Construction projects usually involve high investments. It is, therefore, a risky adventure for companies as actual costs of construction projects nearly always exceed the planned scenario. This is due to the various risks and the large uncertainty existing within this industry. Determination and quantification of risks and their impact on project costs within the construction industry is described to be one of the most difficult areas. This paper analyses how the cost of construction projects can be estimated using Monte Carlo Simulation. It investigates if the different cost elements in a construction project follow a specific probability distribution. The research examines the effect of correlation between different project costs on the result of the Monte Carlo Simulation. The paper finds out that Monte Carlo Simulation can be a helpful tool for risk managers and can be used for cost estimation of construction projects. The research has shown that cost distributions are positively skewed and cost elements seem to have some interdependent relationships.

Key words: risk management, Monte Carlo Simulation, construction, probability distribution

Introduction
Many construction projects are undertaken in a complex and continuous changing environment. Systems and approaches are developed
by theoreticians and used by practitioners to cope with the new challenges. While for some industries these systems have a sufficient number of mathematical models for risk analyses, the construction industry suffers from underdevelopment (Flanagan and Norman 1993). This is partially due to the non-homogeneous and non-serial character and the high dependence of the success of the projects linked to the skills of the individual project manager. Furthermore, the increased national and international competition forces the constructors to focus on their core competence. These effects result in an increasing degree of outsourcing and a reallocation of the risks related to costs, time schedule and quality (Girmscheid and Busch 2014).

According to a survey of the Philipp Holzmann AG, 41% of the losses of construction projects is related to miscalculations in the pre-contract phase and 22% to project risks. 30% of the costs incurs during the construction phase and only 7% is related to force majeure. Philipp Holzmann AG could have increased the margin by 3% in avoiding 10% of the poorest contracts (Linden 1999).

Smith (1999) finds out that expenditure on the appraisal of major engineering projects represent only 10% of the capital costs of the project. However, during this period 80% of the total project costs are frozen. This shows how important the identification of the major risks and estimation of the costs at the beginning of a construction project are.

In 2008, the Boston Globe pointed out that the project ‘Boston’s Big Dig’ ended up costing almost $22 Billion vs. a budget of $2.6 Billion (Murphy 2008). The German Airport Berlin-Brandenburg showed in June 2014 a cost overrun of more than 150% to an amount of 5.4 Billion € (see http://www.flughafen-berlin-kosten.de/). These are only two examples of a huge number of miscalculated projects with cost overruns showing that an adequate cost calculation is more than needed. Such a cost calculation has to consider beside the basic costs also contingencies, which represent the risk of the project. For this, a well-implemented and complete risk management system with mathematical models for the risk analysis is needed. This is not easy, as the practitioners of the construction industry believe that the success of the project is highly dependent on the experience of the project manager gained over many years. They believe that experience cannot be easily transferred to mathematical models (Ashworth 1987).

Within risk management process the risk analysis is seen as the most difficult component, but it is also the most useful (Touran,
Yang, and Lowe 2011). This paper is focused on finding out if the Monte Carlo Simulation can be used to improve risk analysis and hence lead to a better estimation of costs in construction projects.

**Literature Review**

Every venture bears a risk. Therefore, it is important to understand how a risk is defined and what its sources are. In general, a risk can be defined as any uncertainties that, if it occurs, would affect one or more objectives (Hilson 2004). Hence, risk bears threat and opportunities. Usually six types of risks can be defined for the construction area (Girmscheid and Busch 2014):

- legal risks
- scheduling risks
- technical risks
- financial risks
- management risks
- environmental risks

However, a common use and interpretation of the risk types does not exist in the literature. Not all six types of risks may be important for a specific construction project. The dimension of the risks will be defined by factors as project size, environment, skills and experience of the employees, financial factors, technical complexity of the project, etc. Beside the basic costs of a project, which mainly consist of design, production (cost for labour and construction material) and installation, the total costs of a construction project and hence the success of the project is affected by risks.

An overview of the published work on the topic risk and the valuation of construction projects were performed by Touran, Yang and Lowe (2011). The considered literature mainly describes risk models based on the estimation on probabilities and their effect. They remark that a sufficient database is not given for the used stochastic models to analyze the cost and time table risks. Detailed remarks of mathematical evidence according to this statement are not given.

In a study performed 1992 Touran and Wiser (1992) used information from 1,014 low-rise buildings in the us The costs were broken down to 15 different items. After a Test of Goodness of Fit on each cost item, the lognormal distribution was concluded to fit best. The results were used to perform the Monte Carlo Simulation, once with assuming independence of the data and then with recognizing correlations.
A literature review of Baccarini (2005) came to the result that traditional percentage is the most commonly used estimating method in practice for considering project cost contingencies. However, other methods gained more and more interest, of which one is the Monte Carlo Simulation.

Wall (1997) analyzed a number of 216 new built offices from the UK. After the Test of Goodness of Fit was performed, the beta and lognormal distributions were used for the Monte Carlo Simulation. Furthermore, the author concluded that correlations between the cost items have to be taken into account. Ignoring the correlation is more intense than the choice of the distribution, lognormal or beta. Previous studies agree that by considering correlation in simulating and analyzing the risk results in a better estimation of variance of the distribution of total costs of construction projects (Chou 2011; Yang 2005; Flanagan and Norman 1993).

Hollauf (n.d.) reviewed construction projects in the UK. Data for a sample size of 58 construction objects were analyzed. The author found out that a dependency between the different cost elements exists and the correlation between these needs to be taken into account when performing Monte Carlo Simulation.

The German authors Girmscheid and Busch (2007) recommend the Monte Carlo Simulation among others for quantifying risk. In using the Monte Carlo Simulation, the authors agree that experts have to define for each risk a minimum, maximum and most reliable outcome and the corresponding probability. The resulting risk has to be considered as contingency in the general costs of the project.

Some authors define the venture during the construction period as uncertainties and not as risk. This is because the venture is based on subjective estimation and not on statistical investigation. The use of probabilities based on subjective assumptions lead to misinterpretations. Loizou and French (2012) make some critical notices to the use of the Monte Carlo Simulation in the construction industry as for every event subjective probabilities have to be designed. Statistical data from the past is often not available or statistical not significant.

Summarizing the literature, it can be said that for construction projects the analysis of risk and its potential impact is proposed. However, the conditions to determine the input parameters are rarely discussed. The critic of the literature focuses on subjective assumptions for probabilities, which can lead to misinterpretations. An analysis of historical data and their use in the context does exist only rudimental.
Methodology

During the proposal stage, a feasibility study is usually initiated without knowing the exact design and demands of the client. Due to the high risk within this business and to prevent cost overruns, it is common to add a reserve amount to the project costs, the so-called contingency. The calculation methods for such contingencies can be divided into three main categories: deterministic methods, probabilistic methods and modern mathematical methods (Bakhshi and Touran 2014).

Current practice considers risks of the project such as changes in design and project by applying a contingency allowance based on deterministic methods or single point estimation. These methods are easy to handle without demanding a high knowledge of statistics, but as the conditions are not stable, the utility of this approach is reduced. It is highly recommended to use a range estimation rather than single point estimation. This way the variation in the outcomes is reflected (Elkjær 2000). The two methods that can be used to analyze risk in the estimation of project-outrun costs are sensitivity analysis and probabilistic risk analysis (Tan and Makwasha 2010).

These approaches require a big range of data. However, historical data are limited especially in the construction industry. The problem could be solved by using simulations like Monte Carlo Simulation. Simulation based cost analysis requires two sets of data inputs which are, the marginal distribution of the individual cost elements and the correlation matrix consisting of the correlation coefficients between the different pairs of cost elements. Both sets of inputs can be estimated in two ways, (1) using historical data from past projects, (2) subjective judgment or using experience and intuition (Yang 2005).

This research tries to find out how Monte Carlo Simulation can be useful to estimate the costs and determine the contingency for the project, based on historical data. The approach follows the next five steps:

- Collection of historical data
- Definition of the Total Construction Costs (TCC)
- Test of Goodness of Fit
- Determination of correlation
- Monte Carlo Simulation

The used historical data are in accordance with the cost breakdown structure as applied by the bki (Baukosteninformationszen-
The data used in this research was obtained from the Kostenplaner 17 CD. The cost structure follows the DIN 276-1:2008-12. The costs of a building are class-divided into costs for the land, on-site infrastructure works, building construction, external areas, equipment and artwork and incidental costs. This research focuses on the building costs consisting of costs for the building pit, foundation, exterior walls, interior walls, ceiling, roof, fittings and other measures for construction.

Cost data of totally 75 administrative buildings in Germany are analyzed. The buildings were finished between 1976 and 2013. The gross floor area of the buildings is between 269 and 25,134 square meter. 24% of the buildings have a gross floor area of less than 1,000 m² and 33% higher than 5,000 m². While 84% of the buildings is from the private sector, the rest is from public sector.

The sample was assumed large enough to minimize the sampling error that could occur in such studies and is considered a good representation of the population. The sample group, was chosen as it had more data available than other groups. By presenting the data in cost per square meter, the problem of project size or scale is eliminated.

The tcc are defined to be the sum of the elemental costs per square meter. The gross floor space of the buildings is used as the common factor for the different cost elements. This step is followed by the definition of the probability distributions for the cost elements. This will be done by performing the Anderson-Darling Test.

The determination of the correlation between the cost elements is very important. If correlations are ignored, this might result in a significant underestimation of the costs for the job. For the correlation matrix, the Spearman rank correlation coefficients will be calculated. The correlation matrix will then be tested for its feasibility as appropriate (Yang 2005).

Using the results, Monte Carlo Simulation will be performed with the help of the Software Crystal Ball®. Each simulation generates a random set of possible values for each of the cost elements according to the specified marginal distributions. Two sets of 100,000 runs will be performed; one incorporating correlated data and the other assuming independence of elemental cost data.

**Monte Carlo Simulation**

This paper researches in which way historical data of construction projects follow a certain distribution and if these data can be used to model the possible future costs of a project.
### Table 1  Descriptive Statistics for Elemental Cost Data

<table>
<thead>
<tr>
<th>Item</th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
<th>(5)</th>
<th>(6)</th>
<th>(7)</th>
<th>(8)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean</td>
<td>29.68</td>
<td>259.14</td>
<td>472.00</td>
<td>258.94</td>
<td>296.60</td>
<td>330.99</td>
<td>26.20</td>
<td>45.76</td>
</tr>
<tr>
<td>Median</td>
<td>20.15</td>
<td>256.10</td>
<td>467.57</td>
<td>227.91</td>
<td>281.76</td>
<td>292.48</td>
<td>14.86</td>
<td>39.03</td>
</tr>
<tr>
<td>Variance</td>
<td>842.78</td>
<td>10214.52</td>
<td>27433.61</td>
<td>20277.98</td>
<td>7786.95</td>
<td>14984.88</td>
<td>1009.33</td>
<td>811.85</td>
</tr>
<tr>
<td>Std. var.</td>
<td>29.03</td>
<td>101.07</td>
<td>165.63</td>
<td>142.40</td>
<td>88.24</td>
<td>122.41</td>
<td>31.77</td>
<td>28.49</td>
</tr>
<tr>
<td>Min</td>
<td>5.59</td>
<td>11.02</td>
<td>170.42</td>
<td>100.52</td>
<td>157.11</td>
<td>148.29</td>
<td>0.17</td>
<td>10.28</td>
</tr>
<tr>
<td>Max</td>
<td>148.73</td>
<td>686.67</td>
<td>1,040.81</td>
<td>1,249.99</td>
<td>616.62</td>
<td>674.71</td>
<td>145.38</td>
<td>156.30</td>
</tr>
<tr>
<td>Range</td>
<td>143.14</td>
<td>675.65</td>
<td>870.39</td>
<td>1,149.47</td>
<td>459.51</td>
<td>526.42</td>
<td>145.21</td>
<td>146.02</td>
</tr>
<tr>
<td>1st quart.</td>
<td>13.08</td>
<td>194.75</td>
<td>326.17</td>
<td>185.15</td>
<td>239.38</td>
<td>247.41</td>
<td>3.38</td>
<td>25.33</td>
</tr>
<tr>
<td>3rd quart.</td>
<td>29.72</td>
<td>319.13</td>
<td>598.71</td>
<td>300.51</td>
<td>353.39</td>
<td>406.48</td>
<td>4.01</td>
<td>60.71</td>
</tr>
<tr>
<td>Skew</td>
<td>2.44</td>
<td>1.15</td>
<td>0.60</td>
<td>4.78</td>
<td>1.10</td>
<td>0.84</td>
<td>1.99</td>
<td>1.41</td>
</tr>
<tr>
<td>Kurtosis</td>
<td>6.26</td>
<td>3.84</td>
<td>0.53</td>
<td>31.66</td>
<td>2.38</td>
<td>0.11</td>
<td>4.01</td>
<td>2.60</td>
</tr>
</tbody>
</table>

**Notes**  Column headings are as follows: (1) building pit, (2) foundation, (3) exterior walls, (4) interior walls, (5) ceiling, (6) roof, (7) fittings, (8) others.

#### Test of Goodness of Fit

The fitting of the distribution to each cost element was done by using the Software Crystal Ball®. The Anderson-Darling test was used to decide on the best fit for each element. The Anderson-Darling Test measures the distance between the hypothesized distribution \( F \) and the empirical cumulative distribution function \( F_n \) (Anderson and Darling 1952).

\[
A = \int_{-\infty}^{\infty} \frac{(F_n(x) - F(x))^2}{F(x)(1 - F(x))} dF(x). \tag{1}
\]

Table 1 gives an overview of the statistics of the data. For all cost components, the best fitting distribution is the lognormal. It can be noted that the marginal distributions of the cost elements are all positive skewed and all the cost components have a mean, which is greater than the median. This is consistent with past researches (Wall 1997). The exterior walls have the highest costs per square meter with the highest variance. The positive skewness of the distributions does not really surprise, as construction projects rather tend to have costs overruns then lower costs as foreseen.

The tails of the distribution of the cost elements are longer to the right side, which suggests that a major part of the costs fall below the average, but a few extreme cases exceed the average. These extreme cases are subject to risk management and need to be taken into consideration. They can lead to project overruns. The tail probabilities have to be studied to set up the contingency for a project.
Table 2: Anderson-Darling Goodness of Fit Test

<table>
<thead>
<tr>
<th>Item (1) building pit</th>
<th>(2) foundation</th>
<th>(3) exterior walls</th>
<th>(4) interior walls</th>
<th>(5) ceiling</th>
<th>(6) roof</th>
<th>(7) fittings</th>
<th>(8) others</th>
</tr>
</thead>
<tbody>
<tr>
<td>Distribution</td>
<td>L</td>
<td>L</td>
<td>W</td>
<td>L</td>
<td>L</td>
<td>L</td>
<td>L</td>
</tr>
<tr>
<td>A–D</td>
<td>0.46</td>
<td>0.35</td>
<td>0.46</td>
<td>0.41</td>
<td>0.32</td>
<td>0.28</td>
<td>0.57</td>
</tr>
</tbody>
</table>

*Notes: Column headings are as follows: (1) building pit, (2) foundation, (3) exterior walls, (4) interior walls, (5) ceiling, (6) roof, (7) fittings, (8) others. L – lognormal, W – weibull, α = 95%.*

Half of the distributions have a kurtosis value higher than three. This indicates them being more peaked or having taller peaks compared to the normal distribution. This indicates the portion of extreme deviations from the mean value being high. This is mainly for the interior walls, while exterior walls are highly platykurtic with a value of 0.53.

It is crucial to specify the probability distribution of the cost elements. The quality of the results for the better-fit test increases with the available number of data. Previous studies assume the beta, uniform, triangular and lognormal distributions to fit best to the cost data. For historical data, researches suggest the lognormal distribution to fit the best (Touran and Wiser 1992; Wall 1997) This research is in line with past results and finds out that for almost all the given data the lognormal distribution is the best fitting.

Table 2 gives an overview of the results. For the exterior walls, the Weibull distribution fits the best. Nevertheless, the lognormal distribution for the exterior walls has also a good A–D value with 0.5789.

**Correlations of the Cost Elements**

When running a simulation-based cost analysis correlations must be considered if they are significant (Yang 2005). Ignoring correlations might result in a significant underestimation of the costs for the job, and this becomes even more significant if we consider a portfolio of different jobs (Bakhshi and Touran 2012). The rank (Spearman) correlation coefficient was used to reflect the degree of relation between the different cost elements. The advantage of the Spearman correlation coefficient is its use in case of a non-linear relationship between the variables and if both populations are not normally distributed (Yang 2005). The Spearman correlation coefficient is defined by following equation:

\[
 r_{xy} = 1 - \frac{6 \sum_{i=1}^{n} d_i^2}{n(n^2 - 1)},
\]  

(2)
where \( d \) is the difference between the ranks of the corresponding \( x \) and \( y \). The coefficients can range between \(-1\) and \(+1\). A coefficient of \(+1\) shows a perfect positive and \(-1\) a perfect negative relationship. Analysis of the data shows a positive correlation between the different cost elements. This is also in line with past research (Wall 1997).

To find out the significance of the correlation, a null hypothesis was tested against an alternative hypothesis. To test the spearman rank order correlation coefficients for significance at a level of 5%, the \( Z \) score test was performed.

- Null hypothesis: \( C_{ij} = 0 \)
- Research hypothesis: \( C_{ij} \neq 0 \)

\[
z_{ij} = C_{ij} \sqrt{N - 1},
\]

where \( C \) is the correlation between the different cost elements and \( N \) is the sample size.

The statistics transform the correlation coefficients to \( z \) scores on the standard normal probability distribution. The test statistic \( z \) is normally distributed for \( N > 30 \) and therefore can be compared to the critical values \( z \) of the standard normal distribution. To test if the correlation coefficient is significantly different from zero the above test statistics are compared to the 1.96 critical value of \( z \) at the 5% level of significance. Table 4 reflects the outcome.

The \( Z \) score test shows a significant correlation between some of the cost elements. This is consistent with past research (Yang 2005). The correlation is significant mainly between the costs for the main structures of a house: ceiling, walls and the roof. For 17 out of 28 of the coefficients, the critical value of 1.96 is exceeded. The use of a
Claudius A. Peleskei et al.

**Table 4  Z Score for the Correlations of the Cost Elements**

<table>
<thead>
<tr>
<th>Item</th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
<th>(5)</th>
<th>(6)</th>
<th>(7)</th>
<th>(8)</th>
</tr>
</thead>
<tbody>
<tr>
<td>(1)</td>
<td>1.0000</td>
<td>1.7514</td>
<td>1.1808</td>
<td>0.7463</td>
<td>1.4470</td>
<td>1.7288</td>
<td>1.1385</td>
<td>1.0450</td>
</tr>
<tr>
<td>(2)</td>
<td>1.0000</td>
<td>2.3852</td>
<td>2.1177</td>
<td>2.4586</td>
<td>2.1369</td>
<td>0.1279</td>
<td>2.9094</td>
<td></td>
</tr>
<tr>
<td>(3)</td>
<td>1.0000</td>
<td>4.9113</td>
<td>4.2359</td>
<td>4.8355</td>
<td>1.5024</td>
<td>3.6929</td>
<td></td>
<td></td>
</tr>
<tr>
<td>(4)</td>
<td>1.0000</td>
<td>3.9039</td>
<td>4.4520</td>
<td>2.5395</td>
<td>2.7676</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(5)</td>
<td>1.0000</td>
<td>4.8659</td>
<td>3.5707</td>
<td>2.5200</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(6)</td>
<td>1.0000</td>
<td>1.2805</td>
<td>3.1429</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(7)</td>
<td>1.0000</td>
<td>-0.4438</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(8)</td>
<td>1.0000</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Notes**  Column headings are as follows: (1) building pit, (2) foundation, (3) exterior walls, (4) interior walls, (5) ceiling, (6) roof, (7) fittings, (8) others.

correlation matrix in the Monte Carlo Simulation has to be feasible. This restricts the matrix to be positively semi definite, which means that the eigenvalue of the matrix must be non–negative (Yang 2005). A further discussion is not required as in the given situation the used correlations are all positive.

**Simulation of the Total Costs of Construction**

For modelling, the distribution of the $\tau_{cc}$ the Monte Carlo Simulation was applied by using Crystal Ball®. Monte Carlo Simulation generates samples $\{X^{(r)}_i\}_{r=1}^R$ from a given probability distribution $P(x)$. For each cost element, the best fitting probability distribution for its historical data was used. The details of the distribution are described in the above tables one and two. If we talk about simulation, we talk about generating a sample of random numbers. These numbers can be out of a range between 0 and 1. Monte Carlo is used to solve a mathematical or statistical problem. For example, when we throw darts on a figure and determine the relation of the hits to the missed darts. A Monte Carlo Simulation uses the random sampling of an experiment to determine the properties of some phenomenon (Sawilowsky 2003).

It was found out that the distributions of the $\tau_{cc}$ are positively screwed and show a heavy tail to the right. This is not surprising and again in line with past research results (Yang 2005). Furthermore, it was shown that the distribution without correlation of the cost elements has a lower range of possible outcomes. This means that the $\tau_{cc}$ without correlation underestimates the cost risk of a project. It points out that correlation needs to be accounted for when costs are estimated. The potential of running low costs, but also running losses is bigger because correlated construction factors add up each
TABLE 5  Statistics for the Two τcc Distributions

<table>
<thead>
<tr>
<th>Statistics</th>
<th>τcc without correlation</th>
<th>τcc with correlation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Trials</td>
<td>100000</td>
<td>100000</td>
</tr>
<tr>
<td>Mean</td>
<td>962.61</td>
<td>962.67</td>
</tr>
<tr>
<td>Median</td>
<td>946.42</td>
<td>927.71</td>
</tr>
<tr>
<td>Mode</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>Standard deviation</td>
<td>178.58</td>
<td>280.07</td>
</tr>
<tr>
<td>Variance</td>
<td>31.891.93</td>
<td>78.441.83</td>
</tr>
<tr>
<td>Skewness</td>
<td>1.51</td>
<td>1.45</td>
</tr>
<tr>
<td>Kurtosis</td>
<td>17.17</td>
<td>13.46</td>
</tr>
<tr>
<td>Coefficient of variability</td>
<td>0.19</td>
<td>0.29</td>
</tr>
<tr>
<td>Minimum</td>
<td>456.70</td>
<td>322.96</td>
</tr>
<tr>
<td>Maximum</td>
<td>5.154.76</td>
<td>7.789.76</td>
</tr>
<tr>
<td>Range width</td>
<td>4.698.06</td>
<td>7.466.80</td>
</tr>
<tr>
<td>Mean standard error</td>
<td>0.56</td>
<td>0.89</td>
</tr>
</tbody>
</table>

other. Further, it has to be mentioned that the distribution of the τcc without correlation has a greater peak, while both distributions are highly leptokurtic.

The results of the Monte Carlo Simulation can be used to get a better understanding of the range of the costs. Constructors can make visible which effect risks could have on the costs of a project and how the effect can increase or decrease because of correlation between the cost elements. Furthermore it is possible to define which the minimum costs of the project are if a certain probability is demanded. For example the project costs will not exceed 1,400 €/m² with a probability of 90%. The results are shown in table 3.

Conclusion

The research investigated how Monte Carlo Simulation can be used for cost estimation in construction projects. While the major part of past studies conclude that the estimation of the distribution for the cost elements has to be done subjective by experts, this paper analyses if historical data can be used for Monte Carlo Simulation. Furthermore, it analyses if a significant correlation between the input factors of the Monte Carlo Simulation exists.

It was found out that historical data could be used for a Monte Carlo Simulation to give project manager an idea of the variation in costs. It can be implemented into the risk management process to take better decision on the best mitigation strategy. The average cost of the selected office buildings used is about 962 €/m² per square meter, while ignoring quality, technology, location and other factors.
It was shown that the average value could be exceeded by a very large amount.

More than half of the correlations between the cost elements were significant at a coefficient level of 95%. The resulting two probability distributions show that the consideration of the interdependency is important in the risk analysis and must be incorporated in the estimation of the total costs. Ignorance of the correlation might lead to an underestimation of the variance of the project costs. This can lead to inadequate contingencies set up.

It was shown that the lognormal has the best fit compared to other distributions on most of the cost elements. The \( \text{tcc} \) distribution is likely to be lognormal itself due to self-replication property of the lognormal distribution. The \( \text{tcc} \) distribution is heavily reliant on the marginal distributions of the cost elements, which are dependent on the data, used.

The research did not consider the price variance of the buildings resulting from the type of the buildings, quality or location. This might have contributed to the large variation and outliers in the data. A further refinement might result in a higher accuracy of the estimation but this will result in less historical data.

As mentioned at the beginning of this paper, in particular for the construction industry the selection of the projects is high sensitive. Companies within this business are characterized by a small number of high volume projects. The projects have a realization time of some months or even years. These facts imply a high volume of investments for a longer period. The investments are related not only to the construction materials, but also to production facilities, etc. As a result, the economic success and future of a company is high dependent on the success of singular projects. How important the right choice of projects are, shows the example of the Philipp Holzmann AG. Philipp Holzmann AG could have increased the margin by 3% in avoiding 10% of the poorest contracts (Linden 1999).

Companies need to include their experience gained from past projects accurate into their project calculation for new tenders. This way it is easier to offer a price to the client that covers the costs/risks and ensures the company’s future. Finding the right tender price is decisive in competitiveness. A low price would not necessarily cover all risks and a high price would result in losing a tender to a better-prepared competitor. The current practice considers the risks of the project mainly by using single point estimation. These methods are easy to handle but the utility is limited (Elkjaer 2000). The use of the Monte Carlo Simulation by using past data as shown in this paper
reflects the variation in the outcomes. Project manager get a visible range of the possible outcomes of the projects, which ensures a better decision.

References


This paper is published under the terms of the Attribution-NonCommercial-NoDerivatives 4.0 International (cc by-nc-nd 4.0) License (http://creativecommons.org/licenses/by-nc-nd/4.0/).