Catalogue no. 11-522-XIE

Statistics Canada International Symposium Series - Proceedings

Symposium 2005 : Methodological Challenges for Future Information needs

2005

Statistics Canada Statistique Canada

Canada

Proceedings of Statistics Canada Symposium 2005
Methodological Challenges for Future Information Needs

TWO-STEP REGRESSION WITH LATENT VARIABLES, REVISITED

D. Roland Thomas[1] and Irene R.R. Lu[2]

ABSTRACT

This paper examines the biases associated with the use scale scores as proxies for latent variables in structural models. We consider the situation where scores for latent variables can be predicted using responses to specific questionnaire items, and we focus on measured items where the responses are discrete. In particular, we examine the use of IRT scores, and we identify situations where OLS regression based on particular IRT scores leads to consistent estimates of the parameters of a linear structural equation. When this two-step method is not consistent, we use simulation to quantify its bias and compare it to the bias of two-stage least squares and discrete structural equation modeling techniques. The paper ends with a discussion and some recommendations.

KEY WORDS: Latent variables, IRT scores, latent variable regression, OLS, two-stage least squares, SEM

1. INTRODUCTION

1.1 Latent Variable Regression

This paper describes an investigation of the parameter biases associated with some partial-information methods used by researchers for performing linear regression featuring latent variables, i.e., for estimating the parameters of a single linear structural equation. By latent, we refer to variables or traits that cannot be directly observed (e.g., job satisfaction, students' self-esteem, children's reading ability) but for which a measure or *score* could be inferred, or predicted, using responses to specific questionnaire items, typically referred to as manifest (i.e., measured) items. The emphasis in this paper will be on manifest items that are discrete, either binary or ordinal polytomous items. A variety of structural equation modeling (SEM) techniques that yield consistent parameter estimates are now widely available today in a number of software packages such as LISREL (Jöreskog and Sörbom, 2001) and M*plus* (Muthén and Muthén, 2001), and these now include facilities for dealing with discrete items. Nevertheless, a simpler two-step method in which scores for the latent variables are first predicted and then used in ordinary least squares regression as proxies for the unobservable latent variables is still in common use, as documented by numerous authors. The simpler two-step approach persists despite the fact that the biases inherent in simple two-step regression have been extensively documented over a long period in both the social science literature (see the references in Lu, Thomas and Zumbo, 2005) and in the statistical literature (see, for example, Fuller, 1987). Along with its use by data analysts, research interest in the simple two-step method also continues. This includes investigation of scoring methods for continuous manifest items that avoid estimation bias entirely (Skrondal and Laake, 2001), methods for eliminating the two-step bias (Croon, 2002) and studies of the extent of the bias when first-step scores are obtained via Item Response Theory (IRT) for binary and polytomous item responses (Lu, 2004; Lu et al,, 2005). The primary focus of this paper will be two-step regression with IRT scores. It will be shown that Skrondal and Laake's (2001) results for factor analytic scoring can be extended to specific IRT scoring methods, and when two-step IRT biases cannot be avoided, the two-step IRT score bias will be quantified by means of a simulation, and compared to the bias of two-step mean scores, two-stage least squares and discrete structural equation modeling (SEM) methods.

[1] Sprott School of Business, Carleton University, Ottawa, Canada (rthomas@sprott.carleton.ca)
[2] School of Administrative Studies, York University, Toronto, Canada (irenelu@yorku.ca)

1.2 Technical Specifics

We will consider a single linear latent regression or structural model of the form

$$\eta = \vec{\gamma}'\vec{\xi} + \zeta , \tag{1}$$

where η is a latent dependent variable, $\vec{\xi} = (\xi_1, \ldots, \xi_r)'$ is an r-vector of latent explanatory variables, $\vec{\gamma}$ is an r-vector of regression coefficients and ζ represents a vector of disturbances independent of $\vec{\xi}$. Without essential loss of generality, we will assume that the latent variables have zero mean. In the two-step approach, latent variable scores $\hat{\eta}$ and $\hat{\vec{\xi}}$ that are functions of a set of binary or ordinal polytomous manifest items are used as proxies for latent variables η and $\vec{\xi}$ of Equation (1), and the parameters are then estimated using ordinary least squares (OLS). Note that much of the theory of latent regression modeling (e.g. Croon, 2002) assumes continuous manifest items, but in social science applications continuous measurement is rare. The items associated with each of the latent variables in Equation (1), together with the method by which the items in this set are combined to yield the predicted scores, comprise a *scale* for the latent variable. The scale items for η will be denoted by $\boldsymbol{Y} = (Y_1, \ldots, Y_q)'$, and those for $\vec{\xi}$ will be denoted by $\boldsymbol{X} = (X_1, \ldots, X_p)'$. In this paper, it will be assumed that the scales for individual components ξ_k of $\vec{\xi}$ comprise distinct subsets of the X_j's, denoted $\boldsymbol{X}_k = (X_{k1}, \ldots, X_{kp_k})'$, with $\sum p_k = p$. In other words, the items in each vector $\boldsymbol{X}_k$ define a unidimensional latent variable or trait. Note that when we discuss IRT scoring methods in the next section, it will be convenient to drop the subscript k, and refer to a generic single trait with scale items $X_1, \ldots, X_p$.

2. IRT MODELS AND SCORES

In this section, we briefly describe binary and polytomous IRT models and corresponding methods of latent variable scoring. Of the polytomous IRT models available, we will focus on Samejima's (1969) graded response model (GRM) which is well suited to the ordered categorical data commonly used in attitude, perception, and social / behaviour scales

A key aspect of a binary IRT is the *item characteristic curve* (ICC) which is defined as the probability that an individual will respond positively to a given item. In this study we will consider only the "normal ogive" version of the model, for which the ICC can be expressed as

$$P_j(\xi) = P(X_j = 1 \mid \xi) = \Phi[a_j(\xi - b_j)], \quad j = 1, \ldots, p , \tag{2}$$

where ξ denotes the generic unidimensional latent variable or trait of interest, X_j denotes the (random) response to the j'th manifest item, a_j and b_j are referred to as the discrimination and difficulty parameters for the j'th item, respectively, and Φ denotes the normal *cdf.* The GRM is an extension of this binary two-parameter IRT model, which admits $m+1$ category responses, $x_j = 0, 1, \ldots, m$, for each item j. By first defining a set of m ICC's each with its own difficulty parameter b_{x_j}, the probability of a respondent endorsing category x_j for item j, can be defined for the GRM as

$$P^*_{x_j}(\xi) = P_{x_j}(\xi) - P_{(x+1)_j}(\xi) . \tag{3}$$

When the item characteristics are known or have been estimated, the individual's latent trait score can then be predicted by using the individual's item responses. IRT latent variable prediction methods generally fall into two broad classes, the first of which is based on a binary or multinomial likelihood function defined for each individual using the item response probabilities (2) or (3) and includes maximum likelihood estimation (MLE), and a weighted likelihood approach (WLE) proposed by Warm (1989). The second class of the prediction methods is based on the posterior distribution of ξ, obtained by combining the binary or multinomial likelihood with a prior for ξ, typically normal. In this study we will use a score defined as the expected value of this posterior termed the *expected a posteriori* (EAP) predictor by psychometricians. We will show in the next section that conditional expectation character of the EAP predictor leads to consistent two-step estimates in some special cases.

3. TWO-STEP METHODS

As described in the introduction, the two-step approach obtains latent variable scores $\hat{\eta}$ and $\hat{\vec{\xi}}$ by means of one of the techniques described in the previous section (or some other), and then uses these scores in Equation (1) in place of the unobserved latent variables. The parameter vector is then estimated using ordinary least squares (OLS) regression. The score based version of Equation (1) can then be written as

$$\hat{\eta} = \gamma' \hat{\vec{\xi}} + u \,, \tag{4}$$

where the disturbance term now takes the form

$$u = \zeta + \gamma'(\xi - \hat{\vec{\xi}}) - (\eta - \hat{\eta}) \,. \tag{5}$$

OLS estimation of Equation (4) will yield a consistent estimate of γ only if the disturbance term u is uncorrelated with $\hat{\vec{\xi}}$, i.e.,

$$E\hat{\vec{\xi}}[\zeta + (\xi - \hat{\vec{\xi}})'\vec{\gamma} - (\eta - \hat{\eta})] = 0 \,, \tag{6}$$

because then it is clear from Equation (4) that

$$E(\hat{\vec{\xi}}\hat{\eta}) = E(\hat{\vec{\xi}}\hat{\vec{\xi}}')\vec{\gamma} \,. \tag{7}$$

The second moments in Equation (7) can be consistently estimated using sample realizations of the scores, yielding a consistent estimate of $\vec{\gamma}$.

3.1 Sufficient Conditions for OLS Consistency

Convenient sufficient conditions that ensure OLS consistency can be specified as follows:

(i) The vector of manifest items ***X***, conditional on $\vec{\xi}$, is independent of the equation disturbance ζ .

(ii) The manifest items ***X***, conditional on $\vec{\xi}$, are independent of the manifest items ***Y***, conditional on the univariate latent variable η .

(iii) Elements of manifest items ***Y*** are locally independent given the univariate latent variable η .

(iv) Elements of the manifest items $X_1, \ldots, X_r$ are locally independent given, respectively, the latent variables $\xi_1, \ldots, \xi_r$.

(v) $\hat{\vec{\xi}} = (\hat{\xi}_1, \ldots, \hat{\xi}_r)' = E(\vec{\xi} \mid X_1, \ldots, X_r)$, that is, the vector of r latent predictor scores is defined as the expectation of the latent predictor vector conditional on all r sets of manifest item scales.

(vi) $E_{Y|\eta}\hat{\eta} = \eta$, that is, the scores $\hat{\eta}$, which are functions only of the manifest item vector ***Y***, are conditionally unbiased given the true latent response variable.

Proof of the sufficiency of these conditions is omitted for lack of space. The first four conditions are convenient in that they correspond to the assumptions that underlie standard SEM analyses for both continuous and discrete manifest items. The last two conditions imply specific methods of score prediction. Examples are discussed below.

3.2 OLS Regression of a Measured *Y* on a Single Latent Predictor

Unfortunately, none of the commonly known IRT scoring methods are conditionally unbiased (Kim and Nicewander, 1993). However, IRT scores of EAP type for a single latent predictor with polytomous manifest items satisfy sufficient condition (v). Thus when a single exactly measured variable *Y* (which is conditionally unbiased by default) is regressed on a single IRT-EAP scored latent variable, the resulting OLS estimate of the parameter γ will be consistent. This will be illustrated later in the simulation study.

3.3 OLS Regression of a Measured Y on Multiple Latent Predictors

When $r > 1$, the scores specified in sufficient condition (v) comprise multivariate EAP scores. Thus when a single exactly measured variable Y is regressed on a set of multivariate IRT-EAP scored latent variables, the resulting OLS parameter estimates of the parameter vector $\vec{\gamma}$ will be consistent. However, multivariate IRT-EAP scoring is not the approach that is usually taken in surveys such as the National Longitudinal Survey of Children and Youth (NLSCY; Statistics Canada, n.d.), for example. More typical is to compute what we will refer to as univariate scores, namely

$$\tilde{\xi}_k = E(\xi_k \mid \boldsymbol{X}_{\mathrm{k}}), \quad k = 1, \ldots, r. \tag{8}$$

However, $\tilde{\xi}_k \neq \hat{\xi}_k$, in general, and if the univariate scores are used in a two-step latent estimation process, the OLS parameter estimates will not be consistent even when η is exactly measured, or conditionally unbiased. The exception to this is when the latent explanatory variables are independent.

3.4 OLS Regression with Factor Scores

The above analysis can be considered an extension of Skrondal and Laake's (2001) investigation of two-step methods for latent regression equations featuring multivariate normal latent variables coupled with multivariate normal measurement models of continuous confirmatory factor analysis form. They proved that if the latent response variable was predicted using Bartlett scoring, and if the explanatory variables were predicted using multivariate "regression" scoring, then the latent regression parameters would be consistently estimated. The connection with the analysis in this paper is that Bartlett scoring generates scores that are conditionally unbiased (Croon, 2002), and regression scores for a multivariate normal factor model comprise expected values of the posterior distribution of $\vec{\xi}$ namely $E(\vec{\xi} \mid \boldsymbol{X}_1, \ldots, \boldsymbol{X}_{\mathrm{r}})$ as in sufficient condition (v).

3.5 Bias in the General Case of IRT-Based Latent Regression

The preceding theory suggests that except for the special cases studied, the two-step use of IRT-based latent variable scores will generate parameter bias arising from both the latent response variable and the latent explanatory variables. The magnitudes of these biases will be explored in the simulation.

4. COMPARISON METHODS

4.1 Two-Step Regression with CTT Scores

Croon (2002) explored the bias in two-step regression for scores consisting of linear combinations of continuous scale items. He considered manifest items $\boldsymbol{X}$ and $\boldsymbol{Y}$ satisfying unidimensional measurement models of the form

$$\boldsymbol{Y} = \vec{\lambda}_\eta \eta + \vec{\varepsilon} \qquad \text{and} \qquad \boldsymbol{X}_i = \vec{\lambda}_{\xi_i} \xi + \vec{\delta}_i, \quad i = 1, \ldots, r, \tag{9}$$

where the error vectors $\vec{\varepsilon}$ and the $\vec{\delta}_i$ are independent of one another and of the equation disturbance ζ.
Consider classical test theory (CTT) scores consisting of the means of the items $\boldsymbol{Y}$ and $\boldsymbol{X}_i$. Using Croon's (2002) results, it can be shown, for example, that the regression of a response score on a single predictor score yields a biased estimate of γ, with bias factor $(\overline{\lambda}_\eta / \overline{\lambda}_\xi)\rho_\xi^2$, where ρ_ξ^2 denotes the score reliability. This bias factor differs from the classical attenuation result which involves the reliability only. Results for multiple explanatory variable scores can be obtained using the same approach and Croon's results provide a strategy for removing the bias. However, bias removal is not typically used by practitioners who use CTT scores, so the magnitude of the uncorrected bias generated by these CTT scores will be compared to the biases exhibited by the IRT two-step method when the sufficient conditions are not satisfied.

4.2 Two-Stage Least Squares for Latent Regression

Bollen (1996) adapted the method of two-stage least squares (2SLS) to estimate the parameters of one or more linear latent regression models. He assumed measurement models of the form (9), and the crux of his method was to select a reference item from the items in each scale, and to use this as a proxy in the latent model, i.e., a model of the form (1). As with two-step regression, this results in a composite residual term that is correlated with the proxy predictors. The two-stage method differs from the two-step method in that instrumental variables are selected that are correlated with the proxy predictors but uncorrelated with the residual, and which are used to obtain estimating equations that are free of bias arising from the residual. The instruments are selected from the measurement items not chosen as reference items. The method is operationalized by regressing the proxy predictors on the instruments in the first stage, and regressing the proxy response variable on the fitted predictors in the second stage. This results in consistent parameter estimates.

4.3 Discrete SEM Estimation

SEM estimation for continuous items combines a latent regression model such as Equation (1) with measurement models of the type specified in Equation (9). Discrete SEM estimation is based on an extension of the SEM model in which the item vectors $\boldsymbol{Y}$ and $\boldsymbol{X}_i$, $i = 1, \ldots, r$, are themselves regarded as latent, with their discrete counterparts obtained by categorizing the latent items according to threshold levels estimated from the data. Discrete SEM estimation is now implemented in many SEM computer packages, and the simulation results for this paper were obtained using M*plus*. Discrete SEM estimates are consistent.

5. SIMULATION

5.1 Simulation Design

The model on which the simulation is based is precisely as described above for discrete SEM. Complete details are given by Lu (2004). The essential features of the simulation involved: (a) generation of multivariate normal latent variable realizations consistent with the structural model (1), such that $E(\eta) = E(\xi) = 0$ and $\sigma_\eta^2 = \sigma_\xi^2 = 1$; (b) generation of multivariate normal item responses, y_j and x_j, consistent with both model (1) and the measurement models (9); (c) transformation of y_j and x_j into discrete item scores using threshold parameters designed to generate both symmetric and non-symmetric categories; and (d) for the CTT case, re-standardization of these discrete item scores to have mean zero and variance one, using externally determined variances. The simulation results described below illustrate the theoretical results pertaining to IRT-EAP scoring, and also provide a brief comparison of parameter biases between the two-step, two-stage and discrete SEM approaches. Three bias indicators are displayed which are similar to those used by Skrondal and Laake (2001), namely the percentage relative biases in estimates of the structural regression parameter(s), the coefficient of determination of the structural model, ρ_R^2, and the OLS estimate of the parameter standard error(s). Formally, these are defined as

$$B_{\hat{\gamma}} = 100[\frac{\hat{E}(\hat{\gamma})}{\gamma} - 1], \; B_{R^2} = 100[\frac{\hat{E}(R^2)}{\rho_R^2} - 1], \text{ and } B_{SE} = 100[\frac{\hat{E}(SE)}{\hat{V}^{1/2}(\hat{\gamma})} - 1], \quad (10)$$

where $\hat{E}[\cdot]$ and $\hat{V}[\cdot]$ denote empirical expected values and variances based on 500 simulation replications.

5.2 Exact Response Variable and a Single Latent Explanatory Variable, with EAP Scoring

Table 1 displays simulation results for the two-step regression of an exact response variable on a single latent explanatory variable scored by means of IRT-EAP scores. The squared coefficient of determination both of the latent regression equation (ρ_R^2) and of each individual item measurement equation (ρ_M^2) were set equal to 1/2. The number of categories per item (m+1) and the number of items per scale (n) is indicated in the first two columns of the table. The sample size, N, was set at 300, and all items were generated according to a symmetric category distribution.

From the third column of the table it can be seen that the theoretical predictions are borne out, parameter bias being zero (to the nearest percentage point) for 3 and 5 categories per item, and close to zero for the binary items. The fourth column shows that the R^2 estimate of the coefficient of determination is nevertheless negatively biased. This is consistent with theory, since EAP scoring underestimates the true variance of the latent variable. It can also be seen from the table that the bias in R^2 decreases as the number of items increases (also as the number of categories increases), again consistent with theory which predicts that the measurement error variance in EAP scores goes to zero as $n \to \infty$. Finally, from the last column of Table 1 it can be seen that biases in the standard OLS estimates of parameter standard errors are small.

Table 1

Percent Relative Biases for Exact on a Single Latent: IRT-EAP Scoring

($N = 300$; $\gamma = \sqrt{\rho_R^2} = .707$; $\lambda_j^X = \sqrt{\rho_M^2} = .707$)

$m+1$	n	$B_{\hat{\gamma}}$	B_{R^2}	B_{SE}
2	5	-3	-37	-2
	10	-2	-23	-3
	20	-1	-13	-2
3	5	0	-23	-4
	10	0	-14	-6
	20	0	-8	-7
5	5	0	-20	-1
	10	0	-11	-5
	20	0	-6	-4

Note. *m+1* = number of item categories; *n* = scale length; *N* = sample size.

5.3 Exact Response Variable and Two Latent Explanatory Variables: Univariate and Multivariate IRT-EAP Scoring

Table 2 displays the simulation results for a two-step regression of an exact response variable on two latent predictor variables and contrasts the parameter biases obtained using univariate IRT-EAP scores and multivariate IRT-EAP scores. Parameter settings are the same as those in Table 1, except that all results correspond to 5-items scales, and except for the addition of a correlation between the two latent predictors ξ_1 and ξ_2. The third and fourth columns of Table 2 show bias results for univariate IRT-EAP scoring. As predicted earlier, parameter estimates will be biased in this case unless the correlation between ξ_1 and ξ_2 is zero, and this prediction is clearly borne out. Note that bias results for the two regression parameters are averaged in Table 2. Columns 5 and 6 of Table 2 present preliminary results for multivariate IRT-EAP scores. These were obtained using M*plus* (Muthén and Muthén, 2001), and time did not permit completion of the simulation for the binary case. It can be seen that parameter biases are zero or negligible for 5-category items, irrespective of the correlation, as predicted by theory. As before, R^2 is biased irrespective of the type of EAP score used.

5.4 EAP Two-Step Estimation Versus The Identified Alternatives

Table 3 provides a comparison of parameter and R^2 bias for four methods: (1) two-step regression featuring EAP scores; (2) two-step regression featuring CTT scores; (3) two-stage least squares (2SLS);

Table 2

Percent Relative Biases for Exact on Two Latents: Univariate and Multivariate IRT-EAP Scoring

($N = 300$; $\rho_R^2 = 0.5$; $\gamma_1 = \gamma_2$; $n = 5$; $\lambda_{\xi_1} = \lambda_{\xi_2} = 0.707$)

$m+1$	$\rho(\xi_1, \xi_2)$	Univariate EAP		Multivariate EAP	
		$B(\hat{\gamma})$	$B(R^2)$	$B(\hat{\gamma})$	$B(\hat{\gamma})$

2	0.5	9	-27	--	--
2	0.0	-2	-35	--	--
2	-0.5	-27	-52	--	--
5	0.5	7	-14	0	-13
5	0.0	0	-19	0	-19
5	-0.5	-16	-32	-2	-32

and (4) discrete SEM estimation. The simulation uses two latent explanatory variables, with settings as in Table 2, together with a latent response variable having a measurement model identical to that of the explanatory variables. The IRT-EAP two-step method used here is based on univariate scoring only, as this is what is typically used in practice.

It can be seen from Table 3 that the methods that yield smallest relative biases, both in the parameter estimates and in R^2 are two-stage least squares and discrete SEM. In both cases, the biases are negligible. Both two-step methods on the other hand, yield very similar levels of parameter bias, close to 10% in magnitude with 5-item scales, and about twice that for R^2. The parameter biases drop to a magnitude of about 5% when each scale contains 10 items, a level of bias that some researchers might consider tolerable. In general, it appears from these results that the EAP-based two-step method offers no advantage over the CTT two-step approach unless multivariate EAP scoring is employed.

Table 3

Bias Comparisons for Two-Step, Two-Stage and Discrete SEM Estimation

($N = 300$; $\rho_R^2 = 0.5$; $\gamma_1 = \gamma_2$; $\rho(\xi_1, \xi_2) = 0.5$; $n = 5$; $\lambda_\eta = \lambda_{\xi_1} = \lambda_{\xi_2} = 0.775$)

Number of Items	Method	$B(\hat{\gamma})$	$B(R^2)$
5*Y*'s, 5X_1's, 5X_2's	2-Step (EAP)	-9.8	-19.7
	2-Step (CTT)	-9.7	-18.4
	2SLS	0.2	0.5
	Discrete-SEM	1.2	2.6
10*Y*'s, 10X_1's, 10X_2's	2-Step (EAP)	-5.1	-9.9
	2-Step (CTT)	-4.9	-9.5
	2SLS	0.1	0.2
	Discrete-SEM	0.4	1.0

6. CONCLUSIONS AND RECOMMENDATIONS

The requirement for users of Statistics Canada data to apply appropriate techniques for analyzing latent variable models is increasing as these techniques become better known. The results of the current investigation, focused on perhaps the simplest of all latent variable models, are relevant to this requirement and will be summarized in this section. Some practical recommendations will be provided.

The primary conclusion of this and earlier work is that, in general, two-step regression is not a satisfactory method for estimating the parameters of a latent regression equation. We have shown in this paper that Skrondal and Laake's results can be generalized and that for an *exactly measured* response variable and latent explanatory variables with *multivariate* IRT-EAP scoring, a two-step regression will yield consistent estimators. Unfortunately, even this special case this does not apply to latent variables scored using univariate IRT-EAP scoring, which is the technique typically used in the NLSCY (Statistics Canada, n.d.), for example. Furthermore, some of the results of this study suggest that when the sufficient conditions for consistent two-step estimation do not apply, for example when univariate IRT-EAP scores are used for both response and predictor variables, the two-step biases will be at least as large as those obtained using the two-step method with CTT scores, which are of course much easier to compute. Thus we recommend that if univariate IRT-EAP scores (or CTT scores) are to be used at all, they should be computed using at least ten 5-category items in order to keep biases to a minimum; for smaller numbers of categories per item, more items will be needed. As a final caution regarding the use of IRT-EAP scores, it can also be shown that when one or more exactly measured explanatory variables are included in the latent regression model with an exactly measured response variable, multivariate IRT-EAP scores no longer yield consistent parameter estimates.

The results of the simulation study have, not surprisingly, confirmed the utility of both 2SLS and discrete SEM. For the moderate sample size of 300 used in this study, both these techniques produce estimates having negligible bias, in line with their consistency properties. However, it should be noted that 2SLS theory applies to continuous variables, so that the results of this simulation, which used discrete items, are of considerable interest. An additional advantage of 2SLS is its ease of use. Once reference items have been selected for each factor, and instrumental variables identified, software packages such as SAS can automatically implement the two OLS regressions that are required. This procedure can easily be adapted to the complex survey environment, and for the NLSCY, for example, corresponding bootstrapped parameter standard errors can be obtained using standard procedures. Though discrete-SEM techniques also provide a complete estimation solution for latent variable regression, they are not as "user-friendly" and to this point have not been fully adapted to provide bootstrapped variance estimates for complex survey data, though this is likely to happen in the near future. Neither 2SLS nor SEM techniques use pre-calculated scores, but instead directly use the manifest items identified for each scale. Because of the pre-eminence of these techniques, we therefore recommend that all manifest items be made available on each survey database.

REFERENCES

Bollen, K. A. (1996), "An alternative two stage least squares (2SLS) estimator for latent variable equations", *Psychometrika,* 61(1), 109-121.

Croon, M. (2002), "Using predicted latent scores in general latent structure models", In G. A. Marcoulides & I. Moustaki (Eds.), *Latent variable and latent structure models* (pp. 195-223). Mahwah, NJ: Lawrence Erlbaum Associates.

Fuller, W. A. (1987), *Measurement error models*. New York: John Wiley & Sons.

Jöreskog, K. G. and Sörbom, D. (2001), *LISREL: user's reference guide.* Lincolnwood, IL: Scientific Software International.

Kim, J. K. and Nicewander, W. A. (1993), "Ability estimation for conventional tests", *Psychometrika*, 58(4), 587-599.

Lu, I. R. R. (2004), *Latent variable modeling in business research: A comparison of regression based on IRT and CTT scores with structural equation models*. Doctoral dissertation, Carleton University, Ottawa.

Lu, I. R. R., Thomas, D. R. and Zumbo, B. D. (2005), "Embedding IRT in structural equation models: A comparison with regression based on IRT scores", *Structural Equation Modeling*, 12(2), 263-277.

Muthén, L. K. and Muthén, B. O. (2001), *Mplus user's guide.* Los Angeles, CA: Muthén & Muthén.

Samejima, F. (1969), "Estimation of latent ability using a response pattern of graded scores", *Psychometrika Monograph Supplement,* 34(4), 100-114.

Skrondal, A. and Laake, P. (2001), "Regression among factor scores", *Psychometrika*, 66(4), 563-576.

Statistics Canada. n.d. *Microdata User Guide, National Longitudinal Survey of Children and Youth - Cycle 5, September 2002 to June 2003.* http://www.statcan.ca/english/sdds/document/4450_D4_T9_V1_E.pdf

Warm, T. A. (1989), "Weighted likelihood estimation of ability in item response theory", *Psychometrika*, 54(3), 427-450.