict the probability of occurrence for some binary outcome, using one or more continuous or categorical variables as predictors. When the outcome is expressed as the log-odds of the event's occurrence, the logistic regression equation is a linear combination of the predictors, where the regression parameters are typically obtained using maximum likelihood estimation, and where each regression weight indicates the change in the log-odds of the event's occurrence per unit of change in its associated predictor.

Adequacy of fit for a logistic regression model is typically assessed by assessing (1) the significance of the omnibus chi-square test of the model coefficients, which assesses the incremental decrease in the log-likelihood (i.e., deviance) of the regression model containing the full set of predictors when it is compared to the model that contains only the intercept term, and determines whether the former significantly improves prediction over the latter; and (2) the Hosmer-Lemeshow goodness-of-fit test (Hosmer & Lemeshow, 2000), which groups cases into deciles based upon the predicted probability of each, then assesses the degree to which the observed frequencies match the expected frequencies using a chi-square goodness-of-fit test, and where a non-significant test result suggests a well-fitting model. Additionally, when examining individual predictors, the adjusted odds-ratio (i.e., the exponentiated regression coefficient) associated with each predictor can be evaluated as an effect size.

When the predicted probabilities resulting from logistic regression are used for classification purposes, additional indices of model fit are often employed. Simple proportions of correctly classified cases, both for the overall sample as well as for each of the groups in the sample provide one such index. Also, a Receiver Operator Characteristic (ROC) curve, which graphically represents "true positive" and "false positive" classification rates as a function of different classification cutoff values for the predicted probabilities resulting from the logistic regression provides another such index.

In addition, a number of goodness-of-fit indices exist to assess the predictive capacity of the logistic regression model. These "pseudo R^2" indices have been developed that are intended as logistic regression analogs of R^2 as used in ordinary least-squares (OLS) regression. One such index, outlined by Maddala (1983) and Cox and Snell (1989), is:

$$R^2_{MCS} = 1 - \left(\frac{L(Null)}{L(Full)} \right)^{2/N},$$

where $L(Null)$ and $L(Full)$ are the likelihood functions for the intercept-only model and full model, respectively. The R^2_{MCS} statistic is interpretable as the geometric mean square improvement. Because the value of R^2_{MCS} can potentially exceed 1.0, Cragg and Uhler (1970), Maddala (1983), and Nagelkerke (1991) describe a rescaling of this statistic,

$$R^2_{NK} = \frac{1 - \left(L(Null) / L(Full) \right)^{2/N}}{1 - L(Null)^{2/N}},$$

where the rescaling is accomplished by dividing R^2_{MCS} by its maximum possible value. The resulting statistic then has a range that is identical to the range of OLS R^2.

McFadden (1974) outlines perhaps the most straightforward of such pseudo R^2 indices, in the sense of reflecting both the criterion being minimized in logistic regression estimation and the variance-accounted-for by the logistic regression model. This log likelihood ratio R^2 (sometimes referred to as "deviance R^2") is one minus the ratio of the full-model log-likelihood to the intercept-only log-likelihood,

$$R^2_{MF} = 1 - \frac{LL(Full)}{LL(Null)}.$$

This index can also be adjusted to penalize for the number of predictors (k) in the model,

$$R^2_{MFA} = 1 - \frac{LL(Full) - k}{LL(Null)},$$

(Mittlböck & Heinzl, 2004). Lave (1970) and Efron (1978) suggest a direct analog to R^2 as used in OLS regression,

$$R^2_{LE} = 1 - \frac{\sum_{i=1}^{N}(y_i - \hat{p}_i)^2}{\sum_{i=1}^{N}(y_i - \bar{y})^2},$$

where $y_i, \bar{y}_i$, and $\hat{p}_i$ are the observed binary outcomes, mean outcome, and predicted probabilities, respectively. Finally, Adlrich and Nelson (1984) propose a variant of the contingency coefficient defined in Agresti and Finlay (1986),

$$R^2_{AN} = \frac{G(M)}{G(M) + N},$$

where $G(M)$ is the incremental change in the -2 log likelihood function when the full model and the intercept-only model are compared (i.e., the model chi-square statistic).

Menard (2000), in a study comparing five distinct pseudo R^2 indices ($R^2_{MF}, R^2_{AN}, R^2_{MCS}, R^2_{NK}$, and R^2_{LE}) in a logistic regression context, concluded that McFadden's index was preferred due to both its conceptual similarity to OLS R^2 (as used in linear regression), and due its relative independence of the base rate of the binary outcome variable (i.e., the relative occurrence of a "positive/successful" outcome). The author also suggests that additional study is necessary to discern the effects of varying base rates on these indices. Windmeijer (1995), in a comparison of seven pseudo R^2 indices, concluded that an index proposed by McKelvey-Zavoina (1975),

$$R^2_{MZ} = \frac{\sum_{i=1}^{N}(\hat{y} - \bar{\hat{y}})^2}{(N \times 3.29) + \sum_{i=1}^{N}(\hat{y} - \bar{\hat{y}})^2},$$

which reflects decomposition of the variance of the estimated logits, provides the best estimate of the OLS estimator. This index can be difficult to compute, however, and Veall and Zimmermann (1994) propose a correction to the Aldrich-Nelson index that closely approximates its value,

$$R^2_{VZ} = \frac{R^2_{AN}}{\max(R^2_{AN} / \hat{p})},$$

where $\hat{p}$ is the estimated base rate for the binary outcome, and where the maximum value of the index conditional on this base rate is

$$\max(R^2_{AN} / \hat{p}) = \frac{-2[\hat{p}\log(\hat{p}) + (1-\hat{p})\log(1-\hat{p})]}{1 - 2[\hat{p}\log(\hat{p}) + (1-\hat{p})\log(1-\hat{p})]}.$$

The purpose of the present study was to examine, compare, and classify seven distinct pseudo R^2 indices ($R^2_{MCS}, R^2_{NK}, R^2_{MF}, R^2_{MFA}, R^2_{LE}, R^2_{AN}$, and R^2_{VZ}) used in binary logistic regression. Specifically, values for these indices were compared under varying correlational conditions and also under varying underlying base rates of the binary outcome, and these values were also compared to the R^2 values resulting from OLS linear regression based on a continuous outcome. A hierarchical classification of these seven pseudo R^2 indices was then developed based on their pairwise similarity.

Table 1. Correlations among Variates for Simulated Regression Data

	Condition 1 (r = .10)					Condition 2 (r = .30)					Condition 3 (r = .50)				
	IV1	IV2	IV3	IV4	DV	IV1	IV2	IV3	IV4	DV	IV1	IV2	IV3	IV4	DV
IV1	1.0					1.0					1.0				
IV2	.1	1.0				.3	1.0				.5	1.0			
IV3	.1	.1	1.0			.3	.3	1.0			.5	.5	1.0		
IV4	.1	.1	.1	1.0		.3	.3	.3	1.0		.5	.5	.5	1.0	
DV	.5	.5	.5	.5	1.0	.5	.5	.5	.5	1.0	.5	.5	.5	.5	1.0
	Condition 4 (r = .70)					Condition 5 (r = .90)									
IV1	1.0					1.0									
IV2	.7	1.0				.9	1.0								
IV3	.7	.7	1.0			.9	.9	1.0							
IV4	.7	.7	.7	1.0		.9	.9	.9	1.0						
DV	.5	.5	.5	.5	1.0	.5	.5	.5	.5	1.0					

Method

The present study used Monte Carlo techniques to examine, compare, and classify seven pseudo R^2 goodness-of-fit indices. Specifically, samples of size of n = 200 were randomly drawn from a multivariate normal parent population, where each sample consisted of five continuous variables with one of five specified correlational structures representing five distinct multicollinearity conditions. Four of these variables then served as predictors in a logistic regression. The fifth variable was transformed to a binary (0/1) outcome variable by carrying out a split of the cases. The cut point was specified as one five values (50th, 60th, 70th, 80th, or 90th percentile), representing five underlying base rate conditions (p = .50, .40, .30, .20, and .10, respectively) for the "success" (coded as 1) outcome.

Binary logistic regression was carried out on each sample, and seven pseudo R^2 indices were computed ($R^2_{MCS}, R^2_{NK}, R^2_{MF}, R^2_{MFA}, R^2_{LE}, R^2_{AN}$, and R^2_{VZ}). Additionally, R^2 and adjusted R^2 were computed using OLS regression, where the predictors were the same predictors used in the logistic regression, and the outcome variable consisted of values for the fifth, original, continuous variable (prior to dichotomization).

Each simulation consisted of 500 randomly drawn samples, where each set of samples was drawn from one of five multicollinearity conditions (that represented low to high levels of multicollinearity and in which the outcome variable was moderately correlated with these predictors; see Table 1), and one of five split point conditions. We then examined how varying the correlational structure among the generated variables, as well as the split point for dichotomizing the outcome variable (i.e., the effective base rate for the dichotomous outcome) affected the resulting pseudo R^2 values. We also compared the values of the pseudo R^2 indices to the R^2 values resulting from the corresponding OLS linear regression. All analyses were carried out using SPSS v.19.

Results

Table 2 provides the mean goodness-of-fit indices (i.e., pseudo R^2 values from the various indices, OLS R^2, and OLS adjusted R^2) for the complete set of 12,500 simulated samples. The mean values across all conditions are displayed in Figure 1. As can be seen, across all conditions, the values for the pseudo R^2 indices varied widely (from M = .23 to M = .44). These differences were statistically significant (p < .001), but significance here could easily be an artifact of the large number of simulated samples. However, the effect size for these differences was large (η^2 = .79), suggesting meaningful differences existed. Interestingly, the lowest observed mean index values were for the Aldrich-Nelson index, while the highest mean values were for the corrected version of the Aldrich-Nelson index (the Veall-Zimmermann corrected index). The Aldrich-Nelson index also showed the least variability in values (SD = .09), while the "corrected" indices (i.e., the Veall-Zimmerman and Nagelkerke indices) showed the greatest variability (SD = .17). The Veall-Zimmermann index provided the closest approximation to both the OLS R^2 and OLS adjusted R^2 (where the latter two values were based on the continuous outcome variable).

Table 2. Statistics for Goodness-of-Fit Indices

	n	*M*	*(SD)*
Maddala / Cox-Snell	12,500	.26	(.12)
Nagelkerke	12,500	.40	(.17)
McFadden	12,500	.29	(.15)
McFadden adjusted	12,500	.25	(.15)
Lave/Efron	12,500	.31	(.15)
Aldrich-Nelson	12,500	.23	(.09)
Veall-Zimmermann	12,500	.44	(.17)
OLS R^2*	12,500	.47	(.18)
OLS adjusted R^2*	12,500	.46	(.18)

Note. *Based on continuous outcome.

Table 3 provides descriptive statistics for the seven pseudo R^2 indices (in addition to OLS R^2 and adjusted R^2) by base rate of the binary dependent variable, and Figure 2 displays the mean values of these indices. In each of the base rate conditions, the Veall-Zimmermann index again most closely approximated the OLS R^2 statistic. Across base rate conditions, the Aldrich-Nelson pseudo adjusted R^2 index was the least variable, while the two "corrected" indices (Nagelkerke and Veall-Zimmermann) appeared to show the greatest variability across base rate conditions. The main effect of base rate on the mean index (across index types) was statistically significant ($p < .001$), but the effect size was very small, with $\eta^2 < .01$. However, the interactive effect of base rate × index type was moderate in size ($\eta^2 = .09$), suggesting that the effect of base rate on the index values differed by index type. Specifically, base rate had the largest effect on the Maddala / Cox-Snell and Aldrich-Nelson (uncorrected) indices, with low base rates associated with lower pseudo R^2 values. Interestingly, the lowest base rate ($p = .10$) actually increased the value of McFadden's (unadjusted) index.

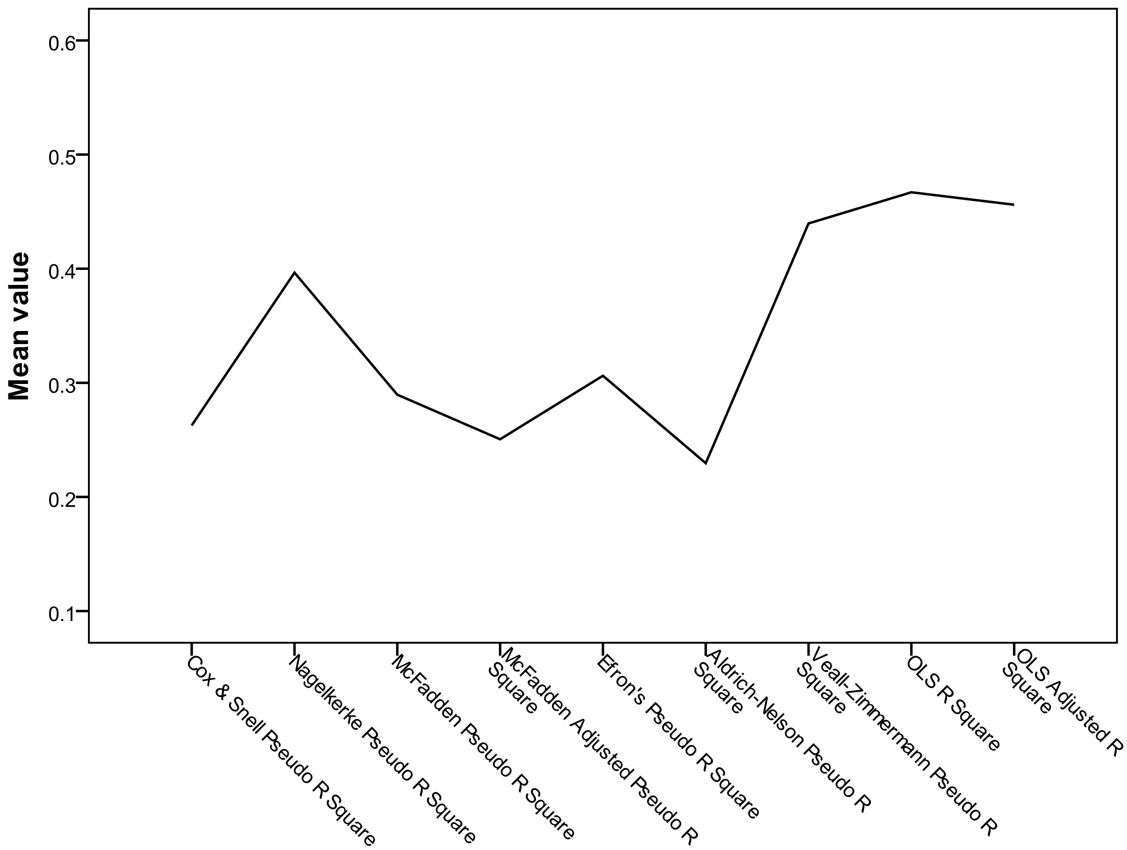

Figure 1. Mean values of regression goodness-of-fit indices for simulated regression data based on 12,500 sample replications.

Table 4 provides descriptive statistics for the indices by multicollinearity condition, and Figure 3 displays these mean values. Here, again, the multicollinarity condition × index type

Table 3. Descriptive Statistics for Regression Goodness-of-Fit Indices by Base Rate

	Base rate											
	p = .10		p = .20		p = .30		p = .40		p = .50		Total	
Index	*M*	*(SD)*	*M*	*(SD)*	*M*	*(SD)*	*M*	*(SD)*	*M*	*(SD)*	*M*	*(SD)*
Maddala / Cox-Snell	.18	(.08)	.24	(.10)	.28	(.12)	.30	(.12)	.31	(.13)	.26	(.12)
Nagelkerke	.38	(.17)	.38	(.16)	.40	(.17)	.41	(.16)	.41	(.17)	.40	(.17)
McFadden	.32	(.16)	.29	(.15)	.28	(.14)	.28	(.14)	.28	(.14)	.29	(.15)
McFadden adjusted	.25	(.16)	.25	(.15)	.25	(.14)	.25	(.14)	.25	(.14)	.25	(.15)
Lave/Efron	.27	(.17)	.29	(.15)	.31	(.15)	.33	(.15)	.33	(.15)	.31	(.15)
Aldrich-Nelson	.16	(.07)	.21	(.08)	.24	(.09)	.26	(.09)	.27	(.10)	.23	(.09)
Veall-Zimmermann	.42	(.18)	.43	(.17)	.44	(.17)	.45	(.16)	.46	(.17)	.44	(.17)
OLS R^2*	.47	(.18)	.47	(.18)	.47	(.18)	.47	(.18)	.47	(.18)	.47	(.18)
OLS adjusted R^2*	.46	(.18)	.46	(.18)	.46	(.18)	.46	(.18)	.46	(.18)	.46	(.18)

Note. *Based on continuous outcome. n = 2500 for each Condition.

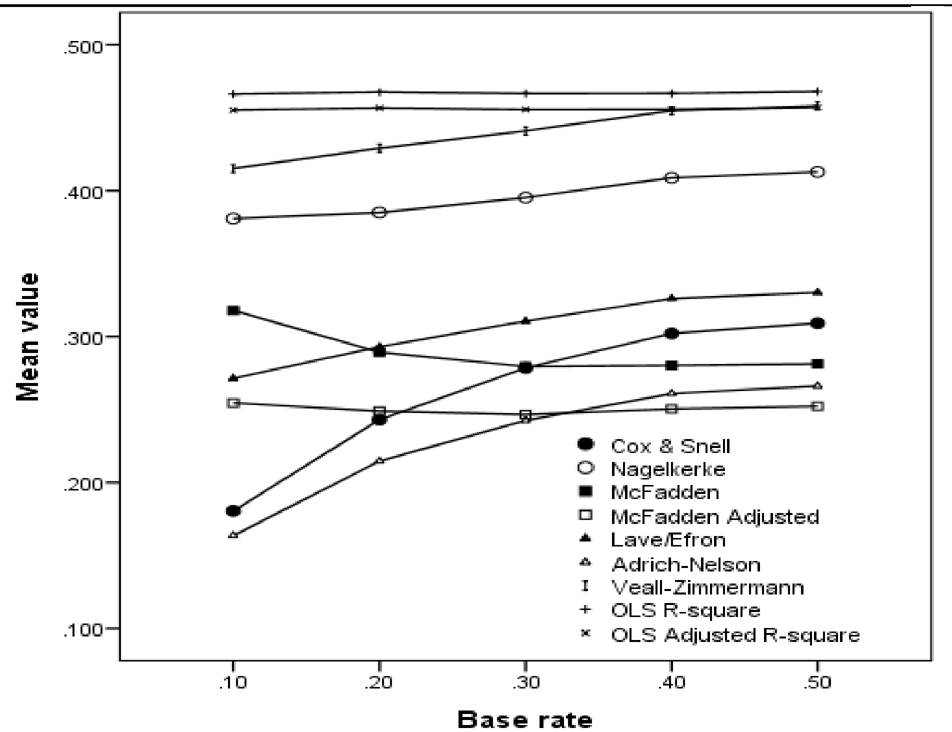

Figure 2. Mean values of regression goodness-of-fit indices by base rate for simulated regression data based on 2500 sample replications per base rate level.

interactive effect was statistically significant ($p < .001$), but the effect size was small (η^2 = .03). Results indicated that the Veall-Zimmermann index once again most closely approximated the OLS R^2 values under each multi-collinearity condition. As would be expected, lower levels of multicollinearity (e.g., r = .10) among the predictors resulted in higher pseudo R^2 values. However, as Figure 3 indicates, lower levels of multicollinearity resulted in greater variability among the mean values of the various pseudo R^2 indices. At the lowest multicollinearity level (r = .10), the mean values of the indices ranged from .37 (Aldrich Nelson index) to .67 (Nagelkerke corrected index).

Finally, Figure 4 displays the mean R^2 indices by both base rate and multicollinearity condition. Although the three-way base rate × multicollinearity condition × index type interaction effect was statistically significant ($p < .001$), this was an artifact of the large number of simulations (500), as the observed effect size was very small (η^2 = .005). That is, and as can be observed in these means plots, the two-way base rate × multicollinearity interaction effect was similar across index types.

We next attempted to develop a hierarchical classification of these seven pseudo R^2 indices. To this end, we first computed the correlations among the indices (see upper diagonal of Table 5). These correlations served as input measures of similarity for complete linkage cluster analysis (McQuitty, 1960). Figure 5 shows the resulting dendrogram. As Figure 5 illustrates, two primary clusters of pseudo

Table 4. Descriptive Statistics for Regression Goodness-of-Fit Indices by Multicollinearity

	Multicollinearity Level											
	Condition 1 (r = .10)		Condition 2 (r = .30)		Condition 3 (r = .50)		Condition 4 (r = .70)		Condition 5 (r = .90)		Total	
Index	*M*	*(SD)*	*M*	*(SD)*	*M*	*(SD)*	*M*	*(SD)*	*M*	*(SD)*	*M*	*(SD)*
Maddala / Cox-Snell	.44	(.09)	.30	(.07)	.23	(.06)	.19	(.06)	.16	(.05)	.26	(.12)
Nagelkerke	.67	(.06)	.45	(.07)	.34	(.07)	.28	(.07)	.24	(.07)	.40	(.17)
McFadden	.54	(.07)	.32	(.06)	.24	(.06)	.19	(.06)	.16	(.05)	.29	(.15)
McFadden adjusted	.50	(.07)	.28	(.06)	.20	(.06)	.15	(.05)	.12	(.05)	.25	(.15)
Lave/Efron	.56	(.07)	.34	(.07)	.25	(.07)	.20	(.06)	.17	(.06)	.31	(.15)
Aldrich-Nelson	.37	(.06)	.26	(.06)	.20	(.05)	.17	(.05)	.15	(.04)	.23	(.09)
Veall-Zimmermann	.70	(.07)	.50	(.08)	.39	(.08)	.32	(.08)	.28	(.08)	.44	(.17)
OLS R^2*	.77	(.03)	.53	(.05)	.41	(.05)	.33	(.06)	.28	(.05)	.47	(.18)
OLS adjusted R^2*	.77	(.03)	.52	(.05)	.40	(.05)	.32	(.06)	.27	(.05)	.46	(.18)

Note. *Based on continuous outcome. Level. n = 2500 for each Condition.

R^2 indices are apparent. The first cluster consists of the Maddala / Cox-Snell and Aldrich-Nelson indices—both of which are log-likelihood based. The second cluster consists of the remaining indices. Within this second cluster, two secondary clusters emerge, consisting of (1) the McFadden and McFadden adjusted indices, and (2) the Lave/Efron, Veall-Zimmermann, and Nagelkerke indices.

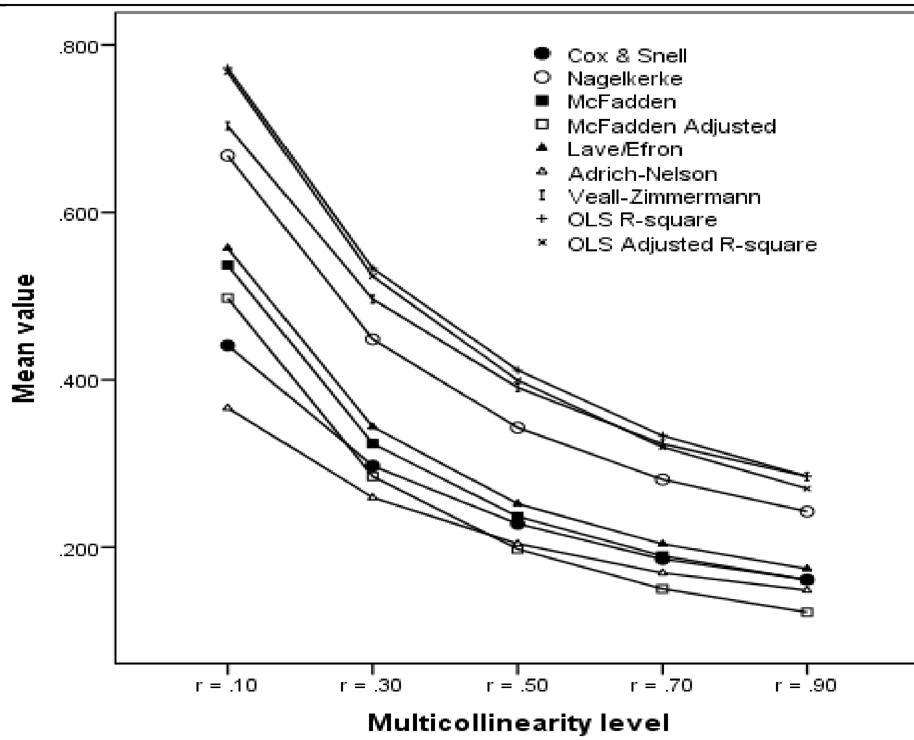

Figure 3. Mean values of regression goodness-of-fit indices by multi-collinearity condition for simulated regression data based on 2500 sample replications per multicollinearity level.

It is important to note that the correlations among the pseudo R^2 indices, and thus the resultant hierarchical cluster analysis of the indices based on these correlations, are based on the profile similarity of the indices. That is, a particular pair of indices is considered to be similar if, for a series of regression analyses, their values closely profile or "track" one another (i.e., when the value of one index increases/decreases, the value of the other index also increases/decreases). A pair of indices may correlate very strongly, however, and yet be quite distinct in magnitude. Therefore, an alternate indicator of (dis)similarity that might be considered is the Euclidean distance between all pairs of indices. Figure 6 shows a complete linkage clustering of the indices using Euclidean distance values (lower diagonal of Table 5) as measures of proximity. As can be

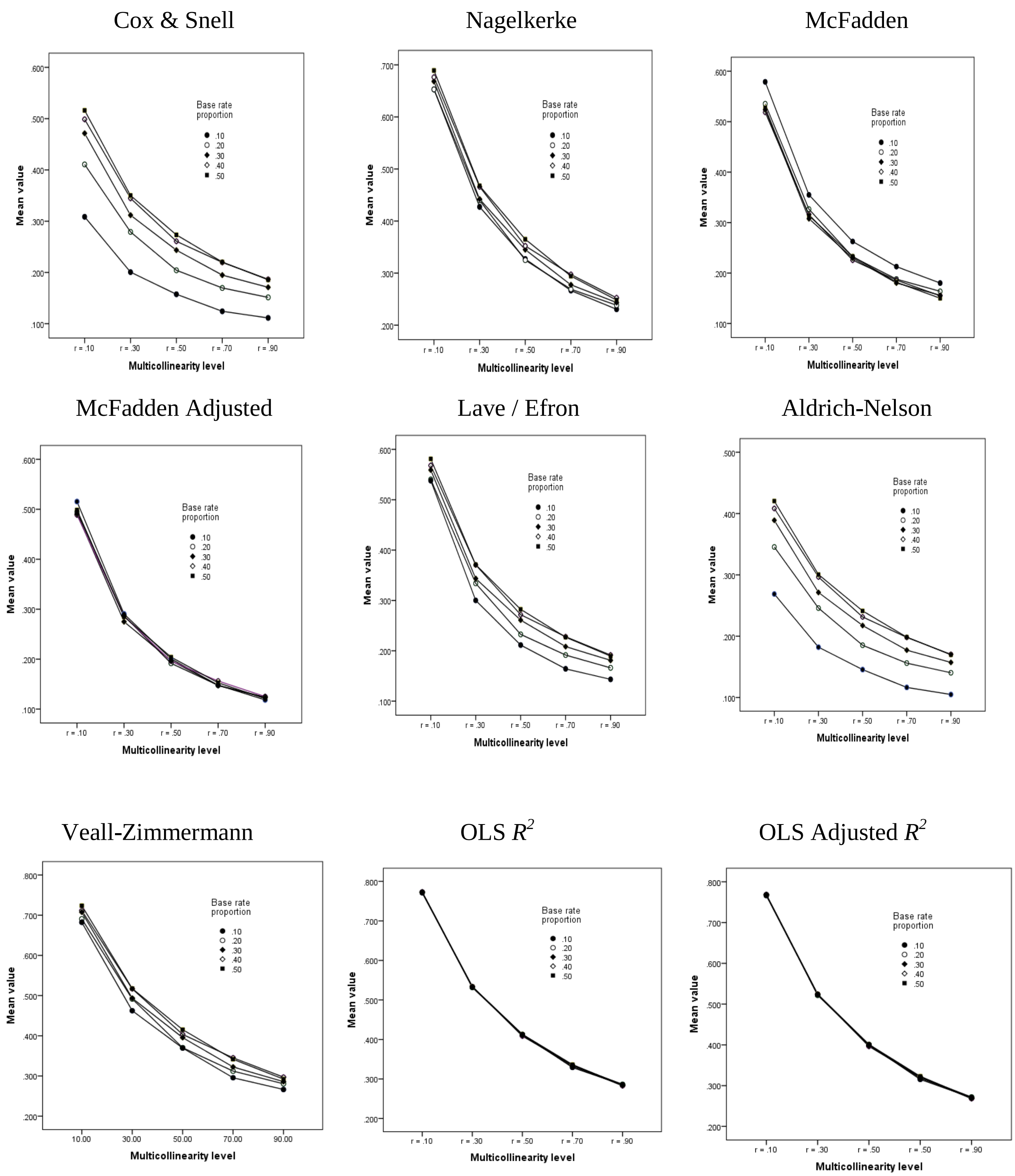

Figure 4. Mean values of regression goodness-of-fit indices by multicollinearity condition and base rate for simulated regression data for Cox & Snell, Nagelkerke, McFadden, McFadden Adjusted, Lave/Efron, and Aldrich-Nelson, Veall-Zimmerman pseudo R^2 index, OLS R^2 , and OLS Adjusted R^2.

Table 5. Correlations (Upper Diagonal) and Euclidean Distances (Lower Diagonal) among Pseudo R^2s

	Maddala / Cox-Snell	Nagelkerke	McFadden	McFadden adjusted	Lave / Efron	Aldrich-Nelson	Veall-Zimmermann
Maddala / Cox-Snell		0.93	0.85	0.89	0.94	0.99	0.94
Nagelkerke	16.92		0.98	0.99	0.99	0.93	0.99
McFadden	9.16	12.60		0.99	0.96	0.85	0.96
McFadden adjusted	7.63	16.61	4.62		0.98	0.89	0.98
Lave/Efron	7.97	10.55	4.84	7.00		0.94	0.98
Aldrich-Nelson	4.77	21.03	11.53	8.83	11.73		0.94
Veall-Zimmermann	21.21	5.43	17.60	21.58	15.38	25.35	

seen in this dendrogram, two primary clusters emerge that are distinct from the primary clusters observed in Figure 5. The first cluster consists of the Nagelkerke and Veall-Zimmerman indices, while the second cluster consists of the remaining indices. Within this second cluster, the Lave/Efron and McFadden indices form one subcluster, while the Aldrich-Nelson and Maddala / Cox-Snell indices form a second subcluster.

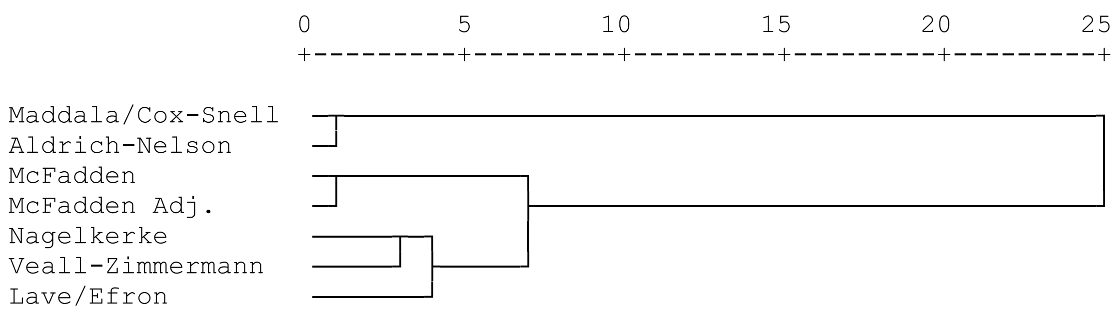

Figure 5. Complete-Linkage Hierarchical Clustering of Pseudo R^2 Indices based on Correlations.

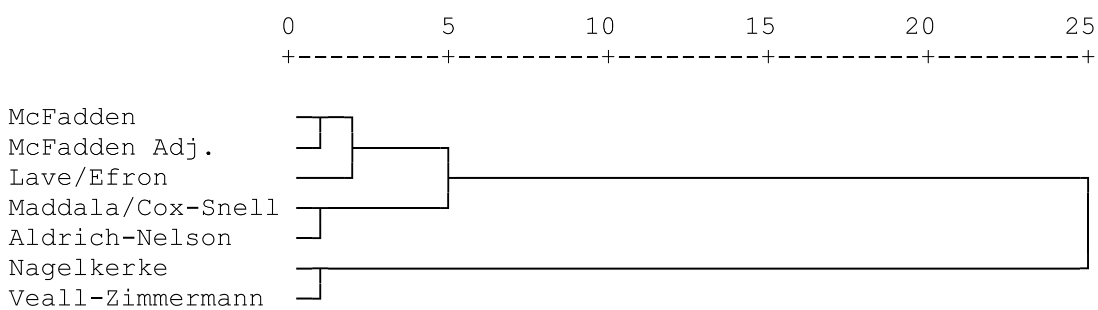

Figure 6. Complete-Linkage Hierarchical Clustering of Pseudo R^2 Indices based on Euclidean Distances.

Discussion

Although pseudo R^2 values for logistic regression are available as output in most statistical packages and are often reported in practice, few if any guidelines exist for their interpretation. The present study suggested that the most commonly used pseudo R^2 indices (e.g., McFadden's index, Maddala / Cox-Snell index with or without Nagelkerke correction) yield lower estimates than their OLS R^2 counterparts, evidence that is consistent with prior simulation studies (e.g., Hagle & Mitchell, 1992; Veall & Zimmermann, 1994). This suggests that the use of guidelines intended for interpretation of the latter (e.g., Cohen, 1988) may not be appropriate for interpreting pseudo R^2 values. One pseudo R^2 index, the Aldrich-Nelson index with Veall-Zimmermann correction resulted in values that most closely approximated the OLS R^2 values, which is consistent with the findings of Veall and Zimmermann. In comparison, values of Nagelkerke's index were somewhat lower, but nearly as close to the OLS R^2 values. It is difficult to assert, however, that the ability to closely mirror the value that an OLS R^2 estimate is necessarily a desirable quality of a particular pseudo R^2 index. That is, OLS linear regression minimizes a least-squares criterion, and the OLS R^2 and adjusted R^2 indices are thus both intended to reflect this optimization. Holding this optimization choice up as the "gold standard" for pseudo R^2 index to emulate may be misguided, particularly when logistic regression parameter estimates are typically not estimated using a least-squares optimization procedure. Instead, perhaps a unique (and less stringent) set of guidelines may be appropriate for the interpretation of pseudo R^2 values.

The present study also found that, as would be expected, increased levels of multicollinearity resulted in decreased values for all pseudo R^2 indices. However, at the lowest levels of multicollinearity, the variability among values produced by the various indices increased. This suggests that, even when optimal conditions exist for predictors in a logistic regression (i.e., situations involve relatively independent predictors), and when typical guidelines for interpretation of OLS R^2 values are used to interpret pseudo R^2 values, widely disparate interpretations may result depending upon which index is chosen.

Although the observed base rate of the binary dependent variable did not substantially affect the values of the indices when the indices were considered collectively, it did affect particular indices. This suggests that a consideration of base rate may be important, particularly when uncorrected pseudo R^2 values are being interpreted.

It is important to consider that there are other conditions that might affect the resultant values of these pseudo R^2 values. Due to the necessity to restrict experimental conditions to a manageable number, the present study, for example, did not vary the sample size, nor did it consider other patterns of multicollinearity, or combinations of predictor variable types (i.e., continuous vs. categorical).

References

Agresti, A., & Finlay, B. (1986). *Statistical methods for the social sciences* (2nd ed.). San Francisco: Dellen/MacMillan.

Aldrich, J. H., & Nelson, F. D. (1984). *Linear probability, logit, and probit models*. Beverly Hills, CA: Sage.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Erlbaum.

Cox, D. R., & Snell, E. J. (1989). *The analysis of binary data* (2nd ed.). London: Chapman and Hall.

Cragg, S. G., & Uhler, R. (1970). The demand for automobiles. *Canadian Journal of Economics, 3*, 386-406.

Efron, B. (1978). Regression and ANOVA with zero-one data: Measures of residual variation. *Journal of the American Statistical Association, 73*, 113-121.

Hagle, T. M., & Mitchell, G. E. (1992). Goodness-of-fit measures for probit and logit. *American Journal of Political Science, 36*, 762-784.

Hosmer, D. W., Jr., & Lemeshow, S. (2000), *Applied logistic regression* (2nd ed.). New York: John Wiley & Sons.

Lave, C. A. (1970). The demand for urban mass transportation. *The Review of Economics and Statistics, 37*, 320-323.

McFadden, D. (1974). Conditional logit analysis of qualitative choice behavior. In P. Zarembka (Ed.), *Frontiers in econometrics* (pp. 104-142). New York: Academic Press.

McKelvey, R. D., & Zavoina, W. (1975). A statistical model for the analysis of ordinal level dependent variables. *Journal of Mathematical Sociology, 4*, 103-12.

McQuitty, L. J. (1960). Hierarchical linkage analysis for the isolation of types. *Educational and Psychological Measurement, 20*, 55-67.

Maddala, G. S. (1983). *Limited dependent and qualitative variables in economics*. New York: Cambridge Press.

Menard, S. (2000). Coefficients of determination for multiple logistic regression analysis. *The American Statistician, 54*(1), 17-24.

Mittlböck, M., & Heinzl, H. (2004). *Proceedings of 1st European workshop on the assessment of diagnostic performance.* Milan, Italy, pp. 71-8.

Nagelkerke, N. (1991). A note on a general definition of the coefficient of determination. *Biometrika, 78*, 691–692.

Veall, M. R., & Zimmermann, K. F. (1994). Evaluating pseudo R^2's for binary probit models. *Quality and Quantity*, *28*, 151-164.

Windmeijer, F. A. G. (1995). Goodness-of-fit measures in binary choice models. *Econometric Reviews*, *14*, 101-116.

Send correspondence to: Thomas J. Smith
Northern Illinois University
Email: tjsmith@niu.edu