Blurring the line between confirmation and exploration: Model comparison of structural equation modeling in JMP/SAS®

Chong Ho Yu, Ph.Ds., Azusa Pacific University, CA

ABSTRACT

Structural equation modeling, as the name implies, aims to confirm pre-determined structural relationships between various factors and variables. Conventionally, path searching, variable selection, and model building should be done at the exploratory stage only. Nevertheless, equipped with the new JMP interface to SAS , today the analyst is able to compare different models with different combinations of paths and variables based on several fitness criteria, such as Akaike Information Criterion and Bayesian Information Criterion. In this sense the distinction between confirmation and exploration is blurred. The objective of this paper is to demonstrate how confirmatory procedures could be "explored" in SAS via the JMP interface. Unlike TETRAD that is capable of automated path searching, model comparison in SAS/JMP is still driven by the analyst. In this paper an example based on a World Bank data set is used to illustrate why it is essential to perform manual exploration on some occasions. The analysis utilizing the World Bank data set indicates that the number of college and university graduates majoring in science in a particular year is a significant predictor of the number of scientific and technical journal articles published two years later, and this variable can predict future productivity, as measured by GDP per worker.

INTRODUCTION

Structural equation modeling (SEM) is one of the most powerful multivariate analysis techniques. But researchers who employ SEM have to know exactly what they want to do by using background information, prior knowledge, and past research. In other words, exploring and revising models "on the fly" without any theoretical foundation is usually discouraged. Some researchers (e.g., Glymour, 2003, 2004; Glymour, Madigan, Pregibon, & Smyth, 1996; Scheines, Spirtes, Glymour, Meek, & Richardson, 1998) have attempted to introduce exploratory elements into SEM by developing path searching software applications, such as TETRAD. Specifically, the objective of path searching is to address a severe threat against the validity of SEM, namely, model equivalency. Model equivalency is a well-known problem in SEM: even if the data and the model could fit each other very well, it doesn't necessarily imply that there are causal relationships among the factors and the variables. It is conceivable that some other equivalent models could fit the data equally well. To counteract this shortcoming, path searching examines many equivalent and even non-equivalent models by exhausting almost all possible combination of factors and variables. For example, a researcher might initially propose a path model like this: A→B→C. In path searching other possible models could be: A→C→B, B→C→A, and C→B→A. Path searchers would not commit themselves to a particular model until many other possibilities are considered and the best fit emerges.

Path searching works well in many situations, but in some cases this approach is not viable at all, as will be explained in a later section. When path searching is not an option, using the JMP interface to run SEM in SAS is highly recommended. In the following a World Bank dataset will be used to illustrate the procedure.

DATA SOURCE

The data source for this project is the archival data set entitled World Development Indictors (WDI) and Global Development Finance (GDF), which is downloadable from the World Bank (2012). This comprehensive data set contains indicators for each country of national well-being, including data concerning a country's education, environment, economic policies, financial sector, health, infrastructure, labor force, social protection, poverty, and international trade. The variables chosen for this project are as follows:

1. Sci 2003 : the percentage of people who graduated from college or university in 2003 with a major in science.
2. EMC 2003: the percentage of people who graduated from college or university in 2003 with a major related to engineering, manufacturing, or construction (EMC).
3. Paper 2005: the number of scientific and technical papers published in peer-review journals in 2005.
4. Patent 2005: the number of patents applied for by residents in 2005.
5. Productivity 2007: Gross domestic product per person employed in 2007.

Only forty nations have complete data that include all of the above variables. The author is aware that this sample size may be insufficient for SEM. However, this limitation is insurmountable because even a large international organization like the World Bank was unable to collect data in certain countries.

Some readers may wonder why the author did not use the data in the same year. It is important to point out that sometimes a concurrent cross-sectional design may be problematic. It is unrealistic to expect that graduates in 2003

could immediately play a role in producing research products, such as scientific journal papers and patents. By the same token, it is unlikely that new findings presented in journal papers and new innovations created by patent holders could lead to instant urge in productivity. Thus, it is reasonable to conduct a time-lag analysis by leaving a two-year gap between different groups of variables. Simply put, some effects must take time to materialize. 2007 productivity was chosen in order to countermeasure one of the major threats against the internal validity of a study, namely, history. Obviously, the 2008 financial tsunami turned many figures upside down, and thus the author avoided using the data collected after 2007. The initial conjecture is that the number of graduates in science and EMC in a given year might positively influence the number of scientific papers published in peer-review journals and the number of patents applied by residents two years later. Patents held by non-residents are not taken into account because their accomplishment might not be attributed to local education. Subsequently, new ideas and new innovations manifested in research papers and patents could eventually improve productivity.
Obviously, path searching cannot be used in this situation because automated path searching algorithm is blind to the temporal nature of variables. Unless one is a Star Trek fan who believes that time travel could happen, it is impossible to build a model like the following: 2007 variables→2005 variables→2003 variables. Hence, in this study path building and model comparison using the JMP interface to SAS will be used as a replacement of automated path searching.

PRIOR RESEARCH AND PRELINMINARY ANALYSIS

As mentioned before, one cannot construct a SEM without any background knowledge or exploratory research about the subject matter under study. In another time-lag study conducted by the author and his colleagues (Yu, DiGangi, & Jannasch-Pennell, 2012), it was found that out of eight potential predictors, only scientific and technical journal articles published previously before2000 could accurately predict economic performance, as measured by GDP per capita in 2007. The variable "scientific and technical journal articles" was reused in the current study, but the indicator of productivity was changed to GDP per worker, because GDP per capita takes non-workers into account, which may skew the result. In the previous study the time-lag is a seven-year interval but in this one the gap is shortened, because the author is interested in examining shorter-term effects.

After the variables are chosen for the current study based on prior research, data visualization and regression modeling were employed for preliminary analysis. All graphs below are vertical histograms showing the frequency counts by binning. Figure 1(a) and 2(a) depict the untransformed distributions of 2005 scientific papers and patents. Both are extremely skewed distributions because scientific research and innovations tend to concentrate on very few developed nations, such as the US and Japan. As a remedy, natural logarithm transformation was utilized to normalize the distributions (see Figure 1(b) and 2(b)).

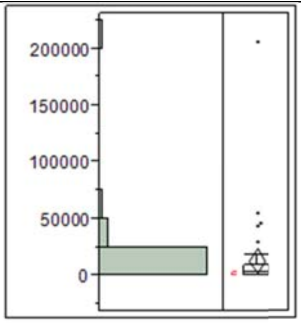
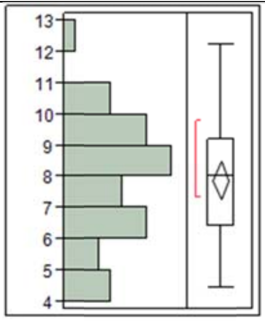
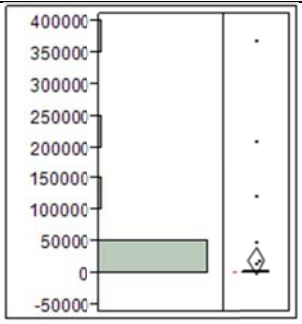
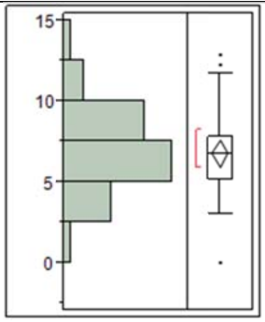
Figure 1(a). Untransformed distribution of 2005 scientific articles.	**Figure 1(b). Natural log distribution of 2005 scientific articles.**	**Figure 2(a). Untransformed distribution of 2005 patents by residents.**	**Figure 2(b). Natural log distribution of 2005 patents by residents.**
			

Figure 3 shows some of the inter-relationships among the variables. Only the pairs that form a linear pattern are included into the panels. The top left panel indicates that when South Korea was included into the data set, the correlation between 2003 science graduates and EMC graduates is insignificant. But it turned to significant after South Korea (the red dot) was removed. Because this analysis is multivariate in nature, the author decided not to delete this bivariate outlier, otherwise the small sample size would get even smaller. After all, the relationship between two sets of graduates is not the focal point of this study.

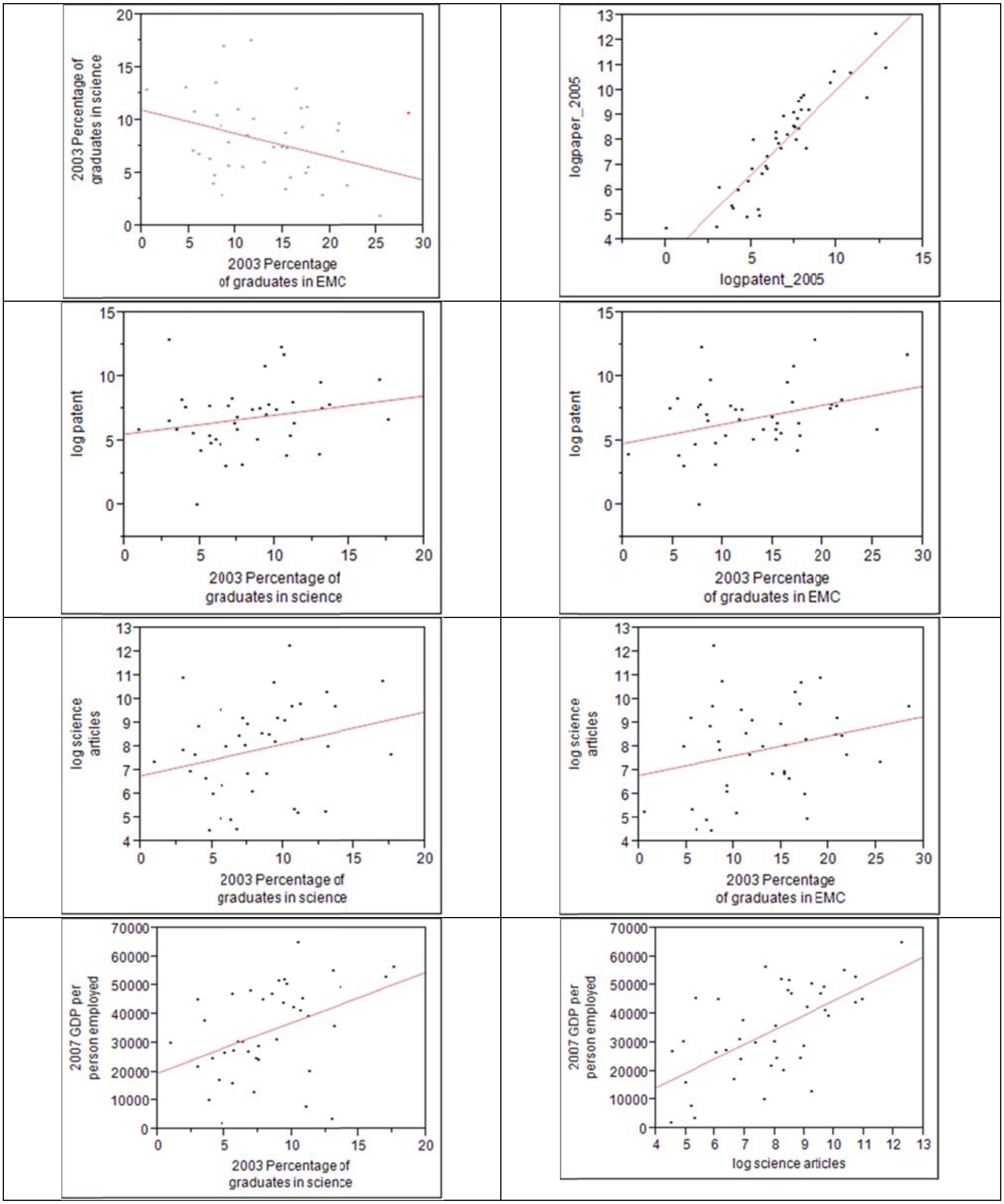


Figure 3. Scatterplot matrix of variables.

Tables 1 to 4 show the regression results. Obviously, unlike SEM, regression modeling disallows the researcher to put the variables in a causal pathway. In each regression there is only one outcome variable and all others are independent variables. This shortcoming necessitates SEM, in which some variables could be endogenous and exogenous variables simultaneously. For example, in the causal pathway A→B→C, B is a predictor because C responds to B, but at the same time B also depends on A. Nonetheless, the isolated pieces of information yielded from regression tell the researcher that while the percentage of 2003 science graduates is associated with 2007 productivity, there are other variables between them. Specifically, the natural log values of 2005 scientific articles and patents are related to the percentages of 2003 science and EMC graduates. And the natural log of 2005 scientific articles could substantially affect productivity in 2007. These scattered pieces could be put together to generate a structural equation model. Due to the tentativeness and exploratory nature of the regression analyses, assumptions for regression, such as independence, normality, and homoscedasticity of residuals, have not been discussed here.

Variable	DF	Parameter Estimate	Standard Error	*t* Value	*p* Value
Intercept	1	8487.17369	8350.05059	1.02	0.3160
2003 percentage of graduates in science	1	2033.42823	621.69946	3.27	**0.0023**
2003 percentage of graduates in engineering, manufacturing, and construction	1	681.85963	382.47475	1.78	0.0828

Table 1. Using 2003 percentage of graduates in science and EMC to predict 2007 productivity.

Variable	DF	Parameter Estimate	Standard Error	*t* Value	*p* Value
Intercept	1	4.93304	1.05877	4.66	<.0001
2003 percentage of graduates in science	1	0.18183	0.07883	2.31	**0.0268**
2003 percentage of graduates in engineering, manufacturing, and construction	1	0.11169	0.04850	2.30	**0.0270**

Table 2. Using 2003 percentage of graduates in science and EMC to predict natural log of 2005 scientific and technical journal papers.

Variable	DF	Parameter Estimate	Standard Error	*t* Value	*p* Value
Intercept	1	2.47084	1.35427	1.82	0.0762
2003 percentage of graduates in science	1	0.22748	0.10083	2.26	**0.0301**
2003 percentage of graduates in engineering, manufacturing, and construction	1	0.18642	0.06203	3.01	**0.0047**

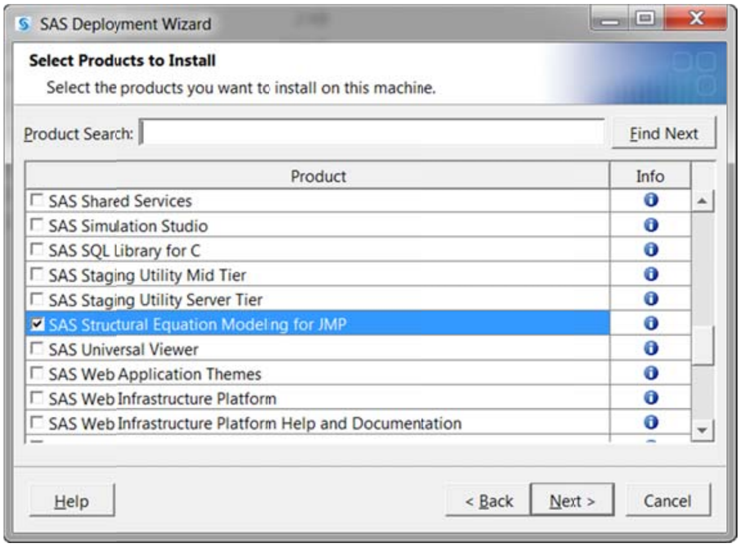
Table 3. Using 2003 percentage of graduates in science and EMC to predict natural log of 2005 patents by residents.

Variable	DF	Parameter Estimate	Standard Error	*t* Value	*p* Value
Intercept	1	-7092.08477	9751.05635	-0.73	0.4716
Natural log of 2005 scientific and technical papers published in peer review journals	1	5600.03931	2413.77859	2.32	**0.0260**
Natural log of 2005 patents applied by residents	1	-433.62321	1829.00917	-0.24	0.8139

Table 4. Using 2005 natural log of 2005 scientific papers and patents to predict 2007 productivity.

JMP INTERACE TO SAS

In order to run SEM, one must have access to both JMP 9 or above, and SAS 9.2 or above. SAS does not have to be installed in the local machine. Rather, it could be located in a remote server. But it is important to point out that when SAS is originally installed, the module “SAS Structural Equation Modeling for JMP” must be checked (see Figure 4).



SAS Deployment Wizard

Select Products to Install

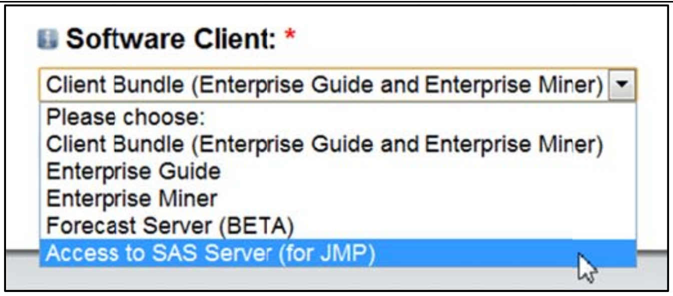
Select the products you want tc install on this machine.

Product Search: [ ] Find Next

Product	Info
□ SAS Shared Services	ⓘ
□ SAS Simulation Studio	ⓘ
□ SAS SQL Library for C	ⓘ
□ SAS Staging Utility Mid Tier	ⓘ
□ SAS Staging Utility Server Tier	ⓘ
☑ SAS Structural Equation Model ng for JMP	ⓘ
□ SAS Universal Viewer	ⓘ
□ SAS Web Application Themes	ⓘ
□ SAS Web Infrastructure Platform	ⓘ
□ SAS Web Infrastructure Platform Help and Documentation	ⓘ

Figure 4. SAS Structural Equation Modeling for JMP.

Although the JMP user can connect JMP to a remote server (se Figure 5a), the easiest way is to use SAS on the local machine. To establish the connection between JMP and SAS, open “SAS Server Connection” from File, choose the SAS version on the local computer, check the radio button “Connect to SAS on this machine,” and then click on the button “Connect” (see Figure 5b).

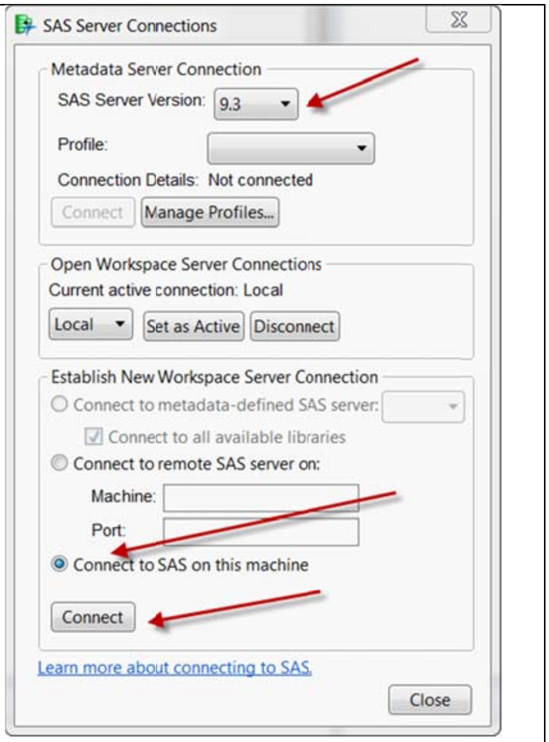


Software Client: *

Client Bundle (Enterprise Guide and Enterprise Miner)

- Please choose:
- Client Bundle (Enterprise Guide and Enterprise Miner)
- Enterprise Guide
- Enterprise Miner
- Forecast Server (BETA)
- Access to SAS Server (for JMP)

Figure 5(a). Access to SAS server.



SAS Server Connections

Metadata Server Connection

SAS Server Version: 9.3

Profile:

Connection Details: Not connected

Open Workspace Server Connections

Current active connection: Local

Establish New Workspace Server Connection

○ Connect to metadata-defined SAS server:

☑ Connect to all available libraries

○ Connect to remote SAS server on:

Machine:

Port:

● Connect to SAS on this machine

Figure 5(b). Establish connection between JMP and SAS.

After the connection between JMP and SAS is established, the user can run SEM via Analyze→Structural Equation Modeling→Single Group Analysis. In the tab “Analysis” the user can click on “Palette” to activate the drawing mode. In this mode the analyst can simply drag and drop the variables into a canvas to construct a model based on the preliminary regression analysis (see Figure 6). In this example all variables are observed and no latent constructs are used. After the proposed model is built, the researcher can click on the “Run” button to execute the computation.

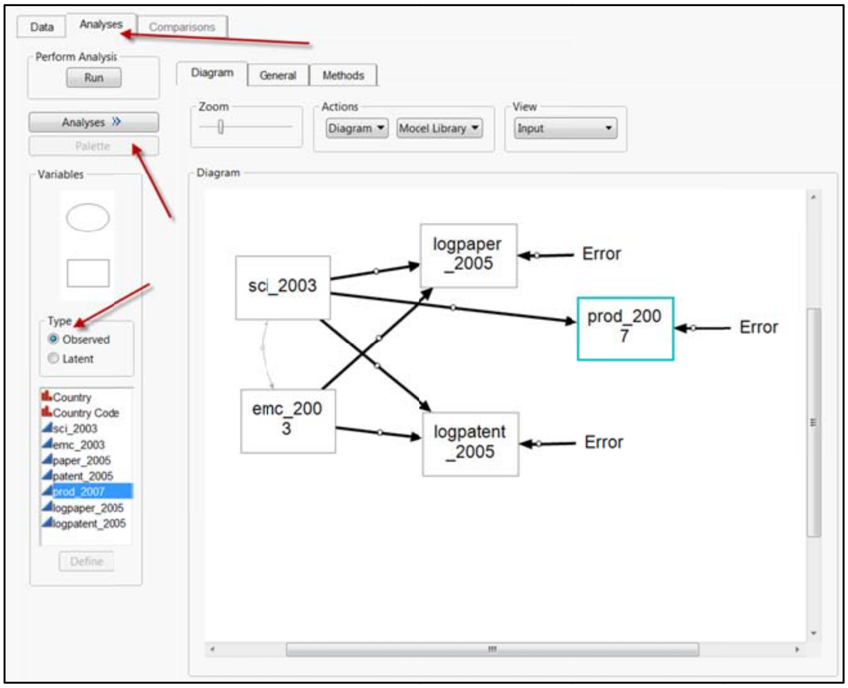


Figure 6. Constructing a SEM by drawing.

The researcher can see the result in just a few seconds. By default the regression coefficient estimates shown on the graph are unstandardized. To obtain a better interpretation, the “View” can be changed to standardized estimates. The parameters with the sign “*” or “**” are considered significant. One asterisk indicates that p < 0.05 whereas two asterisks indicate that p < 0.01 (see Figure 7).

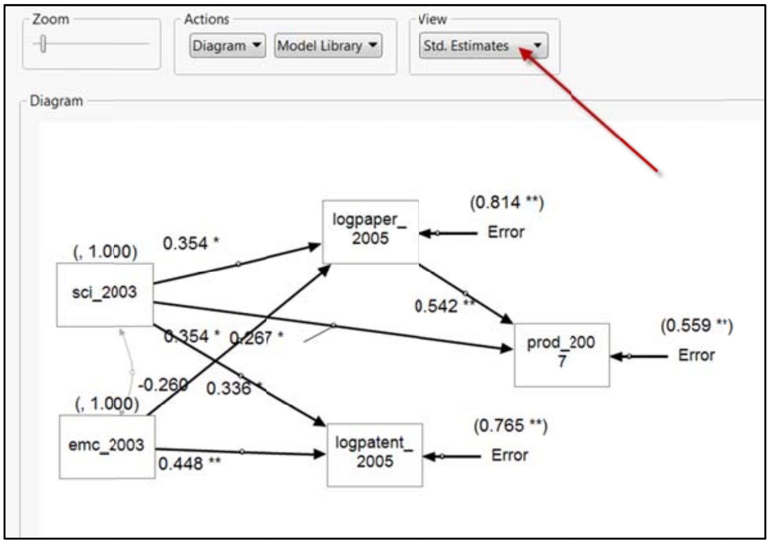
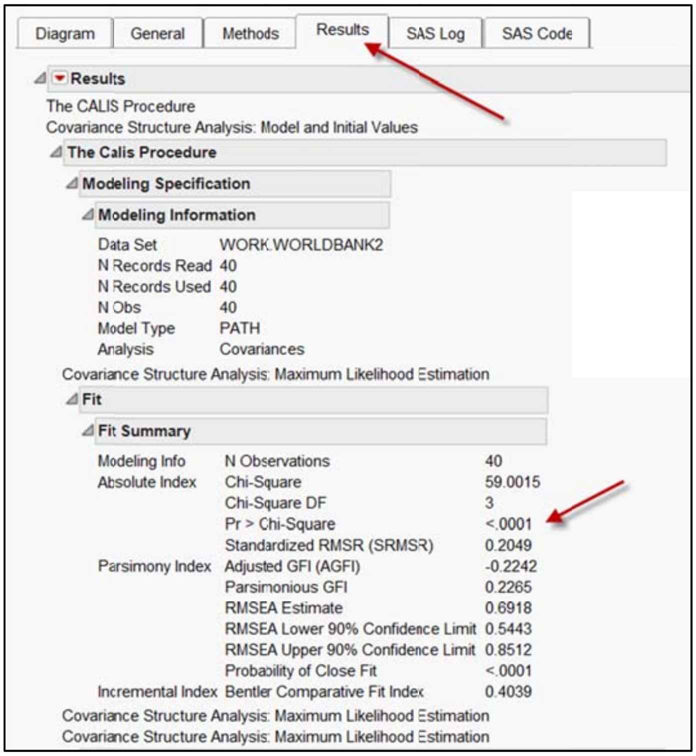


Figure 7. View of standardized estimates.

To unveil more details, the user can open the tab “Results.” It is important to check whether there is any error message. If there is no error, then the convergence is successful. In this model, the chi-square statistics suggest that the model does not seem to be promising (X^2=59.0015, *p*<.0001). The Chi-square test is a measure of the goodness of fit between the observed and the expected. The null hypothesis is that there is no significant discrepancy between the covariance matrix generated by the model and that observed in the data. A *p* value that is small enough to reject the null signifies that the data-model fit is questionable. It is a well-known fact that the significance of the Chi-square is subject to the sample size (Besag, 1980). If the sample size is very large, it is more likely that the model will be unfairly rejected, However, in this small data set (n=40) this is not a concern at all. In addition to the Chi-square, there are many other fitness indices, such as AGFI, RMSEA, Bentler Comparative Fit Index…etc. (see Figure 8).



```
Results
The CALIS Procedure
Covariance Structure Analysis: Model and Initial Values
  The Calis Procedure
    Modeling Specification
      Modeling Information
        Data Set        WORK.WORLDBANK2
        N Records Read  40
        N Records Used  40
        N Obs           40
        Model Type      PATH
        Analysis        Covariances
Covariance Structure Analysis: Maximum Likelihood Estimation
    Fit
      Fit Summary
        Modeling Info      N Observations                    40
        Absolute Index     Chi-Square                        59.0015
                           Chi-Square DF                     3
                           Pr > Chi-Square                   <.0001
                           Standardized RMSR (SRMSR)         0.2049
        Parsimony Index    Adjusted GFI (AGFI)               -0.2242
                           Parsimonious GFI                  0.2265
                           RMSEA Estimate                    0.6918
                           RMSEA Lower 90% Confidence Limit  0.5443
                           RMSEA Upper 90% Confidence Limit  0.8512
                           Probability of Close Fit          <.0001
        Incremental Index  Bentler Comparative Fit Index     0.4039
Covariance Structure Analysis: Maximum Likelihood Estimation
Covariance Structure Analysis: Maximum Likelihood Estimation
```

Figure 8. Poor fitness indices in the initial model.

To rectify the poor fit, an alternate model should be proposed and thoroughly examined. JMP provides the user with a “Copy” button and thus model revision in JMP is relatively easy. The existing model is retained for model comparison in a later stage. Rather than starting over from scratch, the user can drops some variables and redraw the paths in the cloned model (see Figure 9).

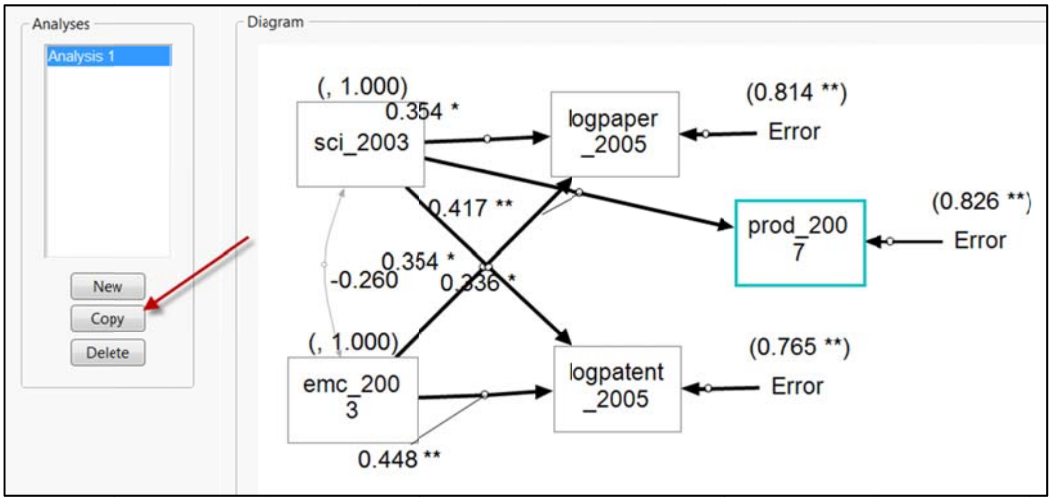


Figure 9. Copy and paste to revise the existing model.

The new model is put into a new analysis (see Figure 10). Because the ultimate goal is to identify the variables that could make contributions to productivity, the natural log of 2005 patents is redundant and thus it is taken out of the equation.

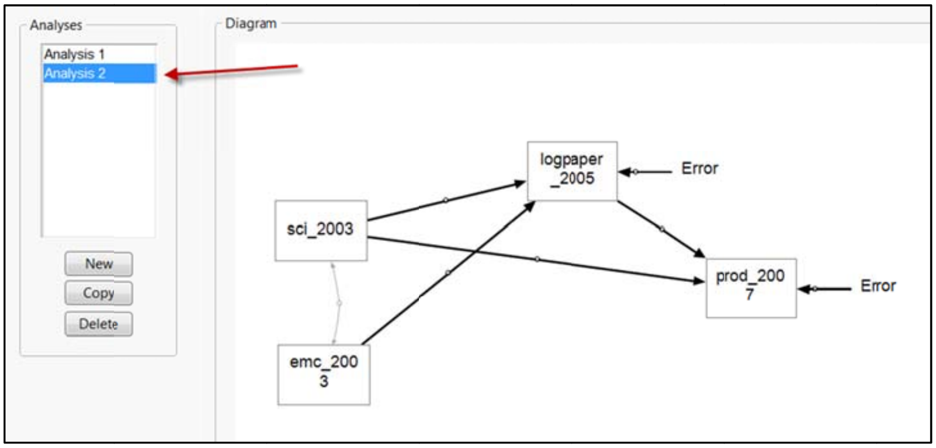


Figure 10. A new model without natural log of 2005 patents.

Figure 11 depicts the new model with standardized estimates. Figure 12 shows that the fitness has been substantively improved. Some researchers might want to stop at this point and accept this one as the final. In the past this decision was understandable because running SEM is very involved and time-consuming. However, being equipped with the JMP interface, today the user can afford further exploration with minimal efforts.

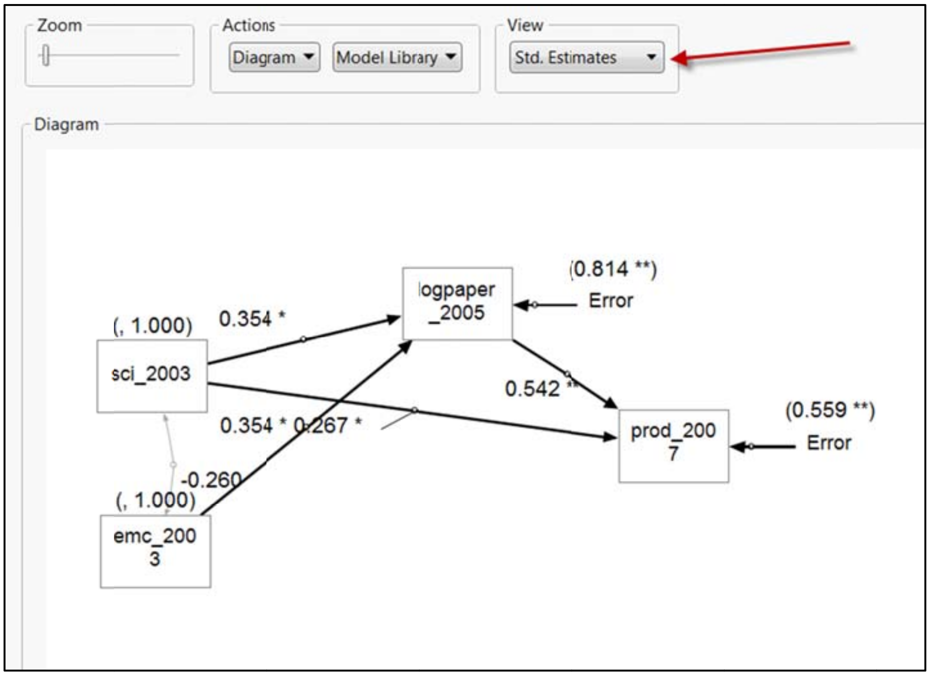
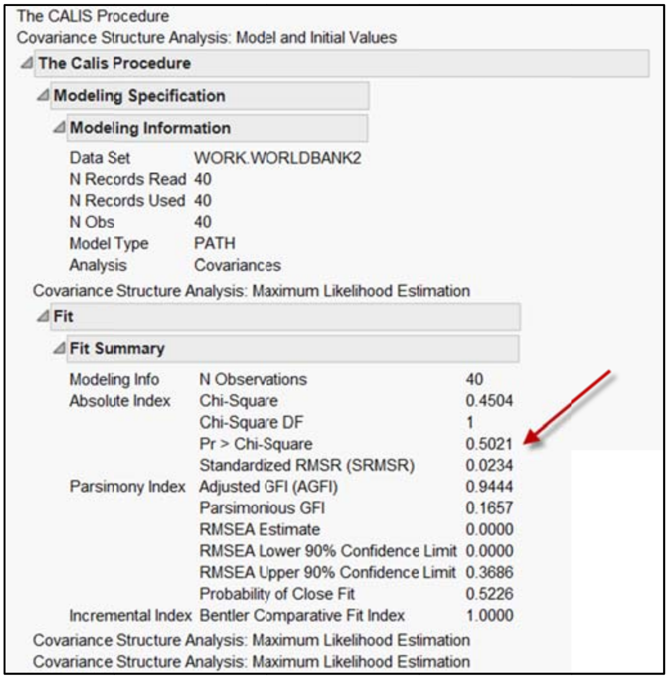


Figure 11. A new model without natural log of 2005 patents.



```
The CALIS Procedure
Covariance Structure Analysis: Model and Initial Values
◿ The Calis Procedure
  ◿ Modeling Specification
    ◿ Modeling Information
      Data Set          WORK.WORLDBANK2
      N Records Read    40
      N Records Used    40
      N Obs             40
      Model Type        PATH
      Analysis          Covariances
  Covariance Structure Analysis: Maximum Likelihood Estimation
  ◿ Fit
    ◿ Fit Summary
      Modeling Info      N Observations                        40
      Absolute Index     Chi-Square                            0.4504
                         Chi-Square DF                         1
                         Pr > Chi-Square                       0.5021
                         Standardized RMSR (SRMSR)             0.0234
      Parsimony Index    Adjusted GFI (AGFI)                   0.9444
                         Parsimonious GFI                      0.1657
                         RMSEA Estimate                        0.0000
                         RMSEA Lower 90% Confidence Limit      0.0000
                         RMSEA Upper 90% Confidence Limit      0.3686
                         Probability of Close Fit              0.5226
      Incremental Index  Bentler Comparative Fit Index         1.0000
  Covariance Structure Analysis: Maximum Likelihood Estimation
  Covariance Structure Analysis: Maximum Likelihood Estimation
```

Figure 12. A new model without natural log of 2005 patents.

Again, using the copy button the user can create Analysis 3 in a new canvas. If EMC graduates have no significant effects on productivity, could this variable be dropped, too? This time only three variables remained in the model, as shown in Figure 13. The standardized estimates and the fitness indices are shown in Figure 14 and Figure 15, respectively.

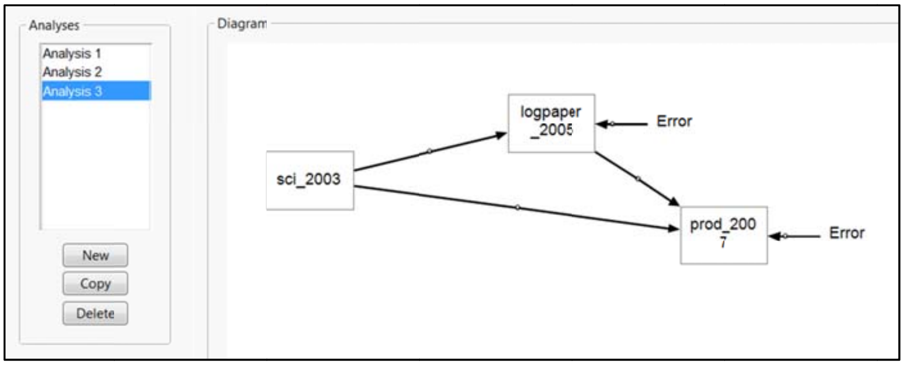


Figure 13. A highly parsimonious model with three variables.

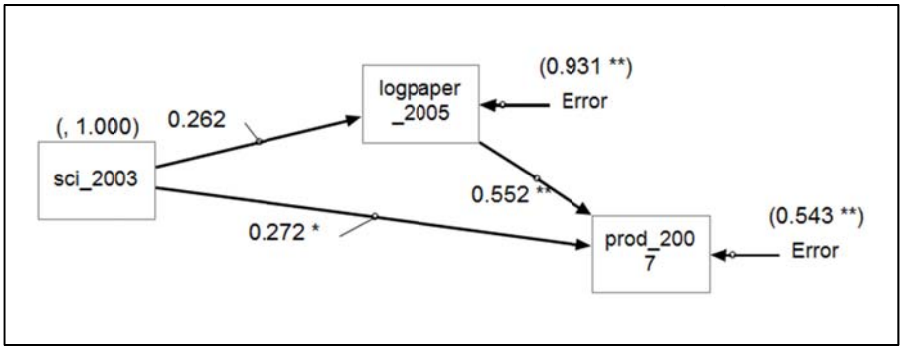
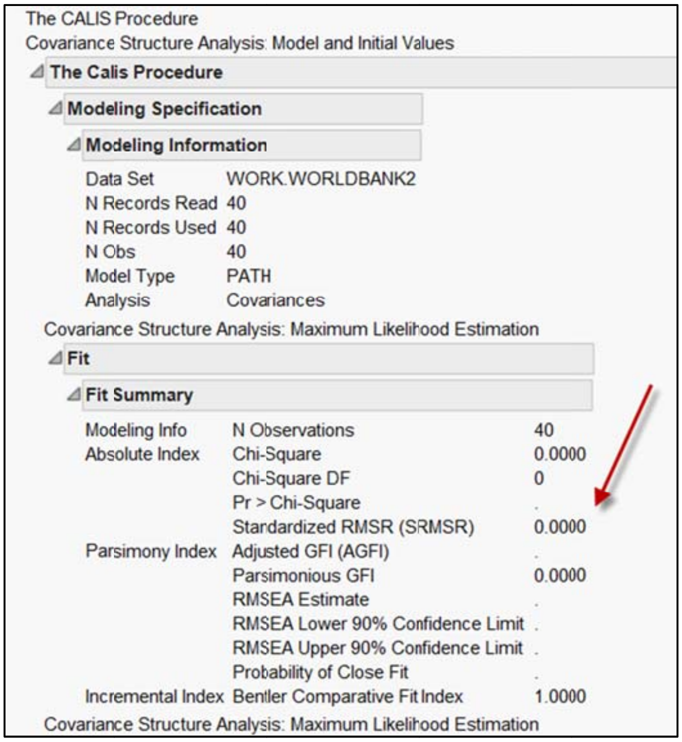


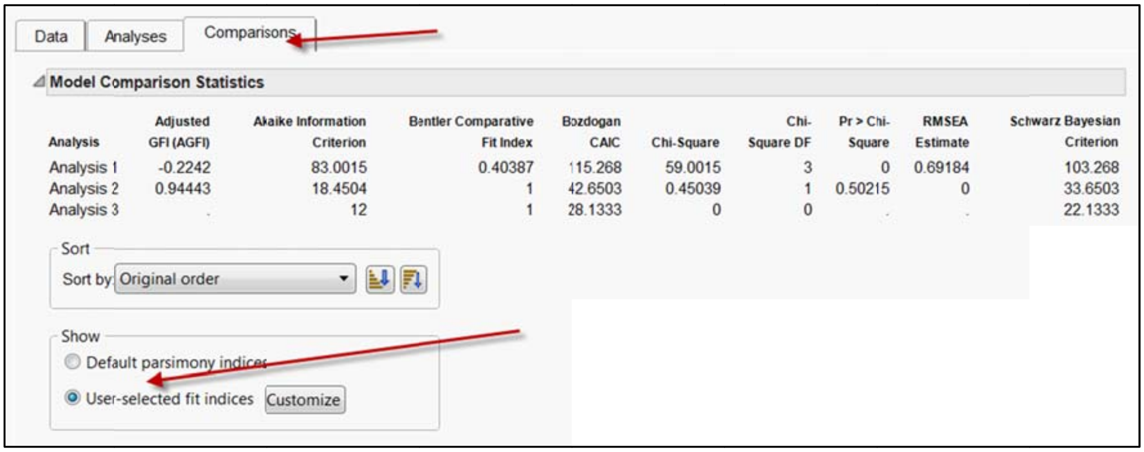
Figure 14. Standardized estimates of the parsimonious model.



```
The CALIS Procedure
Covariance Structure Analysis: Model and Initial Values
 ⊿ The Calis Procedure
   ⊿ Modeling Specification
     ⊿ Modeling Information
       Data Set          WORK.WORLDBANK2
       N Records Read    40
       N Records Used    40
       N Obs             40
       Model Type        PATH
       Analysis          Covariances
  Covariance Structure Analysis: Maximum Likelihood Estimation
   ⊿ Fit
     ⊿ Fit Summary
       Modeling Info      N Observations                      40
       Absolute Index     Chi-Square                          0.0000
                          Chi-Square DF                       0
                          Pr > Chi-Square                     .
                          Standardized RMSR (SRMSR)           0.0000
       Parsimony Index    Adjusted GFI (AGFI)                 .
                          Parsimonious GFI                    0.0000
                          RMSEA Estimate                      .
                          RMSEA Lower 90% Confidence Limit    .
                          RMSEA Upper 90% Confidence Limit    .
                          Probability of Close Fit            .
       Incremental Index  Bentler Comparative Fit Index       1.0000
  Covariance Structure Analysis: Maximum Likelihood Estimation
```

Figure 15. Fitness indices of the parsimonious model.

The information portrayed in Figure 15 may be counter-intuitive. In this panel, both the Chi-square and the degree of freedom show a "zero" while the p value is missing. In the context of a bivariate analysis, zero degree of freedom is problematic. It means that there is no useful independent information to do any meaningful estimation of the relationship (Yu, 2011). Is this the case here? As mentioned before, the Chi-square test is a measure of the goodness of fit between the expected and the observed. If there is no discrepancy between the expected model and the observed data, the Chi-square is zero, of course. In this case, the model is said to be saturated, meaning that the model can perfectly reproduce all of the variances, covariance, and means. Some researchers argue that such a "perfect" model has no explanatory value at all. Nonetheless, a saturated model is still useful for functioning as a baseline model with which other non-saturated models are compared.



⊿ Model Comparison Statistics

Analysis	Adjusted GFI (AGFI)	Akaike Information Criterion	Bentler Comparative Fit Index	Bozdogan CAIC	Chi-Square	Chi-Square DF	Pr > Chi-Square	RMSEA Estimate	Schwarz Bayesian Criterion
Analysis 1	-0.2242	83.0015	0.40387	115.268	59.0015	3	0	0.69184	103.268
Analysis 2	0.94443	18.4504	1	42.6503	0.45039	1	0.50215	0	33.6503
Analysis 3	.	12	1	28.1333	0	0	.	.	22.1333

Sort
Sort by: Original order

Show
○ Default parsimony indices
● User-selected fit indices Customize

Figure 16. Model comparison.

At the last stage of the process, the modeler can perform a model comparison by choosing the "Comparisons" tab (see Figure 16). By default JMP shows Akaike Information Criteria, Bozdogan CAIC, Schwartz Bayesian Criterion, and RMSEA. The user can check the box "User-selected fit indices" to reveal more options. There is no consensus regarding what the best fitness index is. Tomarken and Waller (2005) wrote, "For a number of years, the most common criteria for fit indices that have been used by behavioral researchers are rules of thumb that lack a detailed mathematical or empirical justification" (p.54). The author has no intention to settle this issue once and for all. His suggestions are more philosophical than mathematical. It is recommended that the researcher must focus on the ultimate goal and trim redundant information as much as possible. The Akaike's information criterion (AIC) developed by Akaike (1973) is in alignment to Ockham's razor: Given all things being equal, the simplest model tends to be the best one; and simplicity is a function of the number of adjustable parameters. Thus, a smaller AIC suggests a "better" model. Specifically, AIC is a fitness index for trading off the complexity of a model against how well the model fits the data. The general form of AIC is: AIC = 2k – 2lnL where k is the number of parameters and L is the likelihood function of the estimated parameters. Increasing the number of free parameters to be estimated improves the model fitness, however, the model might be unnecessarily complex. To reach a balance between fitness and parsimony, AIC not only rewards goodness of fit, but also includes a penalty, which is an increasing function of the number of estimated parameters. This penalty discourages over-fitting and complexity. Hence, the best model is the one with the lowest AIC value. Since AIC attempts to find the model that best explains the data with a minimum of free parameters, it is considered an approach favoring simplicity. In this example, the AIC value of the initial model is 83, which is relatively high. The last model, which has a much smaller AIC (12), is simpler but we might not like a saturated model that would have no explanatory value. Thus, we might want to settle down with the middle one.

It is concluded that a high percentage of graduates majoring in science and EMC could lead to better scientific research, indicated by a higher volume of research papers. And better research might eventually benefit productivity. Indeed, even the variable "2003 science graduates" alone is a strong predictor of 2007 productivity. This finding contradicts the popular belief that engineering and applied science is more valuable than pure science in terms of helping the economy. However, the author acknowledges that research based on nation-level, aggregate data is subject to ecological fallacy (Freedman, 1999). It is a bold assumption that some graduates in 2003 contributed to the scientific reports in 2005, such as playing the role of research assistants. Thus, readers should interpret these results with caution.

CONCLUDING REMARKS

PROC TCALIS available in SAS 9.2 or above is said to be a superior procedure (Gu & Wu, 2011; Yung, 2008). For instance, PROC TCALIS is capable of running multi-group SEM while the JMP interface to SAS is confined to single group analysis. However, the SAS syntax may be intimidating to novices and even experienced users. For a long time SEM applications with a graphical user interface, such as EQS and AMOS, have been widely welcome by modelers. Nonetheless, this author found that the JMP interface is better than many other GUI-based SEM packages. For example, JMP uses a contextual menu and dialog system, and thus the user is not overwhelmed by many options in the first screen. When automated path search is inappropriate, manual path searching and model building in JMP is highly recommended.

REFERENCES

Akaike, H. (1973). Information theory and an extension of the maximum likelihood principle. In B. N. Petrov and F. Csaki (Eds.), *International Symposium on Information Theory* (pp. 267–81). Budapest: Akademia Kiado.

Besag, F. (1980). Academic science, policy decision, and Chi-square. *Urban Education,15*, 215-230. DOI: 10.1177/0042085980152006

Freedman, D. (1999). Ecological inference and the ecological fallacy. Retrieved from http://www.stanford.edu/class/ed260/freedman549.pdf

Glymour, C. (2003). Learning, prediction and causal Bayes nets. *Trends in cognitive sciences, 7*, 43-48.

Glymour, C. (2004). The automation of discovery. *Dædalus, 133*, 69-77.

Glymour, C. Madigan, D., Pregibon, D., & Smyth, P. (1996). Statistical inference and data mining. *Communications of ACM, 39*, 35-41.

Gu, F., & Wu, W. (2011). Using SAS PROC TCALIS for multigroup structural equation modelling with mean structures. *British Journal of Mathematical and Statistical Psychology, 64*, 516–537.

Scheines, R., Spirtes, P., Glymour, C., Meek, C., & Richardson, T. (1998). The TETRAD Project: Constraint based aids to causal model specification. *Multivariate Behavioral Research, 33*, 65 – 117.

Tomarken, A. J., & Waller, N. G. (2005). Structural equation modeling: Strengths, limitations, and misconceptions. *Annual Review, 1*, 31-65. doi: 10.1146/annurev.clinpsy.1.102803.144239.

World Bank. (2012). World Development Indictors (WDI) and Global Development Finance (GDF). Retrieved from http://databank.worldbank.org/ddp/home.do.

Yu, C. H. (2011). Degrees of freedom. In Miodrag Lovric (Ed), *International Encyclopedia of Statistical Sciences* (pp.363-365). New York, NY: Springer.

Yu, C. H., DiGangi, S., & Jannasch-Pennell, A. (2012). A time-lag analysis of the relationships among PISA scores, scientific research publication, and economic performance. *Social Indicators Research, 107*, 317-330. doi: 10.1007/s11205-011-9850-5.

Yung, Y. F. (2008, March). *Structural equation modeling and path analysis using PROC TCALIS in SAS® 9.2*. Paper presented at SAS Global Forum, San Antonio, TX.

ACKNOWLEDGEMENTS

Special thanks to Mr. David Dunaetz for his valuable input to this paper.

CONTACT INFORMATION

Your comments and questions are valued and encouraged. Contact the author at:

Chong Ho Yu, Ph.Ds.
Department of Psychology
Azusa Pacific University
901 Alosta Ave.
Azusa, CA 91702
480-567-4782
cyu@apu.edu
http://www.creative-wisdom.com/computer/sas/sas.html