A Flexible Method to Explain Differences in Structural Equation Model Parameters over Subgroups

DL Oberski

Department of Methodology and Statistics, School of Social and Behavioral Sciences, Tilburg University, The Netherlands.

Abstract

Structural Equation Modeling (SEM) is a widely applied technique to model relationships including both observed and latent variables. Recently, the explanation of subgroup differences in SEM parameters has attracted the attention of applied social researchers. Examples are differences in survey measurement error variance and differences in the variance of, possibly only indirectly observed, wages or educational achievement to explain inequality. However, these fields have been hindered by the lack of methods to regress a SEM parameter such as a latent variable's variance on a set of predictor variables. I present a method, "IPC regression", addressing this problem. IPC regression involves three analysis steps: 1) running a pooled SEM, 2) calculating so-called "individual parameter contributions" (IPC's), and 3) regressing these IPC's on the predictors of scientific interest. IPC regression is presented in a general framework that encompasses any type of SEM parameter and prediction model. Large- and small-sample properties are considered, and an example application illustrates the implications for practical research.

All data and analysis code used are available from `http://goo.gl/fstgW`.

Key words: Latent variables, heterogeneity, heteroskedasticity, differential measurement error, variance function regression.

Structural equation modeling (Bollen, 1989) is a widely applied technique in social and psychological research to model relationships among observed and latent variables (MacCallum and Austin, 2000; Bollen, 2002). One drawback of classical structural equation modeling (SEM) is that parameter values are assumed equal over all individuals. Modern extensions to SEM allow for parameter differences over subgroups but not for a way of modeling and explaining such differences. I therefore introduce a method, "individual parameter contribution (IPC) regression", to model subgroup differences in any parameter of a SEM model as a function of explanatory variables.

Explaining differences in SEM parameters is a common but sometimes difficult task. For example, survey methodologists seeking to deal with measurement error in questionnaires have tried to explain differences in survey question reliability and measurement error variance over respondents (e.g., Alwin and Krosnick, 1991; Scherpenzeel and Saris, 1997; Alwin, 2007; Révilla, 2012). Unfortunately, as noted by Alwin (2007, p. 219), "it is not possible to predict reliability of reporting for an individual". Since SEM programs do not allow one to regress, for instance, an error variance parameter value on age, gender and education simultaneously, these studies have been limited to looking at the effect of one explanatory variable at a time using multiple-group structural equation modeling (MGSEM), a point made explictly by Révilla (2012, p. 52). In addition, MGSEM requires a continuous variable such as age to be split up into arbitrary categories. The main disadvantage of this one-variable-at-a-time approach is that one cannot disentangle the effects of, for instance, age and education; two variables that are thought to impact measurement error variance but which tend to correlate. Questions of scientific interest, such as whether the effect of age on measurement error is mediated by cognitive ability (Alwin, 2007, p. 218) have remained unanswered. The study of mechanisms impacting survey measurement error therefore appears to have been inhibited by a technical difficulty. This difficulty can be resolved by IPC regression as a general method to model differences in SEM parameters, including measurement error variances, something I demonstrate by example in this article.

Western and Bloome (2009) also focused on explaining variance, arguing that within-group variance is an important determinant of societal inequality. They suggested modeling the variance of a dependent variable with a so-called “variance function regression”. While useful, their method assumed directly observed and perfectly measured dependent variables, and does not allow the dependent or independent variable to be latent. Potentially, unobserved variables could be incorporated using Structural Equation Models, but there is currently no method to perform variance function regression for SEM with latent variables. Recent applications of variance function regression in the sociological literature have predicted the variance in wages (Western and Rosenfeld, 2011), educational achievement (Montt, 2011), and self-reported health (Zheng, Yang and Land, 2011), all variables that have been shown to contain measurement error (Bound and Krueger, 1991; Crossley and Kennedy, 2002; Schröder and Ganzeboom, 2009). This clear need to study inequality in imperfectly measured variables implies that a method of predicting the variance of latent dependent variables is an important step forwards. I show that IPC regression is a general framework that can be used to address this problem.

The IPC regression method proposed here proceeds in three steps. In the first step a classical structural equation model is fit. The second step is to obtain individual parameter contributions (IPC’s) from the first-step solution. As shown below, this can be done by inverting a Structural Equation Model’s mapping from parameters to individual observations. The third step is then simply to regress the IPC’s obtained in the second step on predictors of interest.

Individual parameter contributions are calculated by a transformation of the casewise gradient of the likelihood with respect to the parameters. Analysis of casewise gradients has a long history in econometrics (Brown, Durbin and Evans, 1975; Nyblom, 1989; Andrews, 1993; Kuan and Hornik, 1995; Hjort and Koning, 2002; Zeileis, 2005; Zeileis and Hornik, 2007). So-called "structural change tests", obtained by cumulating casewise gradients over cases, are well-established for the detection of change points in time series analysis, and have more recently been proposed by Merkle and Zeileis (2013) as an exploration of measurement invariance in factor analysis. Structural change tests allow the researcher to explore subgroups that may have differing parameter values by performing a series of hypothesis tests. IPC regression is different in that it directly models parameter differences as a function of the predictors. The advantage of structural change tests is that they provide well-established hypothesis tests to explore unmodeled heterogeneity. The advantage of IPC regression is that, through regression models, it allows the researcher to explain such heterogeneity. The two methods are therefore related but not identical and serve mutually complementary purposes.

IPC regression can be seen as a generalization of Modification Indices (MI's, or "Lagrange Multiplier tests") and Expected Parameter Changes (EPC's) in multiple-group SEM with equality restrictions (Saris, Satorra and Sörbom, 1987). By viewing the grouping variable in MGSEM as the "predictor variable" and differences between groups in a parameter value as the "dependent variable", the Appendix shows that, asymptotically, EPC's for freeing cross-group equalities correspond to IPC dummy regression coefficients while MI's correspond to the Wald tests on these coefficients. Of course, this correspondence ends for situations that cannot be handled by MGSEM, such as predictors that are continuous, multiple predictors, and mediation models for parameter differences – as discussed above, all situations that occur in practical research. In this sense, IPC regression is a generalization of the MI and EPC in MGSEM to situations that cannot be handled by MGSEM.

The contributions of this article are fourfold. First, it introduces IPC regression as a flexible method of explaining differences between SEM parameters over subgroups, resolving the problems experienced in the survey methodology and within-group social inequality literatures. Second, it provides a basic evaluation of the performance of this method in a small Monte Carlo simulation. Third, it demonstrates the practical implications of IPC regression for applied research by an example survey methodology application. Finally, example code for standard SEM software is provided in an online appendix that can be used by applied researchers wishing to implement the method.

1. INDIVIDUAL PARAMETER CONTRIBUTIONS TO SEM ESTIMATES

In a SEM model, covariances are considered functions of parameters, while covariances are themselves functions of individual observations. This section shows how inverting that relationship yields individual parameter contributions (IPC's), and how these IPC's can be used to investigate the relationship between parameter estimates and predictors.

1.1. Structural Equation Modeling (SEM) Framework

Given a vector $\mathbf{y}$ of p observed variables with population covariance matrix $\mathbf{\Sigma}$, a structural equation model can be viewed as a covariance structure model $\mathbf{\Sigma} = \mathbf{\Sigma}(\boldsymbol{\theta})$, where $\mathbf{\Sigma}(\boldsymbol{\theta})$ is a continuously differentiable matrix-valued (symmetric and positive definite) function of the vector of parameters $\boldsymbol{\theta}$ of the model (see, e.g. Bollen, 1989, for more details).

Given an observed covariance matrix $\mathbf{S}$, based on a sample of size n of $\mathbf{y}$, the vector of parameters $\boldsymbol{\theta}$ is estimated by minimizing with respect to $\boldsymbol{\theta}$ a discrepancy function $F(\boldsymbol{\theta}) = F(\mathbf{S}, \mathbf{\Sigma}(\boldsymbol{\theta}))$ of $\mathbf{S}$ and $\mathbf{\Sigma}(\boldsymbol{\theta})$. This approach encompasses maximum likelihood as well as weighted least squares estimators.

A key matrix is the hessian matrix $\mathbf{V} = \frac{1}{2}\frac{\partial^2 F(\mathbf{S},\mathbf{\Sigma})}{\partial\boldsymbol{\sigma}\partial\boldsymbol{\sigma}'}$, where $\boldsymbol{\sigma} := \mathrm{vech}(\mathbf{\Sigma})$ is the half vectorization of $\mathbf{\Sigma}$. In weighted least squares estimators, $\mathbf{V}$ is a weight matrix determined by the choice of estimator (Satorra, 1989). Sample estimates of $\mathbf{V}$ are generally obtained as a function of either, or both of, the observed or the implied covariance matrices.

To study how parameter values vary as a function of observations, we also define n vectors of size $p(p+1)/2$

$$\mathbf{d}_i := \text{vech}\left[(\mathbf{y}_i - \bar{\mathbf{y}})(\mathbf{y}_i - \bar{\mathbf{y}})'\right], \tag{1}$$

(e.g. Satorra, 1992), so that the average of the individual moment contributions, $\mathbf{d}$, will yield the observed covariances $\bar{\mathbf{d}} = \mathbf{s}$.

1.2. Individual Parameter Contributions (IPC's)

Whether maximum-likelihood or some other estimator is used, SEM parameter estimates $\hat{\boldsymbol{\theta}}$ can be seen as the solution to the estimating equation $\partial F/\partial\boldsymbol{\theta} = 0$, that is, $f(\mathbf{s}, \boldsymbol{\theta}) = f(\bar{\mathbf{d}}, \boldsymbol{\theta}) = 0$, where f is defined implictly. To see how the data vector $\mathbf{d}$ is transformed into parameter estimates $\hat{\boldsymbol{\theta}}$ by SEM analysis, following Bentler and Dijkstra (1984) we observe that parameter estimates can be approximated as a linear function of the data,

$$\hat{\boldsymbol{\theta}} \approx \left(\frac{\partial\boldsymbol{\theta}}{\partial\bar{\mathbf{d}}}\right)\bar{\mathbf{d}} = \mathbf{W}\bar{\mathbf{d}}. \tag{2}$$

Since $\boldsymbol{\theta}$ and $\bar{\mathbf{d}}$ are linked implictly by the estimating equation $f(\bar{\mathbf{d}}, \boldsymbol{\theta}) = 0$, the derivative $\partial\boldsymbol{\theta}/\partial\bar{\mathbf{d}}$ can be obtained by applying the implicit function theorem:

$$\mathbf{W} = \frac{\partial\boldsymbol{\theta}}{\partial\bar{\mathbf{d}}} = -\left(\frac{\partial f}{\partial\boldsymbol{\theta}'}\right)^{-1}\left(\frac{\partial f}{\partial\bar{\mathbf{d}}'}\right) = (\boldsymbol{\Delta}'\mathbf{V}\boldsymbol{\Delta})^{-1}\boldsymbol{\Delta}'\mathbf{V}, \tag{3}$$

where $\boldsymbol{\Delta} = \partial\boldsymbol{\sigma}/\partial\boldsymbol{\theta}$, and the second step follows from the results given by Neudecker and Satorra (1991), who also give $\boldsymbol{\Delta}$ as a function of the model parameter estimates.

Now, since $\hat{\boldsymbol{\theta}} = \mathbf{W}\bar{\mathbf{d}}$, it follows that if we define a transformed data vector,

$$\mathbf{t} := \mathbf{W}\mathbf{d}, \tag{4}$$

then, under the assumption that $\mathbf{W}$ is constant, the mean of this transformed data vector $\mathbf{t}$ will equal the parameter estimates, $\bar{\mathbf{t}} = \mathbf{W}\bar{\mathbf{d}} = \hat{\boldsymbol{\theta}}$. Note that this property will also hold for subgroups. Thus, the transformed data vector $\mathbf{t}$ can be seen as a vector of "individual parameter contributions" (IPC's). Consistent sample estimates of the IPC's, $\hat{\mathbf{t}}$, can be obtained by replacing $\mathbf{W}$ by its sample estimate $\hat{\mathbf{W}}$.

Although not usually described explicitly, IPC's are a common feature in the literature on "robust" standard errors and linearization variance estimation of maximum-likelihood estimates under complex sampling (e.g. Skinner, 1986). Calculating the variance matrix of the observed IPC's $\hat{\mathbf{t}}$ yields the robust covariance matrix, $\hat{\mathrm{var}}(\hat{\boldsymbol{\theta}}) = n^{-1}(\hat{\mathbf{t}} - \bar{\hat{\mathbf{t}}})'(\hat{\mathbf{t}} - \bar{\hat{\mathbf{t}}}) = \mathbf{W}\hat{\mathrm{var}}(\mathbf{d})\mathbf{W}$, which equals Equation [16.10] of Satorra and Bentler (1994). In other words, calculating standard deviations of the $\mathbf{t}$'s will provide exactly the "robust" standard errors. Although here I use the IPC's for a different purpose, this property guarantees that inference about the IPC's is robust to nonnormality.

1.3. Predicting SEM Parameter Estimates

Consider the following hypothetical model predicting a SEM parameter from a vector of predictors $\mathbf{z}$,

$$\boldsymbol{\theta}_{ij} = \boldsymbol{\gamma}\mathbf{z}_i + \boldsymbol{\delta}_i. \tag{5}$$

If the individual parameter values for the j-th parameter, $\boldsymbol{\theta}_{ij}$, were observed, the usual OLS estimator of $\boldsymbol{\gamma}$ would be $(\mathbf{z}\mathbf{z}')^{-1}\mathbf{z}\boldsymbol{\Theta}_j$, where $\boldsymbol{\Theta}_j := (\boldsymbol{\theta}_{1j}, ..., \boldsymbol{\theta}_{nj})'$. However, although individual parameters are, of course, unobserved, since $\hat{\mathbf{t}}_j$ is a consistent estimate of $\boldsymbol{\Theta}_j$, a consistent sample estimate of $\hat{\gamma}$ is the OLS regression of the j-th IPC on the predictors,

$$\hat{\gamma} = (\mathbf{z}\mathbf{z}')^{-1}\mathbf{z}\hat{\mathbf{t}}. \tag{6}$$

Note that, in fact, although we will not consider this issue further, other estimators than for the linear regression model 5 can in principle be obtained in a similar way. For example, choosing the GLM model with Gamma distribution and a log link will yield Western and Bloome (2009)'s variance function regression for SEM with latent variables.

Consider the special case where a multiple-group structural equation model (MGSEM) could be formulated with cross-group equality constraints on all parameters including the parameter of interest. A Modification Index (MI) and Expected Parameter Change (EPC) (Saris, Satorra and Sörbom, 1987) could then be calculated for a cross-group equality constraint on a parameter of interest. Satorra (1989, section 5) discussed "generalized" MI's and EPC's, that are robust to nonnormality. In this simple case, an alternative formulation of the problem is in terms of IPC regression: in the first step, a one-group SEM is estimated, in the second step the IPC's for the parameter of interest calculated, and in the third step, these IPC's are regressed on $\mathbf{z}$, a set of dummy-coded grouping variables. It turns out that the regression coefficients obtained by this IPC regression are asymptotically equal to the generalized EPC's and the Wald tests on the IPC regression coefficients equivalent to the generalized MI. The proof of this interesting result is given in Appendix A. Note that IPC regression is more general than the MGSEM approach, since $\mathbf{z}$ need not represent a single grouping variable, and the relationship between $\mathbf{z}$ and the IPC's can be modeled directly.

Considering the close link between IPC regression and MI's and EPC's, it is not surprising that their primary assumptions should be the same. This assumption is that the second derivative matrix $\mathbf{J} := \mathbf{\Delta}'\mathbf{V}\mathbf{\Delta}$ is approximately constant between the null and the true model (Saris, Satorra and Sörbom, 1987). For IPC regression this implies that

$E_{\mathbf{z}}[\mathbf{J}(\hat{\boldsymbol{\theta}}|\mathbf{z})] \approx \mathbf{J}$. This assumption will not hold in general, since the average conditional parameter estimate is not generally equal to the parameter estimate for a completely pooled model. However, it will hold approximately for parameters depending on the covariance matrix such as latent regression coefficients and variances when the sample covariance matrix is conditioned on $\mathbf{z}$. For this reason, it is recommended, when studying covariance structure parameters, to include the predictors in the first-step SEM analysis with direct effects on all observed variables.

1.4. IPC Regression

In summary, possible differences in structural equation model parameter estimates can be investigated by modeling a transformed data set $\hat{\mathbf{t}}$ as a function of $\mathbf{z}$, for instance by OLS linear regression. This suggests a three-step approach to exploring differences in any set of parameters of a structural equation model with respect to a set of predictors:

1. Estimate the overall single-group model (introducing $\mathbf{z}$ as fixed covariates);
2. From the data and step 1, obtain the transformed data set $\hat{\mathbf{t}}$;
3. Regress the individual parameter contributions $\hat{\mathbf{t}}$ for the parameter(s) of interest on the covariates.

2. MONTE CARLO EVALUATION OF IPC REGRESSION

2.1. Simulation Setup

To evaluate the finite-sample performance of the IPC regression proposed here, we set up a small simulation study. The goal of this study is to demonstrate the results of the previous section in a simple setup, while evaluating the extent to which IPC regression is able to recover true between-group differences and provides valid inferences.

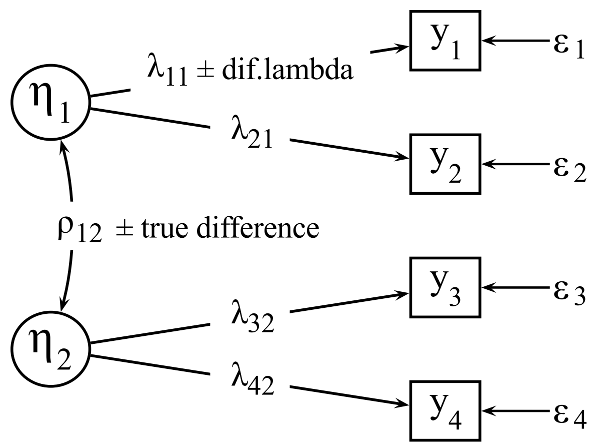

Figure 1: Simulation model. Of interest is the true between-group difference in the factor correlation, ρ_{12}. Depending on the condition, there is also a difference of "dif.lambda" in the size of the first loading. Between-group differences in the intercepts of $y_1 - y_4$ are always present (not shown for clarity).

Figure 1 shows the two-factor model (1 degree of freedom) used for our Monte Carlo simulations. A two-group population is generated, with all loadings equal to 1 in group 1, and $\lambda_{21} = \lambda_{42} = 1, \lambda_{32} = 1/2$, and $\lambda_{11} = 1 +$ *dif.lambda* in group 2. All error variances in both groups equal $\mathrm{var}(\epsilon_j) = 0.8$. The intercepts of all observed variables equal 1 in group 1 and 2 in group 2. The latent variables η_1 and η_2 are standardized in both groups and their correlation is set to 0.5 in group 1 and to $0.5 +$ *true difference* in group 2. We assume that interest focuses on this true difference in latent correlation between groups. IPC regression in this simple example will therefore correspond to a simple dummy variable regression.

The simulation conditions fully cross the following experimental factors:

- The true difference in between-group correlation: {-0.4, -0.2, -0.1, 0, +0.1, 0.2, 0.4}. Note that the zero condition corresponds to no effect;
- The difference in loadings between groups: {0, 0.1, 0.2};
- The sample size per group: n_g = {125, 250, 500, 1000}. The total sample size therefore equaled $n = 2n_g$ = {250, 500, 1000, 2000}.

For each of the $4 \times 7 \times 3 = 84$ resulting conditions, n observations were drawn from a multivariate normal distribution with mean vector and covariance matrix depending on the group. IPC regresssion was then applied to the sample data, and this process was replicated 1000 times for each condition.

2.2. Simulation Results

Figure 2 plots the true difference in latent correlations in each condition against the difference estimated using the differences in individual parameter contributions. For easy reference, the black 45-degree diagonal line corresponds to exact equality of the true and estimated difference. Each point corresponds to a particular simulation condition, connected with lines. The lines lie very close to the ideal of exact equality. When the true difference is strongly positive or negative, it is slightly underestimated in absolute terms. That is, the difference estimate is slightly biased towards zero for very large between group differences in the latent correlation. The differences between the various conditions obtained from crossing sample size with loading differences are hardly discernible. This implies that the IPC difference procedure provides close to unbiased estimates under all of these conditions.

Besides bias in the recovery of the true difference, coverage of 95$\%$ confidence intervals on the IPC regression coefficient is also of interest. Figure 3 shows that coverage is close to the nominal rate under all conditions (dotted reference line). The standard errors obtained by simply regressing the IPC on the group dummy variable therefore provide accurate confidence intervals under these conditions.

Inference may also involve hypothesis tests on the IPC regression coefficients. Table 1 shows the empirical Type-I error across replications, i.e. the proportion of rejections in the conditions with a true difference of zero. Ideally this proportion should approach $\alpha = 0.05$. The Table shows that there is adequate control of Type-I error under the null hypothesis.

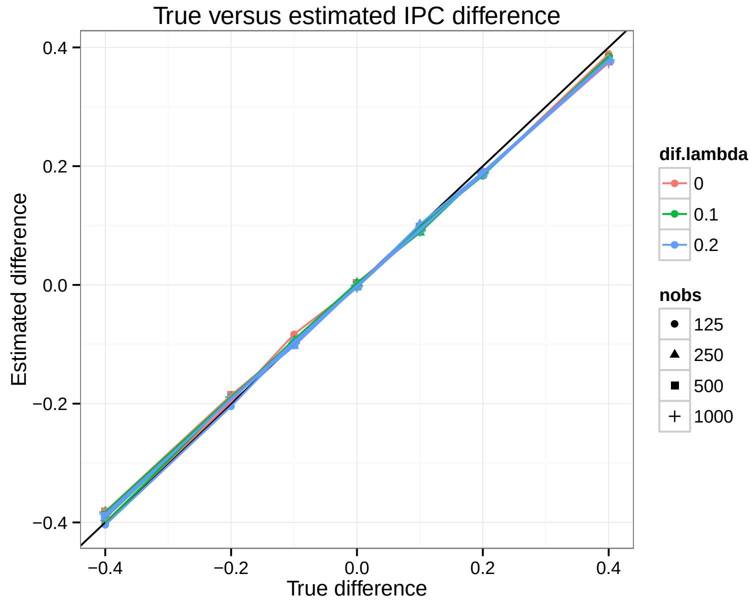

Figure 2: Bias in the mean difference between groups of IPC for the latent correlation. Different conditions are shown as differences in color (loading non-invariance) and point shape (number of sample observations).

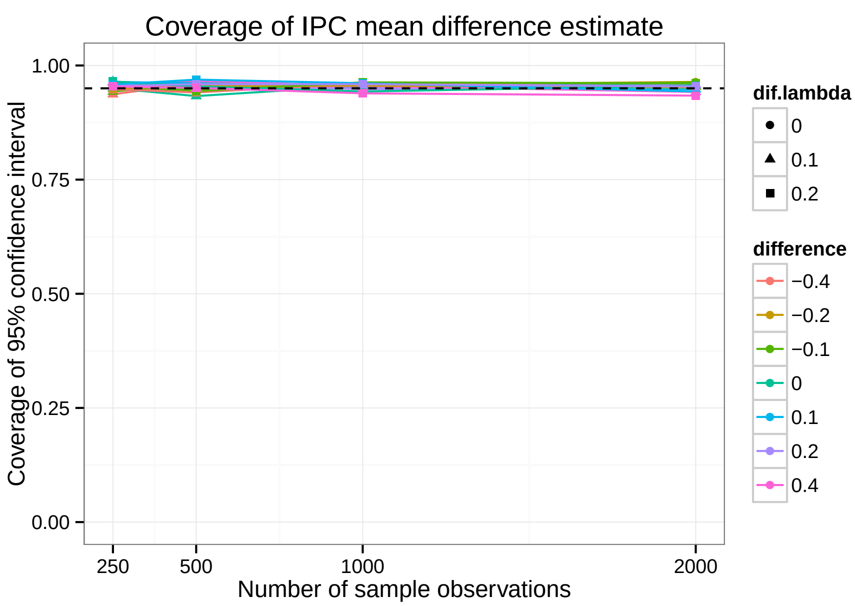

Figure 3: Coverage of 95% confidence intervals of the mean difference between groups of latent correlation's IPC's. Different conditions are shown as differences in color (true latent correlation difference) and point shape (loading difference).

Table 1: Empirical Type-I error rates for $\alpha = 0.05$.

	IPC		
Diff. loading	0	0.1	0.2
No. obs.			
125	0.048	0.050	0.035
250	0.050	0.067	0.046
500	0.049	0.044	0.057
1000	0.048	0.043	0.044

This holds true for the smaller sample sizes as well as when other parameters do differ strongly over the groups.

When the true difference is not zero, the rejection rate is the empirical power of the test to detect a difference in latent correlation. Figure 4 relates the power of the test to sample size and the true size of the difference. For comparison with an established method, the left-hand graph in Figure 4 corresponds to the power using MGSEM, and the right-hand graph to the power using the IPC difference test. The overall median power using multiple group structural equation modeling is 0.0095 higher than that using IPC difference testing. For conditions with a small and medium absolute differences (bottom lines), MGSEM is somewhat more powerful, but with large differences (top lines), IPC difference testing was more powerful for smaller sample sizes.

Overall, in this simple setup the IPC regression method performed very well, demonstrating the results of the previous section. IPC regression was able to recover true between-group differences, and provided accurate confidence intervals and hypothesis tests. There does not appear to be any loss of power relative to multiple-group SEM, which is an alternative method in this particular setup. In general, however, it will not always be possible to compare the two methods, since IPC regression with multiple, possibly continuous, predictors do not have a corresponding MGSEM formulation.

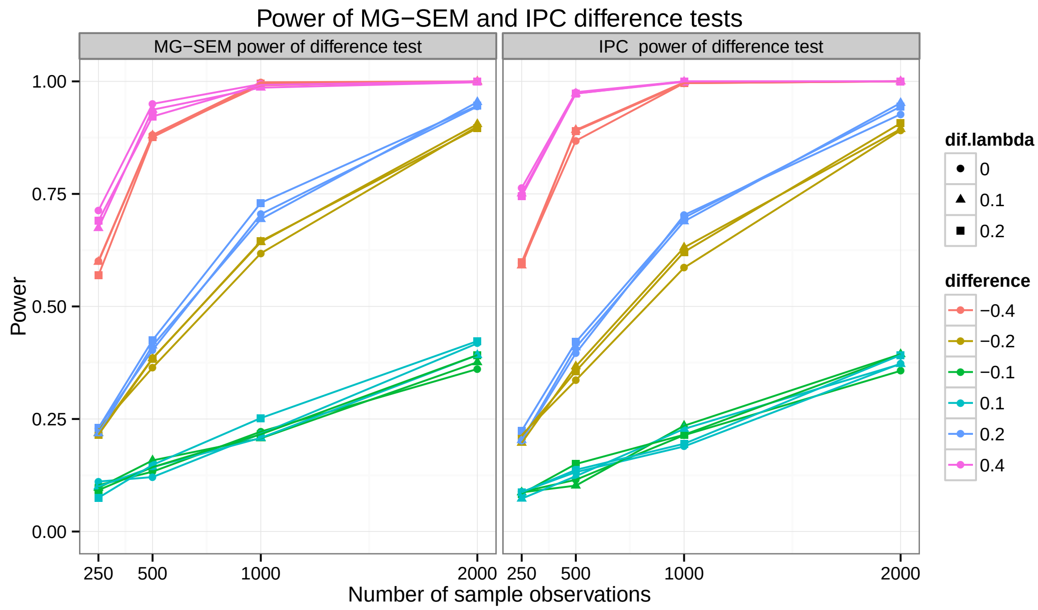

Figure 4: Power to reject mean difference test for latent correlation. Different conditions are shown as differences in color (true latent correlation difference between groups) and point shape (loading non-invariance).

3. EXAMPLE APPLICATION: PREDICTING SURVEY MEASUREMENT ERROR VARIANCE FROM RESPONDENT CHARACTERISTICS

Self-reported internet usage is often studied in relation to other variables (e.g. Hoffman and Novak, 1998; Gross, 2004). As is well-known, when studying the relationship between internet use and other variables, random measurement error may bias the results downwards. Worse, when random measurement error variance differs over groups of sex, age, and education, true differences may not only be obscured, but spurious differences may be erroneously found (Carroll et al., 2006). Therefore a question of high practical relevance is whether measurement error variance in this survey question differs over social groups. Survey methodologists pondering improvements of internet use measurement will, in addition, be interested in the mechanism behind such differences. Here I demonstrate how IPC regression can be used to answer these questions.

Table 2: Descriptive statistics for four years' observations of log(hours of internet use at work per week + 1); $n = 2838$. Covariates not shown.

	Correlations							
	2008	2009	2010	2011	Mean	sd	Skew	Kurtosis
inet_2008	1				0.721	1.029	1.298	0.600
inet_2009	0.668	1			0.743	1.048	1.283	0.543
inet_2010	0.644	0.701	1		0.766	1.073	1.229	0.315
inet_2011	0.609	0.661	0.729	1	0.790	1.094	1.181	0.189

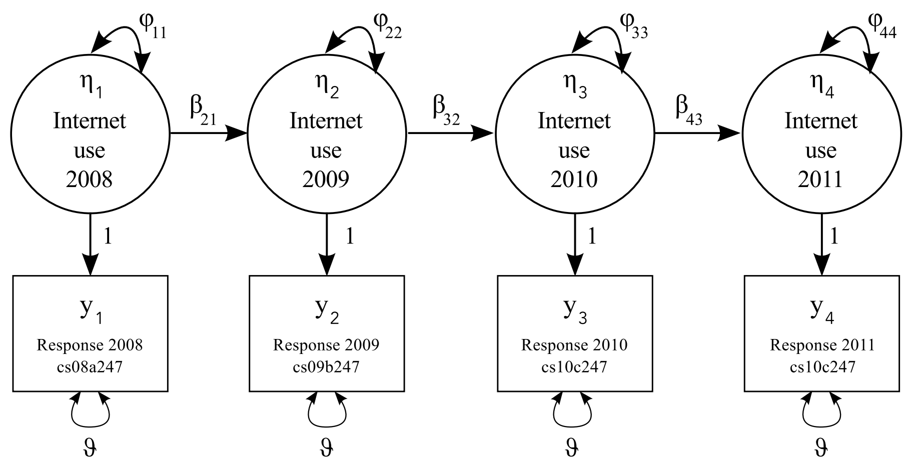

Figure 5: The quasi-simplex model for four consecutive measurement error-prone measurements of internet use in the LISS panel.

All data and code used to produce this example is available online from `http://goo.gl/fstgW`.

I use 2838 observations from the LISS panel study, a simple random sample of Dutch households. For more information on the design of the LISS study, see Scherpenzeel (2011), and see `http://lissdata.nl/` for original data and questionnaires. In this longitudinal panel study design, respondents were asked how often they used the internet in four consecutive years (2008 through 2011). The log-transform of the number of hours per week the respondent claims to use the internet at work is taken to reduce skewness; Table 2 shows descriptive statistics for these measures.

The first step is to formulate a structural equation model estimating measurement error

Table 3: Results of fitting the quasi-simplex model with covariates to the LISS data. Intercepts and fixed covariate main effects not shown. Note: Satorra-Bentler chi-square is 1.5 with 2 degrees of freedom ($p = 0.476$, scaling factor = 2.3). CFI = 1.000, RMSEA = 0.016, $n = 2838$.

		Latent regressions			Variances			
	var(Error)	β_{21}	β_{32}	β_{43}	F1	F2	F3	F4
Est.	0.28	0.91	0.94	0.94	0.58	0.13	0.11	0.08
s.e.	(0.01)	(0.03)	(0.02)	(0.02)	(0.02)	(0.02)	(0.01)	(0.02)
Stand.	0.26	0.89	0.90	0.92	0.74	0.16	0.12	0.09

variance in the answers to the internet use question. Table 2 shows that the consecutive measurements do not all correlate equally: measurements at a greater distance from each other in time correlate less, so that the size of the correlations tapers off towards the lower-left corner of the matrix. This is in line with the so-called "quasi-simplex" model, in which internet use is not only measured with error, but the true internet use may also change over time (Wiley and Wiley, 1970; Alwin and Krosnick, 1991; Alwin, 2007). Figure 5 shows the single-group quasi-simplex model formulated for these data as a structural equation model.

Fitting this model to the LISS data with the maximum likelihood estimator yields a well-fitting model, with estimates shown in Table 3. The model includes the direct effect of self-employment, age, age squared, sex, and education level on each of the four observed variables. These regression coefficient estimates are omitted from Table 3 for clarity. The error variance parameter is estimated at 0.28 ± 0.01. This error variance estimate corresponds to reliability estimates of respectively 0.86, 0.86, 0.87, and 0.88 at the four consecutive years. The overall reliability is therefore high, but not perfect: estimates of the relationship between internet use and social background will be affected.

Does the measurement error variance differ over respondents? Having completed step one of fitting a single-group structural equation model, we now obtain each individual's parameter contribution to the error variance parameter estimate (step two): a $\hat{\mathbf{t}}$ score for each respondent. Step three is then to perform a multiple linear regression of that score on

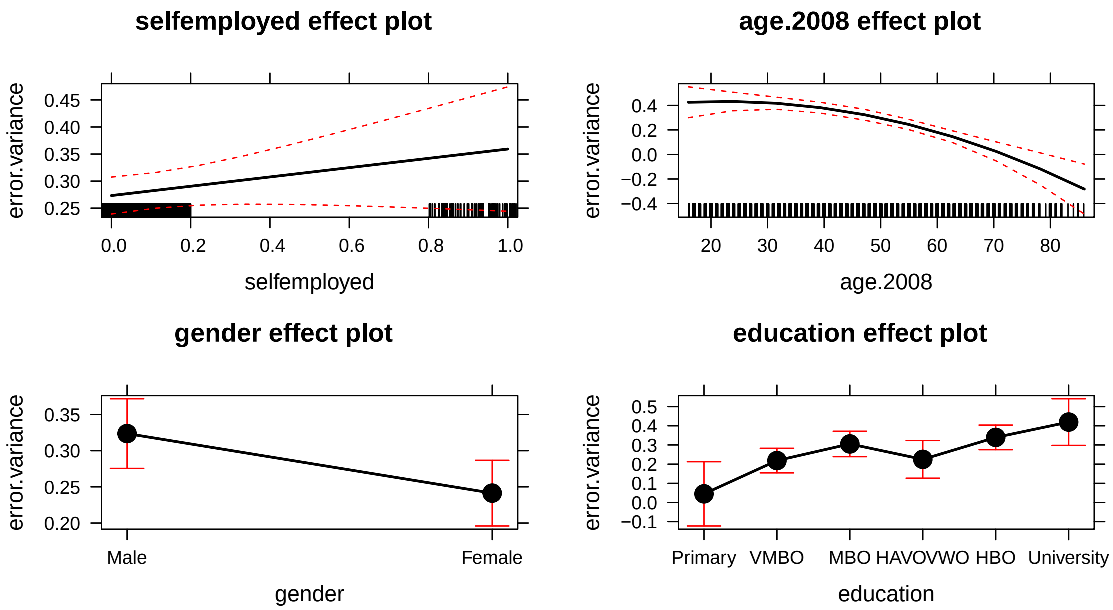

Figure 6: Estimated conditional differences in measurement error variance by respondent characteristics: effect plot for the regression of the IPC of the error variance parameter on self-employed, year of birth, gender, and education.

respondent characteristics using standard routines. We employ the `lm` function in `R`, but linear regression routines such as those in Excel, SPSS, or Stata would serve equally well. The characteristics investigated here are whether the respondent is self-employed or not (0/1), the year of birth and its square, sex of the respondent, and education level – primary, lower secondary (VMBO), middle secondary (MBO), upper secondary (HAVO/VWO), lower tertiary (HBO), and higher tertiary (University).

The multiple linear regression estimates of the relationship between respondent characteristics and error variance are shown as effect plots in Figure 6. The estimates themselves can be found in the appendix. Figure 6 shows that, controlling for the other characteristics, three of the four respondent characteristics (age, sex, and education) have a strong and statistically significant effect on the error variance, and therefore on reliability, which is inversely related to the error variance. The reliability of internet use self-reports *increases* with age, *decreases* with education, and is lower for men than for women.

To survey methodologists and those familiar with the literature on reliability of survey questions, it may seem surprising that reliability would be lower for the young and those of a higher education level. After all, cognitive functioning is higher for these groups and expected to increase the reliability of answers (Alwin, 2007, chapter 10). There may, however, be a simple explanation for this finding: the variance of measurement errors in self-reports of the number of hours spent using the internet at work may simply be related to internet use itself. If one does not use the internet *at all*, presumably there would be little to no measurement error. A possible explanation for the finding that young and educated respondents appear to provide lower-quality answers is that these respondents simply tend not to be the kind of respondents who use the internet at work for zero hours per week.

The effects on error variance of these respondent characteristics may thus be *mediated* by whether the respondent *never* uses the internet. A dummy variable is therefore constructed indicating whether respondents answered "zero hours" on all four occasions. I then formulated a new mediation SEM model in which the error variance IPC is influenced by the "all-zeroes" variable, and this variable, in turn, by the four covariates. This full mediation model fits the data well ($\chi^2_{\text{SB}} = 11$, df = 9, $p = 0.276$, RMSEA = 0.009; full results are found in the Appendix). Thus, it appears that measurement error variance indeed differs over these respondent characteristics, but only because respondents with certain traits tend to be those who never use the internet, strongly reducing their error variance: the difference is 0.50 ± 0.03, corresponding to a fully standardized effect of "all-zeroes" on the error variance of 0.3.

4. DISCUSSION

This article suggested using IPC regression to explain differences in SEM parameter values over subgroups. IPC regression is a three-step method that opens up the possibility of explaining differences in measurement error variances with multiple explanatory variables, and of performing a variance function-like regression for latent variables. While these are important applications with direct benefits for applied researchers, I presented IPC regression in a more general framework that can be used to explain differences in other types of SEM parameters as well.

IPC regression yields consistent estimates of the effect of predictors on parameters, and has connections with established methods such as structural change tests and EPC's and MI's. These results were demonstrated in a small Monte Carlo simulation. This simulation study indicated that true parameter differences are accurately recovered by IPC regression, although a slight bias towards zero was observed for extreme differences. Confidence intervals and hypothesis tests appeared satisfactory, and no loss of power relative to multiple-group SEM was observed. A comparison between IPC regression and MGSEM could only be made due to the simple two-group, single-variable setup, however; IPC regressions with multiple, possibly continuous, predictors do not have a corresponding MGSEM formulation. An example application regressed the measurement error variance of self-reported internet use on respondent characteristics, demonstrating the practical use of the method for applied researchers. Data and program code are provided in the online Appendix.

IPC regression is, of course, by no means the only existing method to study heterogeneity in SEM parameters. There is overlap in certain fields of application with other methods: when there is a single discrete predictor, MGSEM could also be applied; when there is a single discrete predictor whose effect can be assumed random, multilevel SEM is an option (Muthén, 1989; Yuan and Bentler, 2007); when the parameter of interest is a (latent) regression coefficient, SEM with interactions is available (Kenny and Judd, 1984; Bollen, 1995; Klein and Moosbrugger, 2000; Mooijaart and Bentler, 2010); when exploration of possible subpopulations is of interest, structural change tests (Merkle and Zeileis, 2013) or finite mixture (latent class) SEM (Dolan and van der Maas, 1998; Yuan and Bentler, 2010) may be appropriate; and when interest does not focus on one parameter, but on finding subgroups that are maximally different on all parameters, recursive SEM partitioning may be an alternative (Zeileis, Hothorn and Hornik, 2008; Aluja, Lamberti and Sanchez, 2013).

A clear advantage of IPC regression relative to methods that require multidimensional numerical integration or recursive algorithms is its computational simplicity, although with ever more powerful computers this advantage may lose relevance in the future. However, the main attraction of the IPC framework presented here is that can in principle deal with all of these situations. It can, furthermore, deal with situations that are not dealt with by any of these methods, such as the example application discussed.

However, when one of these alternative methods is available and its assumptions are met, IPC regression, being in essence a limited-information approach, is likely to provide asymptotically less efficient estimates and less powerful tests. In the simulation study these differences relative to MGSEM were found to be negligible, and with small samples IPC regression was even found to be more powerful. Nevertheless, how IPC regression compares to the above other methods in the special cases where these are appropriate remains an open question. Two other important limitations of the present discussion are that only SEM for continuous data was discussed, and that missing data remain to be dealt with.

References

Aluja, Tomas, Giuseppe Lamberti and Gaston Sanchez. 2013. Modelling with heterogeneity. In *Advances in Latent Variables*, ed. Brentari E. and Carpita M. Milan, Italy: Vita e Pensiero.

Alwin, D.F. 2007. *Margins of error: a study of reliability in survey measurement*. New York: Wiley-Interscience.

Alwin, D.F. and J.A. Krosnick. 1991. "The Reliability of Survey Attitude Measurement: The Influence of Question and Respondent Attributes." *Sociological Methods & Research* 20:139.

Andrews, Donald WK. 1993. "Tests for parameter instability and structural change with unknown change point." *Econometrica: Journal of the Econometric Society*.

Bentler, P.M. and T. Dijkstra. 1984. "Efficient estimation via linearization in structural models." In *Multivariate Analysis VI*, ed. P.R. Krishnaiah. Amsterdam: North-Holland pp. 9–42.

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

Bollen, K.A. 1995. "Structural equation models that are nonlinear in latent variables: A least-squares estimator." *Sociological Methodology* 25:223–252.

Bollen, Kenneth A. 2002. "Latent variables in psychology and the social sciences." *Annual review of psychology* 53:605–634.

Boos, D.D. 1992. "On generalized score tests." *The American Statistician* 46:327–333.

Bound, John and Alan B Krueger. 1991. “The Extent of Measurement Error in Longitudinal Earnings Data: Do Two Wrongs Make a Right?” *Journal of Labor Economics*.

Brown, Robert L, James Durbin and John M Evans. 1975. “Techniques for testing the constancy of regression relationships over time.” *Journal of the Royal Statistical Society. Series B (Methodological)*.

Carroll, R.J., D. Ruppert, L.A. Stefanski and C.M. Crainiceanu. 2006. *Measurement error in nonlinear models: a modern perspective*. Vol. 105 Boca Raton, FL: Chapman & Hall/CRC Monographs on Statistics & Applied Probability.

Crossley, Thomas F and Steven Kennedy. 2002. “The reliability of self-assessed health status.” *Journal of health economics* 21:643–658.

Dolan, Conor V and Han LJ van der Maas. 1998. “Fitting multivariate normal finite mixtures subject to structural equation modeling.” *Psychometrika* 63:227–253.

Gross, Elisheva F. 2004. “Adolescent Internet use: What we expect, what teens report.” *Journal of Applied Developmental Psychology* 25:633–649.

Hjort, Nils Lid and Alexander Koning. 2002. “Tests for constancy of model parameters over time.” *Journal of Nonparametric Statistics* 14:113–132.

Hoffman, D.L. and T.P. Novak. 1998. “Bridging the Racial Divide on the Internet.” *Science* 280:390–391.

Kenny, David A and Charles M Judd. 1984. “Estimating the nonlinear and interactive effects of latent variables.” *Psychological Bulletin* 96:201.

Klein, Andreas and Helfried Moosbrugger. 2000. “Maximum likelihood estimation of latent interaction effects with the LMS method.” *Psychometrika* 65:457–474.

Kuan, Chung-Ming and Kurt Hornik. 1995. "The generalized fluctuation test: A unifying view." *Econometric Reviews* 14:135–161.

MacCallum, Robert C and James T Austin. 2000. "Applications of structural equation modeling in psychological research." *Annual review of psychology* 51:201–226.

Merkle, Edgar C and Achim Zeileis. 2013. "Tests of Measurement Invariance Without Subgroups: A Generalization of Classical Methods." *Psychometrika* 78:59–82.

Montt, Guillermo. 2011. "Cross-national differences in educational achievement inequality." *Sociology of Education* 84:49–68.

Mooijaart, Ab and Peter M Bentler. 2010. "An alternative approach for nonlinear latent variable models." *Structural Equation Modeling* 17:357–373.

Muthén, Bengt. 1989. "Latent variable modeling in heterogeneous populations." *Psychometrika* 54:557–585.

Muthén, Bengt and A. Satorra. 1995. "Complex Sample Data in Structural Equation Modeling." *Sociological Methodology* 25:267–316.

Neudecker, H. and A. Satorra. 1991. "Linear Structural Relations: Gradient and Hessian of the Fitting Function." *Statistics and Probability Letters* 11:57–61.

Nyblom, Jukka. 1989. "Testing for the constancy of parameters over time." *Journal of the American Statistical Association* 84:223–230.

Révilla, Mélanie. 2012. "Impact of the Mode of Data Collection on the Quality of Answers to Survey Questions Depending on Respondent Characteristics." *Bulletin de Méthodologie Sociologique* 116:44–60.

Saris, W.E., A. Satorra and D. Sörbom. 1987. "The Detection and Correction of Specification Errors in Structural Equation Models." *Sociological Methodology* 17:105–129.

Satorra, A. 1989. "Alternative Test Criteria in Covariance Structure Analysis: A Unified Approach." *Psychometrika* 54:131–151.

Satorra, A. 1992. "Asymptotic Robust Inferences in the Analysis of Mean and Covariance Structures." *Sociological Methodology* 22:249–278.

Satorra, A. and P.M. Bentler. 1994. "Corrections to Test Statistics and Standard Errors in Covariance Structure Analysis." In *Latent Variables analysis: applications to developmental research*, ed. A. von Eye and Clifford C. Clogg. Thousand Oakes, CA: Sage.

Scherpenzeel, A.C. 2011. "Data Collection in a Probability-Based Internet Panel: How the LISS Panel Was Built and How It Can Be Used." *Bulletin of Sociological Methodology/Bulletin de Méthodologie Sociologique* 109:56–61.

Scherpenzeel, A.C. and W.E. Saris. 1997. "The Validity and Reliability of Survey Questions: A Meta-Analysis of MTMM Studies." *Sociological Methods & Research* 25:341.

Schröder, Heike and Harry BG Ganzeboom. 2009. "Measuring and Modelling Education Levels in European Societies." *Paper presented at European Consortium for Sociological Research*.

Skinner, CJ. 1986. "Design effects of two-stage sampling." *Journal of the Royal Statistical Society. Series B (Methodological)*.

Sörbom, D. 1989. "Model modification." *Psychometrika* 54:371–384.

Western, Bruce and Deirdre Bloome. 2009. "Variance function regressions for studying inequality." *Sociological Methodology* 39:293–326.

Western, Bruce and Jake Rosenfeld. 2011. "Unions, norms, and the rise in US wage inequality." *American Sociological Review* 76:513–537.

Wiley, D.E. and James A. Wiley. 1970. "The Estimation of Measurement Error in Panel Data." *American Sociological Review* 35:112–117.

Yuan, Ke-Hai and Peter M Bentler. 2007. "Multilevel covariance structure analysis by fitting multiple single-level models." *Sociological methodology* 37:53–82.

Yuan, Ke-Hai and Peter M Bentler. 2010. "Finite Normal Mixture SEM Analysis by Fitting Multiple Conventional SEM Models." *Sociological methodology* 40:191–245.

Zeileis, Achim. 2005. "A unified approach to structural change tests based on ML scores, F statistics, and OLS residuals." *Econometric Reviews* 24:445–466.

Zeileis, Achim and Kurt Hornik. 2007. "Generalized M-fluctuation tests for parameter instability." *Statistica Neerlandica* 61:488–508.

Zeileis, Achim, Torsten Hothorn and Kurt Hornik. 2008. "Model-based recursive partitioning." *Journal of Computational and Graphical Statistics* 17:492–514.

Zheng, Hui, Yang Yang and Kenneth C Land. 2011. "Variance Function Regression in Hierarchical Age-Period-Cohort Models Applications to the Study of Self-Reported Health." *American sociological review* 76:955–983.

A. IPC REGRESSION AS A GENERALIZATION OF MI'S AND EPC'S IN MGSEM

The expected parameter change for freeing an equality-constrained parameter θ_g in group g is defined as (Saris, Satorra and Sörbom, 1987):

$$\text{EPC}_g = \delta_g := \left(\frac{\partial^2 F}{\partial \boldsymbol{\theta}_g \partial \boldsymbol{\theta}'_g}\right)^{-1} \left(\frac{\partial F}{\theta_g}\right) \tag{7}$$

Note that the hessian with respect to the entire parameter vector is used, not just with respect to the parameter of interest. This takes account (to some degree) of expected changes in other, correlated, parameters due to freeing the constraint (Sörbom, 1989). The EPC's can be used to obtain an estimate of the difference between groups from multiple group SEM as the difference of the EPC's δ_g for each group (Bentler and Dijkstra, 1984): $(\hat{\theta} + \delta_2) - (\hat{\theta} + \delta_1) = \delta_2 - \delta_1$.

Theorem 1. *Formulate an MG-SEM with equality restrictions on all parameters. Order groups by size of θ_g. Define $\hat{\mathbf{t}}$ as in (6). Then: $\delta_2 - \delta_1 = \bar{\hat{\mathbf{t}}}_2 - \bar{\hat{\mathbf{t}}}_1$.*

Proof. The estimated difference between groups from MGSEM is: $J_2^{-1}(\partial F/\partial\theta_2) - J_1^{-1}(\partial F/\partial\theta_1) = J^{-1}[(\partial F/\partial\theta_2) - (\partial F/\partial\theta_1)]$ because $J_1 = J_2$ since all parameters estimates are equal and θ_1 and θ_2 play the same role in different groups: $\Delta_{\theta_1} = \Delta_{\theta_1}$ and $V_1 = V_2$ so $J_1 = J_2 = J$.

Now, $\partial F/\partial\theta_g = \Delta' V[s_g - \sigma(\theta_g)]$ (e.g. Neudecker and Satorra, 1991). So $(\partial F/\partial\theta_2) - (\partial F/\partial\theta_1) = \Delta' V[s_2 - s_1]$, again because $\sigma(\hat{\theta}_2) = \sigma(\hat{\theta}_1)$. Since by definition $s_g = \bar{d}_g$, $\Delta' V[s_2 - s_1] = \Delta' V[\bar{d}_2 - \bar{d}_1]$, leading to $\delta_2 - \delta_1 = J^{-1}\Delta' V[\bar{d}_2 - \bar{d}_1] = \bar{\hat{\mathbf{t}}}_2 - \bar{\hat{\mathbf{t}}}_1$. □

The "modification index" (score test) for freeing this restriction is defined as the difference divided by the variance of the difference. Some standard software packages currently use the inverse hessian J^{-1} as the variance. However, from Satorra (1989, section 5) it is clear that this is only valid when asymptotically optimal (AO) estimation has been used.

For non-AO estimation, such as pseudo-maximum likelihood or in the case of nonnormally distributed $\mathbf{y}$, a sandwich estimator is needed (Satorra, 1989). The resulting hypothesis test is then equivalent to the so-called "generalized" score test (Boos, 1992). From $\bar{\hat{\mathbf{t}}}_2 - \bar{\hat{\mathbf{t}}}_1 = J^{-1}\Delta' V[\bar{d}_2 - \bar{d}_1]$, it is clear that a standard z-test performed on the difference in IPC's fulfills this requirement since it will automatically calculate the sandwich variance $\mathrm{var}(\delta_2 - \delta_1) = J^{-1}\Delta' V \mathrm{var}(\bar{d}_2 - \bar{d}_1) V \Delta J^{-1}$. Thus, squared z-statistics from the regression of $\hat{\mathbf{t}}$ on a group indicator can be interpreted as generalized score tests that are robust to nonnormality.

Complex sampling can be taken into account as described in the text: by adjusting d using the weights and estimating $\mathrm{var}(\bar{d}_2 - \bar{d}_1)$ incorporating clustering and stratification identifiers (Muthén and Satorra, 1995). In general MG-SEM can be expected to yield the most accurate results. However, since some SEM software does not correct the modification indices for nonnormality or complex sampling, in cases of strong nonnormality or complex sampling design effects, the proposed procedure may therefore actually be preferable to MG-SEM.

B. INTERNET USE STUDY EXTENDED RESULTS

Extended results for the application of IPC regression to the internet use example are given in this section.

Table 4: Coefficient estimates for regression of the IPC of the error variance parameter on respondent characteristics ($n = 2838$). The R^2 is 0.03.

	Estimate	s.e.	t-value	p-value
(Intercept)	0.08	(0.09)	0.93	0.35
Self-employed	0.09	(0.06)	1.40	0.16
Year birth	6.71	(0.91)	7.34	0.00
Year birth^2	-2.33	(0.91)	-2.55	0.01
Female	-0.08	(0.03)	-2.42	0.02
Education				
Primary school	0.00			
VMBO	0.17	(0.09)	1.90	0.06
MBO	0.26	(0.09)	2.82	0.00
HAVO/VWO	0.18	(0.10)	1.83	0.07
HBO	0.29	(0.09)	3.21	0.00
University	0.37	(0.11)	3.54	0.00

Table 5: Full mediation SEM: error variance IPC (dependent variable), "all-zeroes" indicator (mediator), and covariates (predictors). Satorra-Bentler chi-square is 11 with 9 degrees of freedom ($p = 0.276$, scaling correction = 0.953); CFI = 0.999, RMSEA = 0.009, $n = 2838$.

Parameter	Est.	s.e.	z
Regression coefficients			
Error variance IPC on			
All-zeroes	-0.50	0.03	-17.0
All-zeroes on			
Self-employed	-0.06	0.03	-2.4
Year birth	-10.07	0.40	-25.1
Year birth^2	6.46	0.41	16.0
Female	0.10	0.02	6.2
Education: VMBO	-0.13	0.04	-3.1
Education: MBO	-0.31	0.04	-7.4
Education: HAVO/VWO	-0.35	0.04	-7.8
Education: HBO	-0.46	0.04	-11.3
Education: University	-0.51	0.04	-11.4
Variance parameters			
Error variance IPC	0.75	0.06	11.8
All-zeroes	0.17	0.00	50.1