Finding Significant Factors to Affect Cost Contingency on Construction Projects Using ANOVA Statistical Method -Focused on Transportation Construction Projects in the US-

Sang Choon Lhee

Director, Gyp-Hyun ENC, Seoul, Korea

Abstract Risks, uncertainties, and associated cost overruns are critical problems for construction projects. Cost contingency is an important funding source for these unforeseen events and is included in the base estimate to help perform financially successful projects. In order to predict more accurate contingency, many empirical models using regression analysis and artificial neural network method have been proposed and showed its viability to minimize prediction errors. However, categorical factors on contingency cannot have been treated and thus considered in these empirical models since those models are able to treat only numerical factors. This paper identified potential factors on contingency in transportation construction projects and evaluated categorical factors using the one-way ANOVA statistical method. Among factors including project work type, delivery method type, contract agreement type, bid award type, letting type, and geographical location, two factors of project work type and contract agreement type were found to be statistically important on allocating cost contingency.

Keywords: Cost Contingency, Construction Project, Empirical Model, Analysis of Variance (ANOVA), Categorical Factor

1. INTRODUCTION

Construction projects, regardless of type and size, include many uncertainties and risks throughout all construction phases from startup to completion. As a result, many construction projects, including transportation projects, have historically experienced significant cost increases (Flyvbjerg et al. 2002; Molenaar 2005). Baccarini (2004) also noted that construction projects are notorious for overrunning budgets from these uncertainties and risks. In order to cope with these invisible and unforeseen uncertainties and risks against underestimating project costs and overrunning budgets, project participants including owners, contractors, and architects have added a cost element of contingency into their base estimates. Cost contingency is defined as an amount of money added to the base estimated amount to achieve a specific confidence level or allow for changes that experience shows (AACE 2000). Gunhan and Arditi (2007) categorized contingency into the following three types depending on the project phase and the party involved: (1) owner contingency, (2) contractor contingency,

and (3) designer contingency. Among them, owner contingency means the budget set aside in order to compensate uncertainties and risks from the owner's point of view during the construction phase although owners hope that contingency would not be needed during a project. In this paper, cost contingency is considered from the owner's perspective and usually refers to additional cost incurred by unforeseen change orders, compensable variations, and excusable delays.

Several contingency estimation models have been developed and proposed using different methodologies. Hollmann (2007) identified four common methods to estimate contingency: (1) expert judgment, (2) predetermined percentage guideline, (3) simulation analysis, and (4) parametric modeling. Among them, the most common method in the construction industry for a long time is to use an arbitrary percentage of the estimated construction cost or bid amount (Thompson and Perry 1992; Touran 2003). Modern estimating textbooks usually represent contingency as a fixed percentage reported to be around 5-10% of the base estimate (Smith and Bohn 1999). In addition, many empirical models for estimating contingency have been proposed using parametric modeling techniques such as statistical regression analysis and artificial neural networks (ANNs). Thal et al. (2010) used multiple linear regression method to develop a model to predict the amount of required contingency funds for U.S. Air Force construction projects. Moselhi et al. (1993) developed a decision-support system that helps contractors preparing competitive bids for building construction projects using artificial neural networks to estimate contingency. Chen and Hartman (2000) developed an ANN-based model to predict the contingency cost at the front-end

Corresponding Author: Sang Choon Lhee,
  Director, Gyp-Hyun ENC, Seoul, Korea
  Tel : +82-2-6203-8800 e-mail : lheesch@naver.com

stage of project development. Lhee et al. (2009) also proposed an ANN approach to predict owner contingency for transportation construction projects, especially asphalt resurfacing works (Lhee et al. 2012).

However, empirically-based models such as regression analysis and ANNs have a big weakness in applications that they do not have a capability of interpreting categorical input variables that cannot be put in some meaningful order (Flood and Issa 2010). In order to solve this problem, contingency estimation models using empirical modeling techniques should be developed using only numerical input variables to affect contingency as output variable, after separating each model type for categorical input variables. From this limitation on developing empirically-based models for estimating contingency that meaningless categorical variables cannot be treated, the purpose of this paper is to identify input variables that affect owner contingency in transportation construction projects and to examine the significance of categorical input variables using the one-way ANOVA (Analysis of Variance) statistical method. From this analysis to treat categorical variables before developing an empirically-based estimation models to use regression analysis or artificial neural network method, the models can estimate cost contingency as accurately as possible and project owners or sponsors like DOTs (Department of Transportation) can consider the effects of categorical input variables in the allocation of contingency in transportation construction projects.

2. IDENTIFICATION OF FACTORS TO AFFECT COST CONTINGENCY AND DATA COLLECTION

Before collecting data, potential input variables that might affect cost contingency on transportation construction projects should be identified and determined. In this study, potential factors were identified from the FHWA (Federal Highway Administration) Major Project Program Cost Estimating Guidance Manual (2007) and other modern estimating textbooks (Popescu et al. 2003).

According to the FHWA manual, input factors to affect cost contingency in transportation construction projects are as follows:

1. Project delivery method type: Design-Build (D/B) contracts on major construction projects have shown little increase from startup to final completion under a negotiated contract amount and therefore may require a smaller contingency due to many reductions in the number of construction claims from design errors and omissions (E&O).

2. Number of concurrent contracts and contract interfaces: On projects where multiple contracts are underway at the same time, close coordination of construction activities and schedules might be required. The potential for one contractor to impact on another contractor's activities is higher and may bring to additional delays or coordination costs during construction. Therefore, a higher contingency may be required on projects under multiple contracts.

3. Contractor proposed construction changes: Contracts include some specifications such as Value Engineering Change Proposal (VECP) to allow contractors to propose construction changes which result in benefits to contractors and owners. Contracts to limit the opportunity for contractors to make these changes may restrict the scope to decrease construction

costs once construction starts. An increased contingency may be appropriate in these situations.

4. Construction time: On projects with longer duration, there is a higher risk for impacts to the schedule and therefore contingency should be higher. In addition, construction scheduled in winter or rainy seasons should be appropriately accounted for contingency because there may be a higher risk in satisfying construction schedules from unforeseen weather delays.

5. Transportation Management Plans for work zones: Major transportation construction projects often have complex construction traffic controls and may have multiple construction contracts underway at the same time. The cost for implementing the Transportation Management Plan (TMP) for work zones must be included in the estimate. Costs may also include incident management, public information and communication efforts, transit demand management and improvements to the local area network for helping improve safety and traffic flow during construction.

6. Environmental impacts: Major construction projects go through an intensive NEPA (National Environmental Policy Act) process. Due to the size and complexity of major projects, there are often greater public and resource agency investigations during construction. This attention results in a greater likelihood that additional environmental mitigations may be required once construction begins.

7. Other factors: As potential factors on cost contingency, there are risks of encountering underground utilities and other obstructions, differing site conditions, contaminated sites, and multi-agency involvements.

In addition, Popescu et al. (2003) mentioned that the magnitude of cost contingency depends on the project contract agreement type, the construction work type, and project geographical location. Based on the FHWA manual and the Popescu's recommendation, potential input variables on contingency for transportation construction projects were summarized and categorized into three types as seen in Table 1.

Table 1. Summary and Categorization of Potential Input Variables (Lhee 2009)

Type	Potential input variable
Project factor	Project work type Project delivery method type Project contract agreement type Project bid award type Project letting (procurement) type Project geographical location Project duration Project amount Number of bidders Project site condition Possibility of construction changes
Market factor	Number of concurrent projects Project letting year
Other environmental / managerial factor	Environmental impact Transportation management plan

Among 15 potential factors, contingency related data were collected from 829 transportation construction projects which were sponsored and completed from 2004 to 2006 by the U.S. FDOT (Florida Department of Transportation) and retrieved by quarterly construction project time and cost reports and bidding tabulation documents. From quarterly time and cost reports, project information such as project delivery method type, project amount, duration, letting type, and geographical location, and project letting year were obtained. And project information such as project work type, contract agreement type, bid award type, geographical location, and the number of bidders were obtained from bidding tabulation documents for each project. Table 2 shows the classification of accessible input variables obtained from the FDOT construction project database depending on the type of variable.

Table 2. Classification of Accessible Input Variables from the FDOT Project Database

	Categorical variable	Numerical variable
Accessible input variables	Project work type Project delivery method type Project contract agreement type Project bid award type Project letting type Project geographical location	Project duration Project amount Number of bidders Project letting year

This study focuses on a statistical analysis on the effect of categorical input variables on cost contingency using the one-way ANOVA (Analysis of Variance) method since previous empirical models to estimate cost contingency have included numerical input variables on the models and treated them for improving the accuracy of the prediction.

3. DESCRIPTION OF ACCESSIBLE INPUT VARIABLES ON COLLECTED PROJECT DATA

(1) Project work type

In general, project work type in transportation construction projects can be represented as roadway, rehabilitation, structure, signage, signalization, lighting, landscape, and other construction activities. Uncertainties and risks will vary depending on the project work type. Normally, roadway, rehabilitation, and structural works include more uncertainties and risks than signing, signalization, lighting, and landscape works in perspective of complexity and variety. For this study, the project work type was categorized into the following five focusing groups: (1) asphalt resurfacing, (2) asphalt paving, (3) bridge work, (4) combination of bridge work and asphalt paving, and (5) other works. Among them, asphalt resurfacing and paving works account for approximately 60 percent of collected project data.

(2) Project delivery method type

Construction delivery method is defined as the set of relationships, roles, and responsibilities of project members and the sequence of activities. It varies on a project-to-project basis depending on project objectives. The FDOT has been using the following two delivery methods for transportation construction projects: Design-Bid-Build (D/B/B) and Design-Build (D/B) methods. Between two delivery methods, the FDOT uses the Design-Build method as innovative delivery one which combines design, construction, and even right-of-way services into a single contract in order to reduce costs and expedite schedule through speedy and coordinated communications between project members. The FDOT Design-Build Guidelines (2007) recommend strong considerations of the Design-Build delivery method on the following types of projects: (1) projects with an expedited schedule and a high possibility on early completion, (2) projects with a minimum right of way acquisition and utility relocation, (3) projects with a well-defined scope for all project members, (4) projects with room for innovation in the design/construction effort, (5) projects with a low risk of unforeseen conditions, and (6) projects with a low possibility of significant changes during all work phases. For the collected FDOT projects, the Design-Bid-Build delivery method was used as a conventional one.

(3) Project contract agreement type

As contract agreement method, the following three types have been generally used for construction projects: Lump sum contract, unit price contract, and cost-plus-fee contract. Among them, the FDOT has usually been using lump sum and unit price contract agreements for transportation construction projects. Specifically, the lump sum contract as innovative method has been used for reducing the costs of design and contract administration related with quantity calculation, verification and measurement on simple projects which have well-defined scope for all project members, a low risk of unforeseen conditions, and a low possibility for work changes during all design and construction phases. Between two contract agreement methods on collected transportation construction projects, the unit price contract method is more popular than lump sum contract one on FDOT projects.

(4) Project bid award type

As bid award type, the FDOT has been using the following three methods for construction projects: Lowest bid method, A+B (Cost+Time) bid method, and Bid Average Method (BAM). The A+B (Cost+Time) bid method enables contractors to determine a reasonable duration required for project completion and includes time bid items with a related cost of the completion duration in determining the lowest bid. In this method, the cost bid item is represented as standard cost and the time bid item is represented by multiplying construction days by predetermined daily road-user costs. In addition, the FDOT has been using the Bid Average Method in order to get contractors to bid a true and reasonable project cost in some 100% state funded or local projects, not federally funded projects. The lowest bid method was used for most of FDOT transportation construction projects and the Bid Average Method was used on only one project among the collected projects.

(5) Project letting type

FDOT projects can be divided into two letting types depending on the procurement agency: Central office letting (CO) and District office letting (DO). The FDOT projects are planned and executed by one main central office and 8 regional district offices. The central office is responsible for purchasing professional

services, contractual services, and commodities related with the state highway system. Each district office, a decentralized agency of the FDOT, is responsible for acquiring commodities, contractual services, road/bridge construction and maintenance, and professional services. Normally, central office letting projects sponsored by the FDOT main office are more large and complex in scope and higher in contract amount than district office letting projects. Between two project letting types, central office letting method accounts for about 1.8 times as many as district office letting one on the collected FDOT projects.

(6) Project geographical location

Geographical locations are important for executing construction projects. Projects near urban areas can easily obtain the supply of labor, materials, and equipments in appropriate ways. Transportation time for them can be also estimated with accuracy. Based on the definition of the United State Office of Management and Budget (U.S. OMB) as seen in Figure 1, the project geographical location was divided into urban and rural groups in this research. For the collected FDOT projects, most of projects were executed in the urban area.
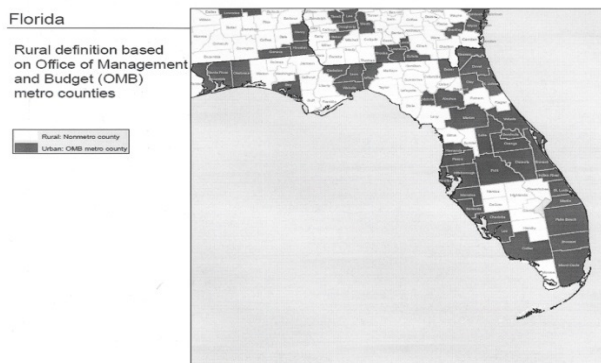

Figure 1. Definition of urban and rural area in the state of Florida (USDA 2000)

4. DETERMINATION ON FORM OF COST CONTINGENCY

Previous researches on empirical models to predict cost contingency used a desired contingency amount or rate as output variable for the prediction models. In ANN-based models, Moselhi et al. (1993) used optimum contingency rates as output variable for building construction projects and Chen and Hartman (2000) predicted total contingency amounts at the front-end stage of project development. Thal et al. (2010) predicted contingency amounts for Air Force construction projects on a developed multivariate linear regression model. However, comparisons of contingency amounts for construction projects in the statistical analysis such as hypothesis testing and ANOVA are meaningless because they are various depending on the base estimate or initial bid amount under the current FDOT contingency practice using the pre-determined percentage method based on original contract amounts. Therefore, comparisons of the contingency rate among projects are effective and meaningful to find the significance of input variables on contingency as output variable. In this study, cost contingency was defined as the cost item that can compensate

for all unforeseen work orders and related risks. Mak and Picken (2000) and Baccarini (2004) mentioned that contingency can be calculated from a comparison between the original predicted cost and the actual final cost. In this way, a desired contingency amount can be calculated from the difference between the initial contract amount and the final contract amount and a desired contingency rate as the target of the ANOVA statistical analysis in this study can be also calculated by dividing the desired contingency amount into the initial contract amount.

5. ONE-WAY ANALYSIS OF VARIANCE (ANOVA) METHOD

The hypothesis testing is used for comparing the means of two samples using the statistic tool of t-test. However, in order to test the means of more than two samples, the t-test cannot handle the comparison. In this case, the one-way analysis of variance (ANOVA) method can be used to compare many means of two or more samples using the statistic tool of F-test, a key element in the ANOVA. The null hypothesis in the ANOVA is that all sample means are equal, while the alternative hypothesis is that at least one sample mean is different. The ANOVA tests the null hypothesis with two estimates about population variance (i.e. between-sample variance and within-samples variance) and produces an F-statistic, the ratio of the between-sample variance to the within-sample variance. If the sample means are drawn from the same population, the between-sample variance should be lower than the within-sample variance following the central limit theorem. Therefore, a higher F-ratio means that all sample means are equal and fails to reject the null hypothesis. However, this statistical method can be used only for numerical data (Sall et al. 2005).

In this study, the one-way analysis of variance (ANOVA) method was used to analyze the influence of categorical input variables on cost contingency in transportation construction projects. The data were analyzed using the JMPIN statistical software. The ANOVA method compares the mean of the response (output) variables (i.e. desired contingency rate) for several groups as categories of a qualitative explanatory (input) variable. The software provides the F-ratio which is used to determine the significance of each input variable. All variables were tested at the 95% confidence level ($\alpha=0.05$) because the statistical analysis done within this range is considered to be acceptable in the construction industry (Hale et al. 2009). In the discussion of results, the p-value observed from the ANOVA is shown as a measure of significance level since the p-value is the probability of rejecting a null hypothesis. Since the ANOVA assumes a null hypothesis that the means of more than two samples are equal ($\mu1=\mu2=\cdots=\mu n$), the p-value must be less than or equal to 0.05 for the null hypothesis to be false. Given that the null hypothesis is true, the p-value represents the probability of observing a random sample that is at least as large as the observed sample. If the p-value is below 0.05, the difference in the means is considered to be statistically significant (Weistein 2007).

6. RESEARCH FINDINGS

Table 3 provides a summary of ANOVA statistical results for categorical factors to influence cost contingency. With respect

to contingency rate, the comparison among the five groups for the work type showed that there was significant difference (p-value=0.0003) at the chosen confidence level. Figure 2 shows representative box plots for contingency rate metric of the five groups by work type and a good example of the variation evident within each group. Horizontal lines within the box represent mean contingency rate metric values (respectively 0.057, 0.078, 0.099, 0.099, and 0.078 in bold and shading mark) presented in Table 3 and another horizontal line across the boxes represents gross mean contingency rate metric value of the five groups. For the contract agreement type, the difference was also significant in contingency rate (p-value=0.0002) from lump sum and unit price methods with mean contingency rate from lump sum and unit price groups as 0.056 and 0.082. Figure 3 shows the box plot for contingency rate metric by the contract agreement type. However,

there was no difference in contingency rate for the delivery type (p-value=0.0656) with gross mean contingency rate from design-bid-build and design-build methods as 0.071 and 0.046. There were no significant differences in contingency rate from the two specification groups for the bid ward type (p-value=0.8834) and the letting type (p-value=0.4680). For the bid award type, the mean contingency rates for lowest bid and cost+time bid methods were 0.071 and 0.069, respectively. Similarly, for the letting type, the mean contingency rates for central office letting and district office letting methods were 0.072 and 0.067, respectively. In addition, the comparison of mean contingency rates between urban and rural groups for the geographical location showed that there was no difference (p-value=0.4184). The mean contingency rates from urban and rural groups were 0.071 and 0.064, respectively.

Table 3. ANOVA Statistical Results for Categorical Factors to Influence Contingency Rate

| Categorical factor	Group	Sample size	Mean	Std. dev.	F-ratio	p-value	Reject null hypothesis?
Work type	Asphalt resurfacing	371	0.057	0.077	5.337	0.0003	Yes
	Asphalt paving	106	0.078	0.076			
	Bridge work	76	0.099	0.148			
	Bridge work + asphalt paving	51	0.099	0.089			
	Other works	178	0.078	0.108			
Delivery method type	Design-Bid-Build	779	0.071	0.096	3.398	0.0656	No
	Design-Build	50	0.046	0.070			
Contract agreement type	Lump sum	326	0.056	0.094	13.99	0.0002	Yes
	Unit price	458	0.082	0.095			
Bid award type	Lowest bid	750	0.071	0.096	0.021	0.8834	No
	Cost+Time bid	33	0.069	0.073			
Letting type	Central office letting	528	0.072	0.084	0.527	0.4680	No
	District office letting	301	0.067	0.111			
Geographical location	Urban	686	0.071	0.097	0.656	0.4184	No
	Rural	139	0.064	0.079			
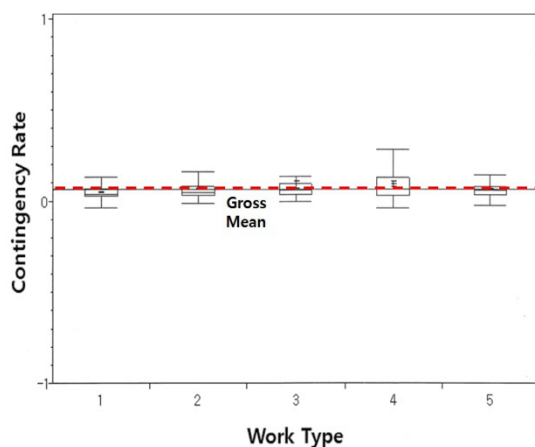
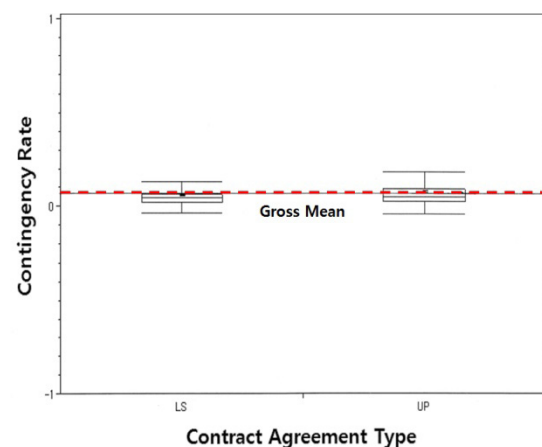Figure 2. Box plot of contingency rate for work type

Figure 3. Box plot of contingency rate for contract agreement type

7. SUMMARY AND CONCLUSIONS

This study was motivated by the need to determine which categorical input factors affect contingency since they cannot be directly put in empirical prediction models using regression analysis or ANN method. The study has collected contingency-related data in transportation construction projects and evaluated the effect of categorical factors which might influence on the cost contingency using the one-way ANOVA statistical method. Among six categorical factors including work type, delivery method type, contract agreement type, bid award type, letting type, and geographical location, factors of "work type" and "contract agreement type" was found to be statistically significant. From the comparison of contingency rate among groups for work type and contract agreement type, work type and contract agreement type was significantly different showing respectively 0.0003 and 0.0002 as p-value at the 95% confidence level. Therefore, construction owners or sponsors like departments of transportation (DOTs) should strongly consider effects of work type and contract agreement type in allocating contingency in their base estimates.

Future works should focus on identification and evaluation of categorical input factors on cost contingency on another construction type since the statistical result from this study was obtained only on transportation project data. And then, empirical models developed by separating each type for these effective categorical factors should be systematically planned and proposed for improving predictions of cost contingency.

REFERENCES

American Association of Cost Engineers (AACE)'s Risk Management Committee (2000). "AACE International's risk management dictionary." Cost Eng., 42(4), 28-31.

Baccarini, D. (2004). "Accuracy in estimating project cost construction contingency: A statistical analysis." The Int. Constr. Research Conf. of the Royal Institution of Chartered Surveyors, Headingley Cricket Club, Leeds. Australia.

Chen. D. and Hartman, F. T. (2000). "A neural network approach to risk assessment and contingency allocation." AACE International Transactions, RISK.07.01-RISK.07.06.

Federal Highway Administration (FHWA) (2007). Major project program cost estimating guidance, Washington, DC

Flood, I and Issa, R. R. A (2010). "Empirical modeling methodologies for construction." J. Constr. Eng. Manage., 136(1), 36-48.

Florida Department of Transportation (FDOT) (2007). Design-Build guidelines, Tallahassee, FL.

Flyvberg, B., Holm, M. S., and Buhl, S. (2002). "Underestimating costs in public works projects: Error or lie?" J. Amer. Plan. Assoc., 68(3), 279-295.

Gunhan, S. and Arditi, D. (2007). "Budgeting owner's construction contingency." J. Constr. Eng. Manage., 133(7), 491-497.

Hale, D. R., Shrestha, P. P., Gibson, G. E., and Migliaccio, G. C. (2009). "Empirical comparison of design/build and design/bid/build project delivery methods." J. Constr. Eng. Manage., 135(7), 579-587.

Hollmann, J. K. (2007). "The monte-carlo challenge: A better approach." AACE International Transactions, RISK.03.01-RISK.03.07.

Lhee, S. C., Issa, R. R., and Flood, I. (2009). "Predicting owner's contingency for transportation construction projects using intelligent computing techniques." Proc., ASCE Int. Workshop on Comput. in Civ. Eng., Austin, TX, 442-451.

Lhee, S. C. (2009). "Artificial neural network model to predict financial contingency for transportation construction projects" Ph.D. dissertation, University of Florida.

Lhee, S. C., Issa, R. R., and Flood, I. (2012). "Predicting of financial contingency for asphalt resurfacing projects using artificial neural networks." J. Constr. Eng. Manage., 138(1), 22-30.

Mak, S. and Picken, D. (2000). "Using risk analysis to determine construction project contingencies." J. Constr. Eng. Manage., 126(2), 130-136.

Molenaar, K. R. (2005). "Programmatic cost risk analysis for highway megaprojects." J. Constr. Eng. Manage., 131(3), 343-353.

Moselhi, O., Hegazy, T., and Fazio, P. (1993). "DBID: Analogy-based DSS for bidding in construction." J. Constr. Eng. Manage., 119(3), 466-479.

Popescu, C. M., Phaobunjong, K., and Ovararin, N. (2003). Estimating building costs, Marcel Dekker, New York.

Sall, J., Creighton, L., and Lehman, A. (2005). JMP start statistics, Third edition. Toronto, Canada: SAS Institute/Thomson.

Smith, G. R. and Bohn, C. M. (1999). "Small to medium contractor contingency and assumption of risk." J. Constr. Eng. Manage., 125(2), 101-108.

Thal, A. E., Cook, J. J., and White, E. D. (2010). " Estimation of cost contingency for Air Force construction projects." J. Constr. Eng. Manage., 136(11), 1181-1188.

Thompson, P. A. and Perry, J. G. (1992). Engineering construction risks: A guide to project risk analysis and risk management, Thomas Telford, London. U.K.

Touran, A. (2003). "Calculation of contingency in construction projects." IEEE Trans. Eng. Manage., 50(2), 135-140.

U.S. Department of Agriculture (USDA) Economic Research Services (2000), Rural Definitions: State-Level Map (Florida), Washington, DC.

Weistein, F. (2007). Wolfram math world, the web's most extensive mathematics resources