Structural Equation Model of Estimating Final Construction Cost of Road Tunnel

Kleopatra Petroutsatou
Department of Construction Engineering and Management, School of Civil Engineering, National Technical University of Athens
kpetrou@tee.gr

Sergios Lambropoulos
Department of Construction Engineering and Management, School of Civil Engineering, National Technical University of Athens

Abstract

In the preliminary design phase of technical projects, the selection of the best alternative solution requires reliable early estimates of construction cost. This is particularly true in the case of tunnels, where underground risks are involved and cost overruns are frequent. To estimate final construction cost of road tunnels in a more realistic and systematic manner a structural equation model (SEM) was developed. SEM is a systematic combination of confirmatory factor analysis, multiple regression analysis and path analysis. More importantly, SEM has an error variable that capitalizes on both measurement errors and structural errors by accurately reflecting the actual phenomena. SEM allows the user to visually depict the paths of how several variables affect the final construction cost of an underground project. A set of geotechnical parameters influencing the final construction cost was selected, based upon an extensive literature review and consultations with a number of experts. Detailed information regarding 49 km of constructed tunnels was gathered in an extensive database and exploited by SEM. "Goodness of Fit" of the model was tested for accuracy and robustness.

Keywords

Cost estimation, Tunnels, Structural equation model

1. Introduction

Social and financial needs for safe and rapid transportation of people and goods impose the construction of secure high speed motorways. In order to overcome problems associated with urban regions, environmentally sensitive areas or archeological sites the construction of tunnels becomes imperative.

It is very important to establish the early cost estimate of technical projects because affordability determines to a large extent the best alternative solution. This is particularly the case for tunnels where underground risks are involved.

This study aims to provide an organized and accurate model that estimates the final construction cost of road tunnels by utilizing Structural Equation Model (SEM). The data was gathered from an extensive database that was developed in previous work (Petroutsatou and Lambropoulos, 2007). The tunnels are constructed across the new motorway Egnatia Odos that crosses a large variation of geological conditions.

2. Literature Review

Various analytical methods have attempted to model complicated estimating processes in construction engineering and management. The widely used cost estimating methods in the preliminary stage of a project are: (i) Statistical methods based on multiple regression analysis (MR) and (ii) neural networks (NNs).

Although, MR techniques have proven successful, there is a significant flaw with their use; they ignore all the potential measurement errors of the observed variables (Du Toit M. and Du Toit S., 2001). The assumption that independent variables can be measured without error is not true most of the times (Molenaar *et al*., 2000). SEM can be thought as an extension to MR that deals with poorly measured independent variables.

SEM recognizes the measurement error and offers an alternate method for measuring prime variables of interest through the inclusion of latent and observed variables. The former refers to a hypothetical concept that cannot be directly observed nor measured; the latter is a substitute variable that can be measured in lieu of the latent variables.

SEM is a combination of factor analysis, MR, and path analysis. It has two components, a measurement and a structural model. The former incorporates a factor analysis that is concerned with how well latent variable are represented by observed variables. The latter represents MR and path analysis that models the relationships between latent variables and the final outcome (Kline, 2005). The variables of SEM are also classified in to an exogenous and endogenous variable depending on whether they influence, or are influenced by, others.

According to Bae (2005), SEM is able to: (i) identify a casual relationship between an independent variable and a dependent variable by taking the measurement error of the observed variables into consideration; (ii) model a concept that is difficult to directly measure or quantify explicitly; and (iii) represent indirect effects as well as the direct causal or correlation relationships diverse and hierarchical variables. Moreover, SEM makes it possible to visualize the complex relations through a graphical representation that shows the directional paths between the variables.

For instance, in a study of college performance, a researcher might want to determine the impact of the latent variable *Intelligence* on *Graduation Probability* (Molenaar *et al*., 2000). Since *Intelligence* itself is not directly observable, the researcher might collect a *Scholastic Aptitude Test (SAT)* score and *Grade Point Average (GPA)* in high school as a measure of *Intelligence*. Anyone of these variables alone has error associated with it and is not a perfect measure of the latent variable *Intelligence*. In a standard regression model, these errors associated with indirect measurements are not well addressed, thereby propagating the model's error (Myers, 1990).

Figure 1 shows this example of SEM, where *Intelligence* is marked in an oval and the observed variable are marked in a quadrangle and the errors shown in circles.

SEM modeling has developed in the construction engineering and management area, where complex phenomena and dynamic relationships are often involved. Molenaar *et al*., (2000) used SEM to predict the possibility of disputes at the early stage of a project. Mohamed (2003) used SEM to model the joint venture performance of overseas construction projects. Islam and Faniran (2005) modeled the impact of project conditions on project effectiveness through SEM applications. Wong and Cheung (2005) used SEM modeling to test the hypothesis that partners' trust is the most critical factor to partnering success.
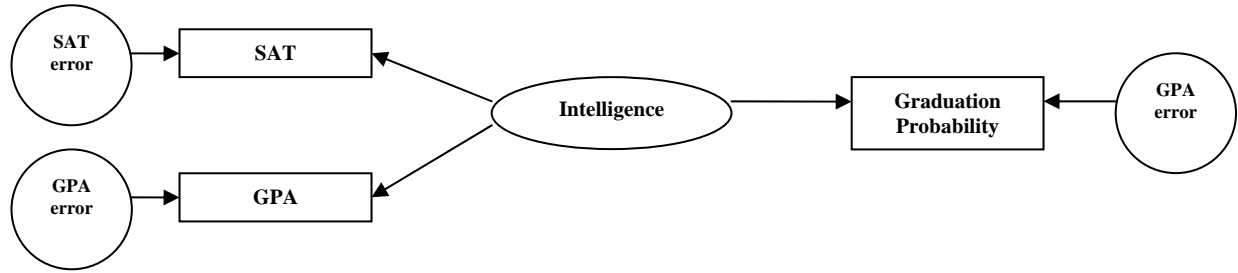
**Figure 1: Example SEM – Graduation Probability as Function of Intelligence
(Molenaar *et al*., 2000).**

3. Data Collection

A database was created (Petroutsatou and Lambropoulos, 2007) using information collected from 109 (N=109) different tunnel sections in 33 Egnatia twin bore tunnels with a total single bore length of 46 km. The tunnel excavation has a width of 11 metres and a height of 9 metres approximately, depending on the ground conditions encountered.

A structured questionnaire and a sample answer were sent to the site engineers. The collected data were scrutinized and tested in terms of accuracy and coherence by a series of site visits. The variables were selected from: (i) a literature review on the characterization of the rockmass, and (ii) interviews with academics, experienced designers and site engineers regarding the main geotechnical parameters that determine the support work quantities.

The database contained 149 different tunnel sections. Each tunnel section was described by 33 different elements concerning the geological and geomechanical conditions. 46 km of motorway was depicted in terms of underground conditions and geometry by 17 different elements per meter.

The data collection contained 14 variables representing the three latent variables that were used in this study to evaluate the project performance of a tunnel in terms of final cost variation. These original variables were reduced to 9 variables. The reduction of the number of variables was performed on the basis of a Pearson product-moment correlational analysis. Variables found to have significant associations with other variables within a latent variable construct were excluded from the model.

4. Research Model

This study uses the LISREL 8.80 program (LInear Structural RELation), which is widely used as the representative program of SEM (Kline, 2005). A base model was developed incorporating the latent variables with their corresponding measures into an initial structural equation model on the basis of theoretical expectations, past empirical findings and previous analysis (Petroutsatou and Lambropoulos, 2007). Model improvements were then performed over several iterations to arrive at a final model specification by using a modification of indices provided by the program. After refinements, the model that performed well with regard to both Goodness Of Fitness measure (GOF) and the theoretical expectations was selected as the final SEM model (Molenaar *et al*., 2000).

Finally, the authors developed the logical relationships among the 3 latent variables (one exogenous in green and two endogenous in yellow) and 8 observed variables, shown in Table 1, in the form of a path diagram shown in Figure 2, which represents the best performing model according to GOF results (Table 2).

Table 1: A Sample Table

Latent variables	Observed variables
GeoEnvr.: Geological Environment	H: height of overburden (m)
	GSI: geological strength index
Design	AREA: excavated section (m^2)
	SHOTCRET: shotcrete (m^3)
	STEELSET: steel sets (kgr)
	CONCRETE: concrete (m^3)
	STEEL: steel (kgr)
Project	TOTAL_CO: total cost of the excavation category (€, prices 2004)
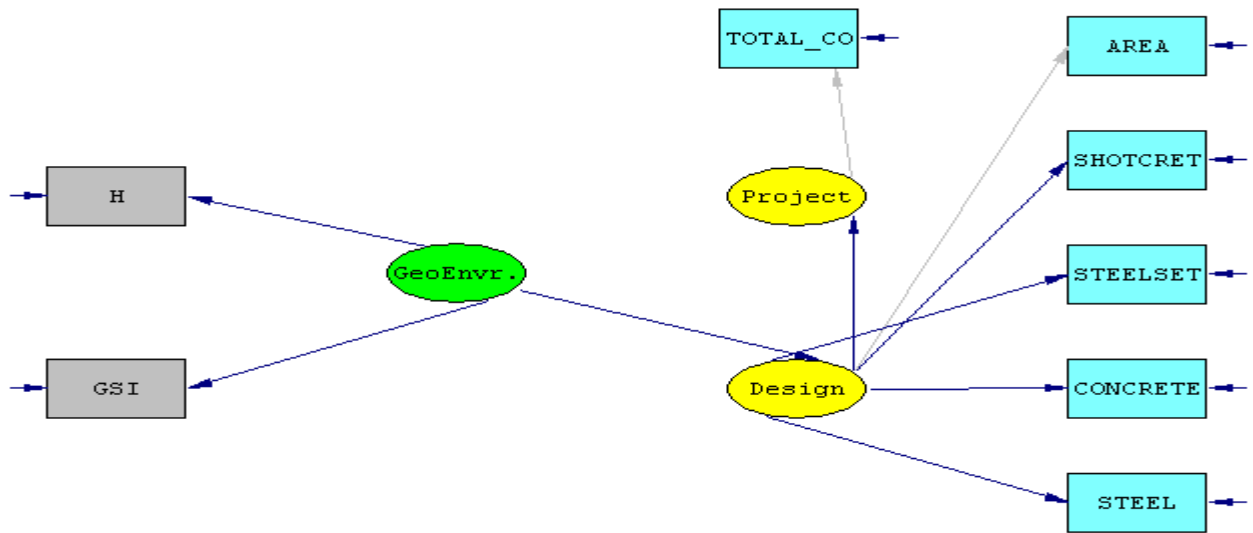
Figure 2: Conceptual Model of Project Performance

This model solidifies/transforms the conceptual model (Figure 2) into a prediction tool, revealing the quantitative impacts that variables have on tunnel final cost. The level of project performance in this study is measured by one observed variable, the total cost of the excavated category. The geological environment is represented by the height of overburden earth and the geological strength index. The design is represented by the five observed variables shown in Figure 2. Design has the most impact on project performance (coefficient = 0.83), followed by geological environment (indirect coefficient = -0.21).

The SEM model shows how the project performance is influenced by the various variables. The path diagram is based on causal or correlational relationships and graphically depicts how the latent variables interact with the observed variables to influence the cost variation of total cost.

5. Discussion of Results and Conclusions

As for the accuracy of the model, the goodness of fit index (GFI) is an apposite indicator to test how well data fits. It generally ranges from 0 (no fit) to 1 (perfect fit). In contrast with *R-squares* of the regression analysis, the GFI is rarely affected by a change in sample sizes or violation of multivariety normality. According to Kline (2005), the universally recommended level of GFI is 0.8 and above. The model developed in this study reflects the data properly with a GFI of 0.81, as shown in Table 2. In addition, the root-mean-square error of approximation (RMSEA) is estimated at 0.065 (Table 2), indicating a reasonable error of approximation.

Table 2: Analysis of SEM Results

Goodness-of-fit (GOF) measure	Final SEM model
Sample size	109
Goodness of Fit Index (GFI)	0.81
RMSEA	0.065
Normed Fit Index (NFI)	0.76

Conclusively, this work contributes to the identification of key variables and their interrelations that significantly affects the tunnel performance terms of its final cost.

Future research may further investigate the project performance in terms not only of total cost but also in terms of schedule programming and profit variation.

6. References

Bae, B. R. (2005). *LISREL Structural Equation Model: Understanding and Application,* Korea: Chunglam Publishing Company.

Du Toit, M., and Du Toit, S. (2001). *Interactive LISREL: User's Guide,* Scientific Software International Inc.

Islam, M. D. M., and Faniran, O. O. (2005). "Structural equation model of project planning effectiveness", *Construction Management and Economics,* Vol. 23, No. 2, pp. 215-223.

Kline, R. B. (2005). *Principles and Practice of Structural Equation Modeling,* 2nd edition, NY, The Guilford Publications, Inc.

Mohamed, S. (2003). "Performance in international construction joint ventures: Model perspective", *Journal of Construction Engineering and Management, ASCE,* Vol. 129, No. 6, pp. 619-626.

Molenaar, K., Washington, S., and Diekmann, J. (2000). "Structural equation model of construction contract dispute potential", *Journal of Construction Engineering and Management, ASCE,* Vol. 126, No. 4, pp. 268-277.

Myers, R. (1990). *Classical and Modern Regression with Applications,* 2nd edition, Belmont, California, Duxbury Press.

Petroutsatou, K., and Lambropoulos, S. (2007). "Early estimating of road tunnel construction cost", *Technica Chronica, Scientific Journal of Technical Chambers of Greece*, Vol. I, No. 1-2, pp. 99-110.

Wong, P. S. P., and Cheung, S. O. (2005). "Structural equation model of trust and partnering success". *Journal of Management in Engineering, ASCE,* Vol. 21, No. 2, pp. 70-80.