

# Using Mplus individual residual plots for diagnostics and model evaluation in SEM

*Tihomir Asparouhov and Bengt Muthén*

Mplus Web Notes: No. 20

October 31, 2017

## **1 Introduction**

A variety of plots are available in Mplus 7.2 that can be used to visually evaluate the fit of the model. We can inspect the fit on a more detailed level where individual observations are plotted. Such level of detail usually can not be obtained when standard SEM fit statistics are used such as chi-square statistics and CFI/TLI indices. These standard SEM fit statistics are usually based on a global model fit, i.e., a fit for the entire population as a whole rather than fit for individual observations. In this paper we provide several examples where the plots can be used to quickly discover patterns in the data that are not accounted for in the model.

Consider for example a SEM model that contains the following relationship between the dependent variables  $Y$ , latent variables  $\eta$  and covariates  $X$

$$Y = \nu + \lambda\eta + \beta X + \varepsilon$$

$$\eta = \alpha + \gamma X + \zeta$$

where  $\nu$ ,  $\alpha$ ,  $\lambda$ ,  $\gamma$ , and  $\beta$  are model parameters where  $\nu$  and  $\alpha$  are vectors and  $\lambda$ ,  $\gamma$  and  $\beta$  are matrices. Mplus can be used to compute the factor scores  $\hat{\eta} = E(\eta|Y, X)$  and the model estimated/predicted values for  $Y$ , which we denote by  $\hat{Y}$ . Mplus computes two versions for the predicted values of  $Y$ . The first one uses both the factor score and the covariates, i.e.,

$$\hat{Y} = \hat{\alpha} + \hat{\lambda}\hat{\eta} + \hat{\beta}X.$$

The second method for evaluating the predicted  $Y$  values is

$$\hat{Y}' = E(\widehat{Y}|X).$$

The second version of predicted value is commonly used in statistical practice while the first version is more useful in a SEM setting because it is more directly related to the actual estimated model and uses also the factor scores to predict the  $Y$  values. The first method however has a slight caveat. Because the factor score  $\hat{\eta}$  already incorporates the information contained in the dependent variable  $Y$ , then  $\hat{Y}$  will also contain information from  $Y$ . Therefore we will be using  $Y$  to predict  $Y$  which can in itself be problematic

as it lacks the rigor of pure statistical conditional expectation. However, when the latent variables  $\eta$  are measured well with a sufficient number of accurate measurements, the dependence between  $\hat{\eta}$  and a single  $Y$  variable, i.e., a particular indicator, will be sufficiently small so that it can be ignored. On a more technical level the purity of the  $\hat{Y}'$  definition guarantees that

$$Cov(\hat{Y}', \varepsilon) = 0$$

while

$$Cov(\hat{Y}, \varepsilon) \neq 0.$$

In most cases we assume that the systematic/predicted part of the variable is independent of its residual. Note however that if the measurement instrument is good

$$Cov(\hat{Y}, \varepsilon) \approx 0.$$

Thus any violation of the above equation could be interpreted as a model misspecification.

Estimates for the individual level residuals  $\varepsilon$  can also be formed using either of the estimated  $Y$  values

$$Y_{res} = Y - \hat{Y}$$

$$Y'_{res} = Y - \hat{Y}'.$$

Thus a common check that can be done using graphing utilities is to inspect that  $Y_{res}$  and  $\hat{Y}$  are independent.

In this paper we provide some simple examples where inspecting visually the scatter plots between the dependent variables  $Y$ , the estimated values  $\hat{Y}$ , the residual values  $Y_{res}$ , the estimated factor scores  $\hat{\eta}$ , and the covariates  $X$  can lead to useful discoveries and model modifications that could be overlooked otherwise. For our examples we use simulated data so that the illustration plots are somewhat clearer than those that might be found in real practical examples. The methodology, however, applies to real examples as well.

The plots do not infer statistical significance. This should come as usual from formal statistical testing. The plots however can be used to more meaningfully adjust the model. Model modification indices can sometimes produce large values for model parameters without clear justification. Usually model modifications that originate from a scatter plot are quite easy to interpret and explain. The power to detect model misspecifications with plots naturally will depend on the person analyzing the scatter plot and thus a formal statement regarding plot power is not possible. Presumably, however, the power will be lower than with other purely numerical statistical tools. Usually the sample size does little to improve the effect of a particular scatter plot. For example doubling the number of plotted points, i.e., replicating each plotted point, will not increase or make any difference in the plot, while any statistical test will increase its power. Our ability to detect visually a missing effect depends on the size of the effect rather than its statistical significance.

The examples provided in this paper are by no means exhaustive. These are just a few simple illustrations. Other scatter plots can be examined to

cross-validate the estimated model.

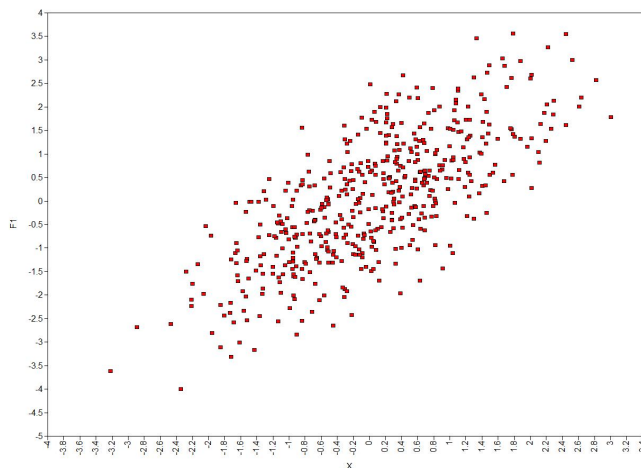
## 2 Non-linear trend

In this example we generate data according to a factor analysis model with one factor with 10 indicator variables and one covariate. All loadings  $\lambda$  are set to 1, all intercept parameters are set to 0, all residual variances are set to 1, all direct effect coefficients  $\beta$  are set to 0. We generate the data with a non-linear effect on the factor

$$\eta = \gamma X + F(X) + \zeta,$$

where  $X$  is a standard normal variable,  $\gamma = 1$  and  $F(X)$  uses of the these 3 forms: 0,  $X^2$ ,  $exp(X)$ . As we vary the function  $F$  we get 3 different samples. We generate the data sets with 500 observations. Next we estimate a standard factor analysis model where the non-linear trend is not accounted for, i.e., we estimate the same model as the data generating model but without the predictor  $F(X)$ . We plot the estimated factor score  $\hat{\eta}$  against the predictor  $X$ . The plots for the 3 samples are given in Figures 1-3. The results are very clear. In the first sample the model is adequate, i.e., the linear trend between  $\eta$  and  $X$  is correct and the plot reflects the linear trend that is in the model. In the second sample Figure 2 shows a quadratic relationship between  $\eta$  and  $X$ . Thus the model is insufficient and there is a need to add  $X^2$  as a predictor of  $\eta$ . In the third sample Figure 3 shows an exponential relationship between  $\eta$  and  $X$ . Thus the model is insufficient and there is a need

Figure 1: Sample 1 - linear trend,  $F(X) = 0$



to add  $exp(X)$  as a predictor of  $\eta$ . None of the standard fit statistics, such as the chi-square test, CFI, and TLI detected the misfit/inadequacy of the model in the second and the third samples. In addition the plots can actually provide guidance regarding the type of non-linearity that should be explored. The quadratic v.s. the exponential trend are clearly distinguishable.

### 3 Using residual values to detect direct effects

In this section we generate a sample as in the previous section with 10 indicator variables, one factor and 5 standard normal covariates  $X_1, \dots, X_5$ . All

Figure 2: Sample 2 - quadratic trend,  $F(X) = X^2$

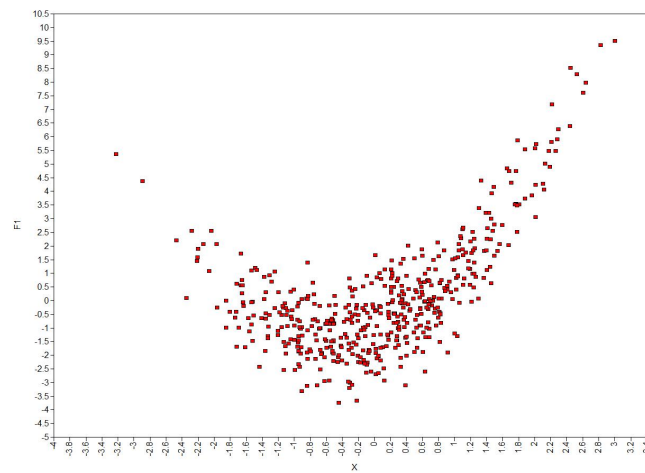
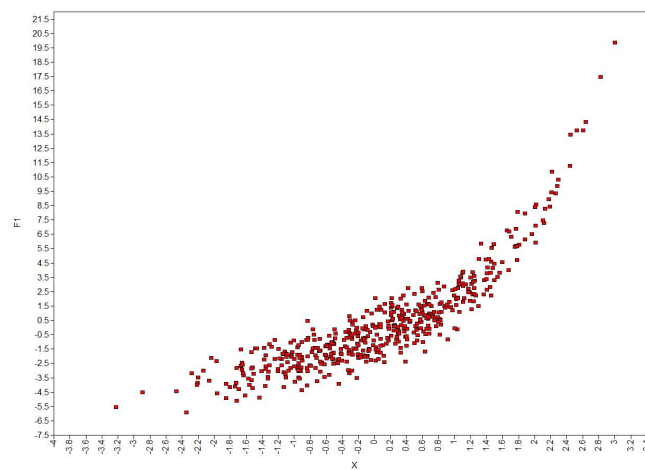


Figure 3: Sample 3 - exponential trend,  $F(X) = \text{Exp}(X)$



$\gamma_i$  coefficients are set to 1. We also add one direct effect to the model generation from  $X_1$  to  $Y_1$ , i.e.,  $\beta_{11} = 1$ . We estimate the MIMIC model without the direct effect, i.e., assuming that  $\beta_{11} = 0$ . We consider 4 different scatter plots:  $Y_1$  v.s.  $X_1$  given in Figure 4,  $Y_2$  v.s.  $X_1$  given in Figure 5,  $Y_{1,res}$  v.s.  $X_1$  given in Figure 6,  $Y_{2,res}$  v.s.  $X_1$  given in Figure 7.

Both Figure 4 and 5 show a linear trend between  $Y_i$  and  $X_1$ . These trends do not indicate any misspecification because they are implied by the estimated model. However, according to the estimated model which didn't account for the direct effect from  $X_1$  to  $Y_1$  the residual variables should be independent of  $X_1$ . Figure 6 shows a linear dependence between  $Y_{1,res}$  and  $X_1$ . This implies an unaccounted direct effect from  $X_1$  to  $Y$ . Figure 7 shows that  $Y_{2,res}$  is independent of  $X_2$  as expected and implies no misspecification. Therefore it is important to examine scatter plots for the residual variables and not just the observed variables. This type of misspecification is easily detectable however with other methods such as modification indices as well as by examining the model estimated variance covariance matrix for  $Y_i$  and  $X_i$  and the observed variance covariance matrix for these variables. This comparison is obtained in Mplus with the RESIDUAL option of the OUTPUT command.



Figure 4:  $Y_1$  v.s.  $X_1$

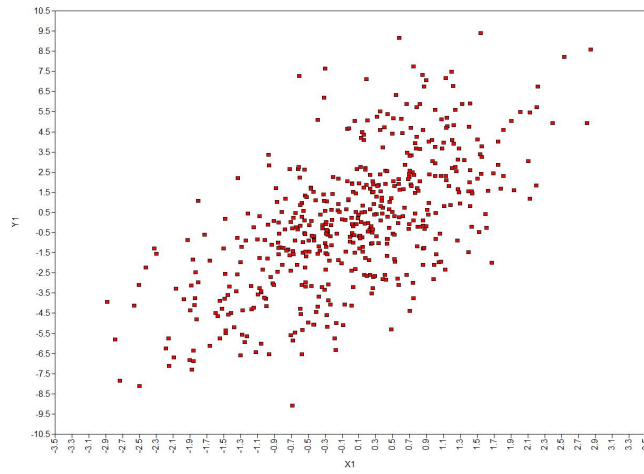


Figure 5:  $Y_2$  v.s.  $X_1$

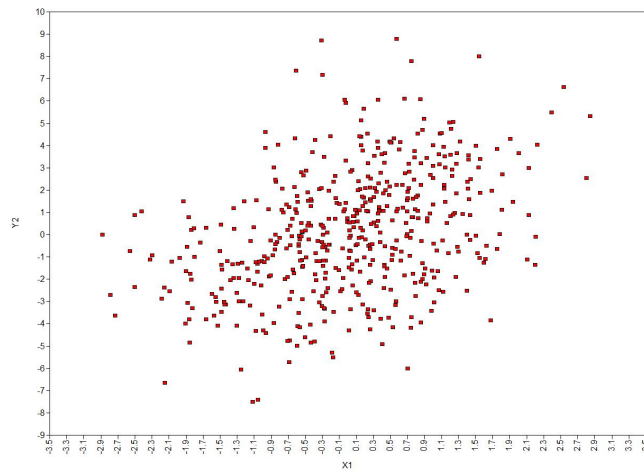


Figure 6:  $Y_{1,res}$  v.s.  $X_1$

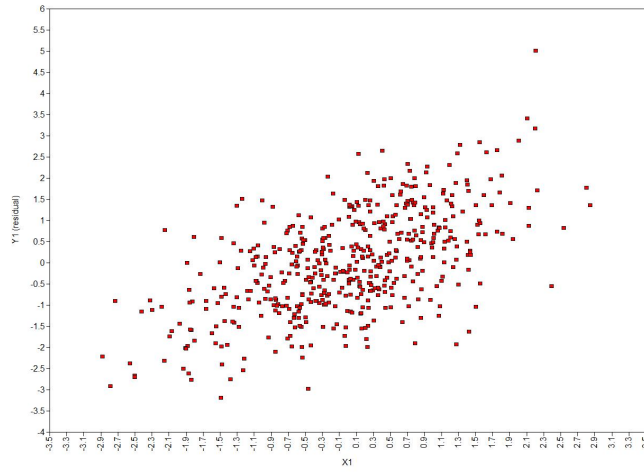
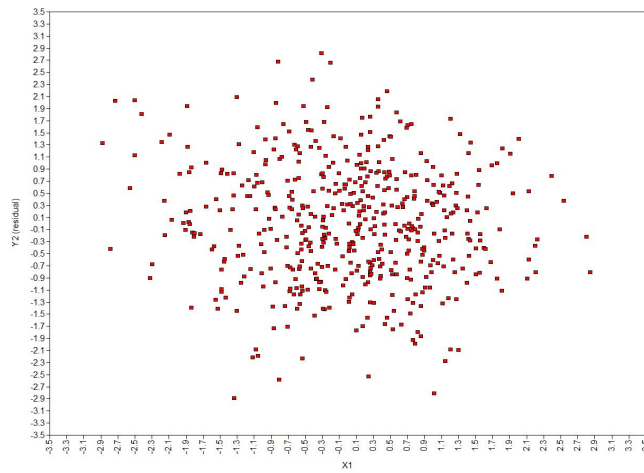


Figure 7:  $Y_{2,res}$  v.s.  $X_1$



## 4 Using residual scatter plots to detect residual covariances

In this section we generate a sample as in the previous section with 10 indicator variables, one factor and without any covariates. We introduce to the data generation model a residual correlation of 0.5 between  $Y_1$  and  $Y_2$ . The data however is analyzed without the residual correlation. We consider the scatter plot between  $Y_{1,res}$  and  $Y_{2,res}$  to see if the misspecification will be detected. This scatter plot is presented in Figure 8, while in Figures 9 we present the scatter plots between  $Y_{1,res}$  and  $Y_{3,res}$  for comparative purposes. Figure 8 shows that  $Y_{1,res}$  and  $Y_{2,res}$  are positively correlated and there is a linear trend between the two with a positive slope. This clearly suggest that the residual correlation between  $Y_1$  and  $Y_2$  should be added to the model. On the other hand Figure 9 shows that  $Y_{1,res}$  and  $Y_{3,res}$  appear to be independent which is correctly reflected in the model.

In this example the misspecification was detected by using scatter plots between the residual variables. A scatter plot between the observed variables would not be able to detect this misspecification because both the model and the scatter plot indicate a positive correlation between any pair of indicator variables.

This type of misspecification can also be detected using modification indices or by comparing the estimated and observed variance covariance matrices given in the residual output.

Figure 8:  $Y_{1,res}$  v.s.  $Y_{2,res}$

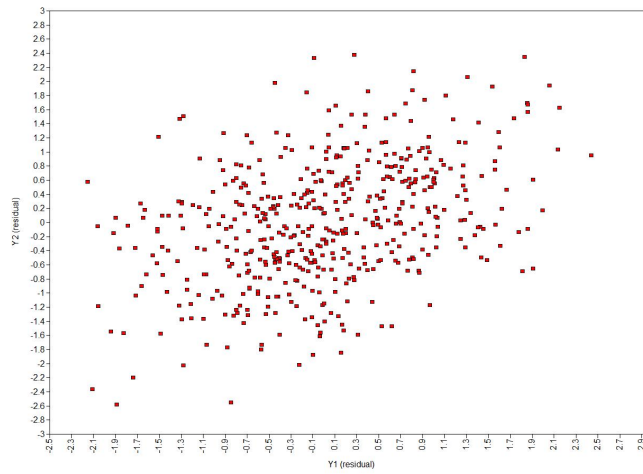
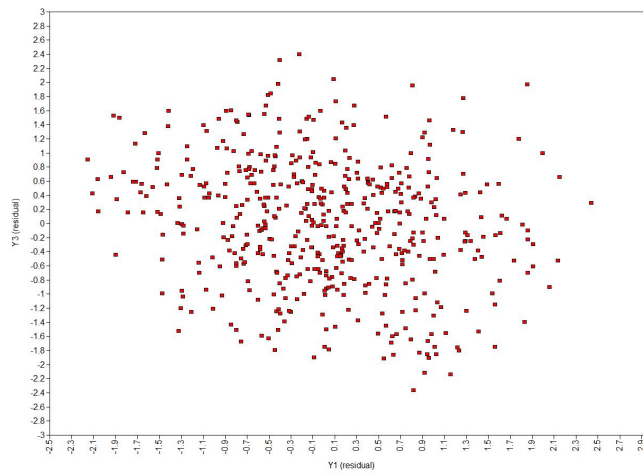


Figure 9:  $Y_{1,res}$  v.s.  $Y_{3,res}$



## 5 Using scatter plots to detect heterogeneity in the population

For this example we generate the data set as in the previous section using a factor analysis model with 10 indicators and one factor. The parameters are as in the previous sections except for the residual variances of the indicator variables which we set now at 0.1. In addition we generate the data from a two-class model with two equal classes where the factor loadings for  $Y_2$  in the second class is 0. All other loadings are 1. We estimate a model where the latent class variable is ignored, i.e., we estimate a simple factor analysis model with the 10 indicators and a single factor. We consider the scatter plot of  $Y_i$  v.s.  $\hat{Y}_i$ . Figure 10 contains the scatter plot  $Y_1$  v.s.  $\hat{Y}_1$  and Figure 11 contains the scatter plot  $Y_2$  v.s.  $\hat{Y}_2$ . In Figure 10 we see a good match between the observed and the estimated values. In Figure 11 however that is not the case. We can clearly see the two patterns in the data, indicating an underlying heterogeneity in the population and the potential for Mixture modeling to improve the model and provide a better fit for the data. In this example as well, the usual fit statistics such as the chi-square test, CFI and TLI are misleading because they indicate a nearly perfect fit and overlook the drastic discrepancy illustrated in Figure 11.

Figure 10:  $Y_1$  v.s.  $\hat{Y}_1$

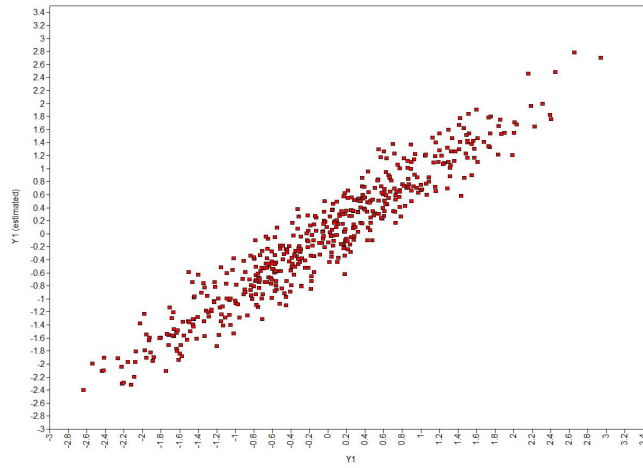
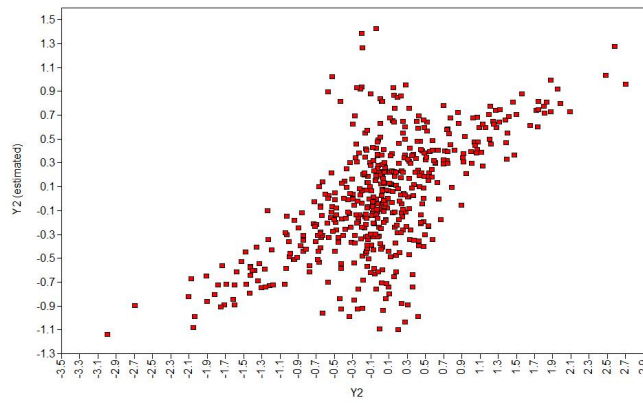


Figure 11:  $Y_2$  v.s.  $\hat{Y}_2$



## 6 Discussion

Residual analysis in regression is a well established statistical tool, however, in structural equation modeling this is not so. Because the latent variables in the model are not observed, to construct the residuals  $Y_{res}$  we need to use the factor score estimates for the latent variables. In this note we show that despite this complication, the residuals can successfully be used to evaluate model fit and discover needed model modifications. Bollen and Arminger (1991) explore other applications of SEM residual analysis such as outlier detection and individual observation test of fit using standardized individual residuals. Wang, Brown and Bandeen-Roche (2005) explore residual analysis in a Mixture context for model diagnostics including determining the number of latent classes. Raykov (2005) uses residual analysis for evaluating local goodness of fit in SEM. Residual analysis can also benefit from the use of plausible values for the latent variables, see Asparouhov and Muthén (2010), based on Bayesian estimation. Plausible values would allow us to use the residuals for further modeling that goes beyond standard SEM techniques and expands the utility of this methodology.

## References

- [1] Asparouhov, T. & Muthén, B. (2010). Plausible values for latent variables using Mplus. Technical Report.  
<http://statmodel.com/download/Plausible.pdf>
- [2] Bollen, K.A. & Arminger, G. (1991). Observational residuals in factor

analysis and structural equation models, *Sociological methodology*, 235-262

- [3] Raykov, T. (2005). Residuals in Structural Equation, Factor Analysis, and Path Analysis Models. *Encyclopedia of Statistics in Behavioral Science*.
- [4] Wang, C.P., Brown, H. & Bandeen-Roche, K. (2005). Residual Diagnostics for Growth Mixture Models, *Journal of the American Statistical Association*, 100 , 1054-1076.