Conditional Design Effects for SEM estimates

DL Oberski

Tilburg University, The Netherlands
doberski@uvt.nl

Abstract

Discusses "conditional design effects" of different survey error components on structural equation model estimates. Estimates effects of clustering, measurement error, and nonnormality for an application involving reciprocal effects between latent variables and correlated error terms.

Keywords: latent variable, clustering, stratification, weighting, measurement error, nonnormality.

1 Introduction

Structural equation modeling (SEM) is a popular framework for formulating, fitting, and testing an variety of models for continuous data in a wide range of fields. Special cases of SEM include factor analysis, (multivariate) linear regression, path analysis, random growth curve and other longitudinal models, errors-in-variables models, and mediation analysis (Bollen, 1989).

This paper discusses a method to evaluate the relative impact of various aspects of the complex sampling design and data on structural equation model parameter estimates. Such "conditional design effects" could serve different purposes:

- Given a pilot study, to judge what to improve so as to attain a given standard error for parameters of interest or given power of a hypothesis test;
- Given candidate weights, to evaluate their effects conditional on other factors.
- To evaluate which "total survey error" components (Groves, 2005) are most relevant to particular multivariate models.

In particular, we consider the effects of stratification, clustering, unequal probabilities of selection, nonnormality, and measurement error.

The remainder of this paper is structured as follows. Section 2 introduces structural equation models (SEM) under complex sampling. It is shown how variance estimators can be defined that omit consideration of each of clustering, stratification, weighting, nonnormality, and measurement error. Conditional design effects based on these variance estimates are then defined in section 3. Section 4 presents an application to a structural equation model for European Social Survey data.

2 SEM under complex sampling

Given a p-vector of observed variables y, let Σ denote its population covariance matrix, and S_n a sample estimator of it. A structural equation model (SEM) is a covariance structure model $\Sigma = \Sigma(\theta)$ expressing the population covariances Σ as a function of a parameter vector θ, an often used parameterization of SEM being the "LISREL all-y" model, consisting of a "structural" model and a "measurement" model

$$\eta = B\eta + \zeta \qquad\qquad y = \Lambda\eta + \epsilon, \tag{1}$$

where η is a vector of latent variables, ζ a vector of latent residuals, and ϵ is a vector of measurement errors. In what follows we will focus on applications in which the "structural" parameters of the model are of primary interest.

Model (1) implies the covariance structure

$$\Sigma(\theta) = \Lambda(I - B)^{-1}\Phi(I - B)^{-1'}\Lambda' + \Psi, \tag{2}$$

where $\Phi := \mathrm{Var}(\zeta)$ and $\Psi := \mathrm{Var}(\epsilon)$. The diagonal of the matrix Ψ contains the measurement error variances, so that the "reliability ratio" (Fuller, 1987) of the j-th observed variable can then be defined as $\kappa_j := 1 - \Psi_{jj}/\Sigma_{jj}$. In other words, without measurement error (when $\kappa_j = 1$ for all j), $\Psi = 0$ and $\Lambda = I$, in which case the model reduces to a path analysis with observed variables.

SEM parameter estimates $\hat{\theta}_n$ are obtained by minimizing a fitting function $F(s_n, \sigma(\theta))$, where $s_n := \mathrm{vech}(S_n)$, $\sigma := \mathrm{vech}(\Sigma)$, and the vech operator denotes half-vectorization. An important matrix is the hessian $V_n := \partial^2 F/\partial\sigma\partial\sigma'$. In the case of WLS estimation, V_n consistently estimates a symmetric estimation weight matrix. The most common choice for F is, however, the normal-theory maximum likelihood (ML) fitting function; in this case (Fuller, 1987, appendix 4.B), $V_{\mathrm{NT}} \overset{a}{=} 2^{-1}D'(\Sigma^{-1} \otimes \Sigma^{-1})D$, where D is the duplication matrix.

Let $\bar{x}$ denote a design-consistent estimator of $E_\pi(x)$. The estimator $\bar{x}$ possibly but not necessarily involves weighting. Define

$$d_{hi} := \sum_{ct} \mathrm{vech}[(y_{hict} - \bar{y})(y_{hict} - \bar{y})'],$$

where y_{hict} is the vector associated with the t-th third-stage unit of the c-th second-stage unit of the i-th PSU of stratum h, with the summation going over all the units within the i-th PSU (Satorra, 1992, 260). This device essentially redefines the observed data matrix to d, simplifying the estimation of the (co)variances s to that of estimating the mean vector $s_n = \bar{d}_n$.

The design-consistent estimator of the (co)variances $\bar{d}$ should be used for s_n in the fitting function. This will then guarantee consistency of the estimator $\hat{\theta}(s_n)$. It can be shown that minimizing F_{ML} with $\bar{d}$ as an estimate of s_n is equivalent to the "pseudo-maximum likelihood" (PML) estimator introduced by Skinner et al. (1989, pp. 80–3).

2.1 Variance estimation

The asymptotic variance of the PML parameter estimates is

$$\mathrm{AVAR}_{\theta,\pi}(\hat{\theta}) = (\Delta' V \Delta)^{-1}\Delta' V \Gamma V \Delta(\Delta' V \Delta)^{-1}, \tag{3}$$

where $\Delta := \partial F/\partial\sigma(\theta)$, and $\Gamma := \mathrm{VAR}_\pi(s_n)$. Equation 3 can be recognized as the "sandwich" estimator of variance.

The redefinition of s_n as the linear estimator $\bar{d}_n$ implies that the usual theory of estimators for means may be applied to obtain a consistent estimate of $\Gamma := \mathrm{VAR}_\pi(s_n) = \mathrm{VAR}_\pi(\bar{d}_n)$. Assuming that this variance is finite, variance estimators can be obtained under various conditions.

Using standard linearization techniques (Wolter, 2007; Skinner et al., 1989, p. 49), a nonparametric estimate $\hat{\Gamma}_{\mathrm{NN,complex}}$ can be constructed that takes into account each of stratification, clustering, and weighting (Muthén and Satorra, 1995). It is equally possible to selectively disregard each of these factors.

The linearization variance estimator will also take nonnormality into account. To allow for the evaluation of the effect of nonnormality, a variance estimator is needed that takes the complex sampling design into account but does assume normality. Oberski (frth) suggested using

$$\hat{\Gamma}_{\text{NT,complex}} := \hat{\Gamma}^{(b)}_{\text{NT}} + n^{-1}\hat{\Gamma}^{(w)}_{\text{NT}}, \tag{4}$$

where $\hat{\Gamma}^{(b)}_{\text{NT}}$ and $\hat{\Gamma}^{(w)}_{\text{NT}}$ denote the between-cluster and within-cluster normal-theory covariance matrices respectively, and the normal-theory variance estimator of a variance matrix S_n is $\hat{\Gamma}_{\text{NT}} = 2D^{+}(S_n \otimes S_n)D^{+'}$, with D^{+} the Moore-Penrose inverse of D.

Finally, the effect of measurement error on standard errors requires evaluation while still taking complex sampling and nonnormality into account. From the model, it is clear that if there had been no measurement error, the observed sample covariance would have estimated $\Sigma - \Psi$. Let

$$y^* := y' S_n^{-\frac{1}{2}}(S_n - \hat{\Psi}_n)^{\frac{1}{2}}. \tag{5}$$

Then $\text{VAR}(y^*) = (S_n - \hat{\Psi}_n)$ (Mair et al., 2012). Therefore, analyzing the transformed data y^* as though these had been directly observed will provide consistent point estimates, but standard errors under the (incorrect) assumption that the factor scores were directly observed without measurement error (Skrondal and Laake, 2001). We use this fact to evaluate the "design" effect of measurement error.

3 Conditional design effects

We define the "conditional design effect" of a design factor as the variance estimate taking account of all factors, divided by the variance estimate taking account of all factors except one. Skinner (1986, pp. 89-91) and Skinner et al. (1989) remark that this quantity is not the same as the well-known Kish (1965) design effect and is more accurately labeled conditional "misspecification effect", since it quantifies the effect of not taking account of one design factor. In practice, sample estimators of design effects and misspecification effects tend to correspond, justifying the term "design effect". We follow this argument, while stressing the difference with Kish's design effect.

3.1 Special cases: design effect of weighting or clustering alone

When considering only the influence of weighting, the "design effect" for means is often taken to equal $\text{deff}_{\text{Kish}} := 1 + \text{cv}^2(w)$, where $\text{cv}(w)$ is the coefficient of variation of weights (Kish, 1965). Stapleton (2006) suggested using this adjustment in SEM, but found in simulations that it did not accurately reproduce the true relative variance increase. However, there is a special case in which d^2 can be expected to be reasonable measure of the design effect: namely, when the model is scale-free and unstandardized regression coefficients and loadings are of interest. In the context of dealing with nonnormality, Satorra and Bentler (1994, p. 410-1) noted that for such parameters it is the case that $\text{AVAR}(\hat{\theta}_n) = c \cdot (\Delta' V \Delta)^{-1}$, where c is the Satorra-Bentler "scaling correction factor" (assuming *df* > 0). Furthermore, c is proportional to $\text{VAR}(\Delta' \bar{d})$ (p. 409); and applying standard arguments for design effects of means to $\Delta' \bar{d}$ the design effect $\text{VAR}_{\text{NN,complex},(jj)}/\text{VAR}_{\text{NN},(jj)}$ should be close to $\text{deff}_{\text{Kish}}$ for such parameters.

Skinner (1986, p. 95) showed that, when considering only the effect of clustering on multivariate parameters, the design effect can be derived as

$$\text{deff} = 1 + (m^* - 1)\rho_g, \tag{6}$$

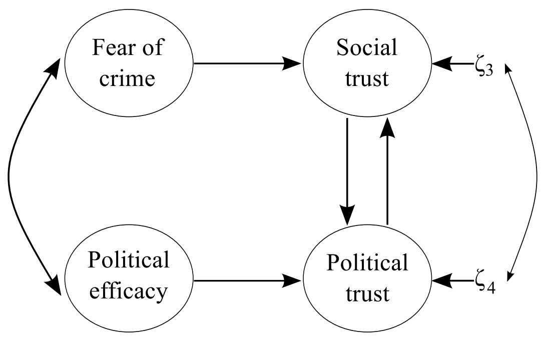

Figure 1: Structural model for the Danish ESS data.

where $m^* := n^{-1} \sum_i m_i^2$ and m_i is the size of the i-th cluster. Here, ρ_g is the intracluster correlation of the "individual parameter contributions" to the g-th parameter, IPC $:= (\partial\hat{\theta}/\partial s_n)d$. For structural equation models, it can be shown that (Oberski, frth) IPC $= (\Delta' V \Delta)^{-1} \Delta' V d$. Thus, after calculating the IPC's, the design effect of two-stage sampling for the g-th parameter can be obtained by calculating the intracluster correlation coefficient for the g-th column of IPC and applying Equation 6. Oberski (2013) implements the IPC for structural equation models in the open source SEM library `lavaan` (Rosseel et al., 2013).

4 Application to ESS 2008 data

Saris and Gallhofer (2007) analyzed a structural equation model of social and political trust. Figure 1 shows a simplified version. The model in Figure 1 cannot be estimated with ordinary linear regression because it contains a reciprocal effect between social and political trust, which is identified by excluding the regression coefficient from "fear of crime" to "political trust" and from "political efficacy" to "social trust".

The variables shown in the model are not observed variables, but latent variables defined as influencing the answers to certain survey questions. For each of these constructs we can obtain at least two measures from the European Social Survey[1] (ESS) round 4, conducted in 2008t. As an illustration we select only data from Denmark. where respondents were selected with equal inclusion probabilities, which simplifies the discussion needed below. Composite scores $\hat{\eta}$ were calculated by summing each latent variables' indicators. The reliability coefficients were 0.73, 0.77, 0.57, and 0.64 for "social trust, "political trust", "fear of crime" and "efficacy", respectively.

Fitting the model in Figure 1 to the Danish ESS data yields the structural parameter estimates shown in Table 1. Main interest focuses on the regression coefficients, although other parameters may also be of interest, for instance when calculating standardized coefficients. The main finding, discussed by Saris and Gallhofer (2007), is that social trust appears to influence political trust but not the other way around.

Using the methods described above, four standard error estimates were computed, taking account of:

- σ_1: Interviewer clustering, nonnormality, and measurement error;
- σ_2: Nonnormality and measurement error;
- σ_3: Clustering and measurement error;
- σ_4: Clustering and nonnormality.

[1] See http://ess.nsd.uib.no/ess/round4/.

In all cases, the point estimator remained the same. Table 1 shows the square-root design effect (deft) of clustering, conditional on nonnormality and measurement error defined as σ_1/σ_2, and similarly for the other factors.

Table 1 shows that the relative standard error increase due to clustering and nonnormality are very close to each other and around 20%. The similarity of these increases is partially due to clustering and nonnormality effects' strong interactions (Oberski, frth). The relative variance increase due to measurement error is much larger, nearing 95%.

	Est.	σ_1	σ_2	σ_3	σ_4	$\text{deft}_{\text{clus}}$	deft_{nn}	$\text{deft}_{\text{measerr}}$
pol $\rightarrow$ soc	0.303	0.261	0.212	0.210	0.142	1.230	1.242	1.835
soc $\rightarrow$ pol	0.771	0.170	0.166	0.160	0.086	1.025	1.068	1.983
fear $\rightarrow$ soc	-0.678	0.271	0.247	0.238	0.142	1.100	1.140	1.908
eff $\rightarrow$ pol	0.510	0.232	0.175	0.168	0.117	1.322	1.383	1.980

Table 1: Point and standard error estimates from fitting the model under different conditions. "deft" columns show relative standard error increase due to each factor, σ_1/σ_k.

In this application, correction for measurement error therefore had the strongest impact on standard errors. This may lead one to wonder whether the mean square error will be lower without that correction. Table 2 shows that it is not. Calculating bias from the uncorrected regression coefficients and plugging in the obtained standard error, the table shows that the $\sqrt{\text{MSE}}$ is still higher than the standard errors of unbiased estimates (from Table 1). Under the model assumptions, in spite of the considerable variance increase due to correction for measurement error, corrected estimates provide the lowest MSE.

	$\hat{\theta}_n$	$\hat{\text{b}}^2$	$\hat{\sigma}^2$	$\text{MSE}^{\frac{1}{2}}$	$\hat{\sigma}_1$
pol $\rightarrow$ soc	0.44	0.02	0.04	0.26	0.26
soc $\rightarrow$ pol	0.83	0.01	0.02	0.17	0.17
fea $\rightarrow$ soc	-0.31	0.14	0.02	0.39	0.27
eff $\rightarrow$ pol	0.28	0.06	0.02	0.27	0.23

Table 2: Square-root mean square error (MSE) without correction for measurement error.

5 Summary

Conditional design effects can be useful measures of the effect of clustering, stratification, unequal sampling probabilities, and measurement error on structural equation model parameters variances. The idea is to evaluate what would happen to standard errors if a particular "design" factor were not considered. This article discussed an implementation for SEM and methods of overcoming certain difficulties in estimating conditional design effects.

An application to a SEM analysis of Danish data from the 2008 European Social Survey identified measurement error as the most influential. In spite of the large increase in parameter uncertainty due to correction for measurement error, however, the mean square error of corrected parameter estimates still outperformed that of the uncorrected estimates, a result that emphasizes the importance of corrections for measurement error.

Acknowledgements This work was supported by The Netherlands Organization for Scientific Research (NWO) [vici grant number 453-10-002].

References

Bollen, K. (1989). *Structural Equations with Latent Variables*. John Wiley & Sons, New York.

Fuller, W. (1987). *Measurement Error Models*. John Wiley & Sons, New York.

Groves, R. M. (2005). *Survey Errors and Survey Costs*. Wiley-Interscience.

Kish, L. (1965). *Survey Sampling*. John Wiley & Sons, New York.

Mair, P., Satorra, A., and Bentler, P. M. (2012). Generating nonnormal multivariate data using copulas: Applications to SEM. *Multivariate Behavioral Research*, 47(4):547–565.

Muthén, B. and Satorra, A. (1995). Complex sample data in structural equation modeling. *Sociological methodology*, 25:267–316.

Oberski, D. (2013). *Individual Differences in Structural Equation Model Parameters*. ArXiv. `http://arxiv.org/pdf/1304.33`.

Oberski, D. (frth). *Estimating the effect of nonnormality and clustering on covariance structure model parameter variances*. Retrieved from `http://daob.nl/wp-content/uploads/2013/03/Oberski-Normality.pdf`.

Rosseel, Y., Oberski, D., Byrnes, J., Vanbrabant, L., and Savalei, V. (2013). lavaan: Latent variable analysis. [Software]. Available from `http://lavaan.ugent.be/`.

Saris, W. and Gallhofer, I. N. (2007). *Design, evaluation, and analysis of questionnaires for survey research*. Wiley-Interscience, New York.

Satorra, A. (1992). Asymptotic robust inferences in the analysis of mean and covariance structures. *Sociological Methodology*, 22:249–278.

Satorra, A. and Bentler, P. (1994). Corrections to test statistics and standard errors in covariance structure analysis. In von Eye, A. and Clogg, C. C., editors, *Latent Variables analysis: applications to developmental research*. Sage, Thousand Oakes, CA.

Skinner, C. (1986). Design effects of two-stage sampling. *Journal of the Royal Statistical Society. Series B (Methodological)*, pages 89–99.

Skinner, C., Holt, D., and Smith, T. (1989). *Analysis of Complex Surveys*. John Wiley & Sons, New York.

Skrondal, A. and Laake, P. (2001). Regression among factor scores. *Psychometrika*, 66(4):563–575.

Stapleton, L. M. (2006). An assessment of practical solutions for structural equation modeling with complex sample data. *Structural Equation Modeling*, 13(1):28–58.

Wolter, K. (2007). *Introduction to Variance Estimation, Second Edition*. Springer-Verlag, New York.