Sensitivity Analysis for Cash Flow Simulation Based Real Option Valuation

Tero Haahtela
Aalto University, School of Science and Technology
P.O. Box 15500, FI-00076 Aalto, Finland
+358 50 577 1690 tero.haahtela@aalto.fi

Abstract

Sensitivity analysis identifies the critical aspects of the investment model that affect model output uncertainty. Common sensitivity analysis on options considers how the solution changes as a result of change in one of the key parameters (underlying asset value, volatility, exercise price, interest rate, time to maturity, dividends). In case of cash flow simulation based real options, these are mostly indirect variables that are computed based on the primary input variables - demand, unit selling price, and unit costs - in the cash flow calculation. The method presented in this paper combines the uncertainties in the underlying asset together by simulation, and then uses regression analysis approach for estimating how a change in a primary variable or variables affects simultaneously the stochastic process parameters of expected value and volatility during different time periods. As a result, the presented method shows how a change in a primary variable affects project value with real options. This information is also essential as it shows decision makers which uncertainties to follow, and thus this mitigates the common black box syndrome of consolidated cash flow simulation based methods.

Keywords: real options, sensitivity analysis, cash flow simulation, volatility

JEL Classification: G31, G13
1. Introduction

Sensitivity analysis is essential in order to identify the critical aspects of the investment model that affect model output uncertainty. The purpose in sensitivity analysis is to determine which input parameters are important and contribute most to the output variability (Helton, 2000). Based on this relative importance, sensitivity analysis provides guidance for further data collection and modeling (Manache & Melching 2008). Sensitivity analysis reveals under which conditions the model or with which parameter values the solution is robust, i.e. under what circumstances the model is insensitive to the changes in input parameters. It helps in defining under what circumstances the change in operational activities is required, which in case of real options is required in order to be able to decide on option exercise timely. Sensitivity analysis may also reveal what is the accuracy of the final model and the degree of uncertainty (Saltelli et al., 2004). This may be related to data accuracy and model uncertainty, and this uncertainty is reflected both as ambiguity and volatility uncertainty.

Traditional sensitivity analysis on options considers how the solution changes as a result of change in one of the parameters of underlying asset value (delta), volatility (vega), exercise price, time to maturity (theta), and interest (rho). With financial options, these aspects can be computed using well-known equations and available market information, allowing precise state-space structure for different underlying asset values for each time period. However, in case of real options analysis, the parameters affecting the Greeks directly are not the direct parameters that are changing but rather indirect or secondary parameters, for example volatility, that are derived from the actual changing primary cash flow calculation input parameters. The parameters that truly change are those related to cash flow model, e.g. product demand, sales unit price, and variable unit costs.

Typical for these primary parameters is that when their values change in a cash flow simulation, they affect simultaneously several of the previously mentioned indirect real option parameters. Most likely a change in a significant parameter in a cash flow calculation affects at the same time both the expected underlying asset value and some other parameters used to describe its stochastic process. Because of the interactions with serial and cross-correlations in the simulated cash flow calculation, also the stochastic process used to describe the underlying asset movements may need to be changed. Commonly assumed geometric Brownian motion may not be appropriate and robust enough, because it does not handle negative underlying asset values and multiplicativity is not necessarily the most reliable description of the stochastic process. Also the amount of uncertainty in terms of underlying asset value ambiguity in the beginning may change.

Sensitivity analysis with simulated cash flow options should be conducted in a way that shows how the change in a primary parameter, e.g. sales unit price, affects the total value of
the case. This can be conducted with ordinary least squares regression based method, which calculates how the underlying asset value and its stochastic process (expected value, ambiguity, and volatility) change simultaneously as a result of change in a single or multiple primary variables of cash flow calculation. These regression estimator values are then used in the real option valuation method (e.g. in a binomial lattice). The procedure is illustrated in Figure 1.

![Figure 1: Illustration of the sensitivity analysis approach. The most important primary cash flow simulation parameters are used as regressors on estimating several underlying asset properties that determine the binomial tree parameterization used in project valuation.](image)

The structure of this paper is as follows. After the introduction, section 2 discusses real option valuation based on a consolidated cash flow simulation based on the ideas of Copeland & Antikarov (2001; 2005), and why this approach may be better in some cases instead of applying directly Longstaff & Schwartz (2001) least squares regression method. Section 3 discusses sensitivity analysis, focusing mostly on the sampling based approaches. It also describes the suggested sensitivity analysis method and illustration of contribution to variance chart with grouped variables. This provides information of the most important primary cash flow simulation input parameters, and along with other additional analysis methods, suggests what primary input parameter could be used as regressors on estimating secondary parameters describing the underlying asset properties. These secondary parameters are then used in ROV sensitivity analysis. Section 4 presents a stylized case to demonstrate the use of explained sensitivity analysis. Section 5 shows construction of regression estimators. Section 6 discusses the results, and Section 7 finalizes the paper with conclusions.

2. Real options valuation with cash flow simulation

When there is no tractable asset in the market, one alternative is to use Monte Carlo simulation on cash flows for estimating the underlying asset value and consolidated volatility for valuation. Most comprehensive discussion and description of the method and valuation framework (Figure 2) is presented in Copeland & Antikarov (2001; 2005). The cash flow valuation parameters are modeled so that their uncertainty is presented with distributions,
time series, cross- and auto-correlations. The cash flow is simulated and based on that, both the expected value and its stochastic process are estimated. Therefore, this procedure consolidates all the continuously evolving risks into a single deviation measure, consolidated univariate volatility estimate. The approach is based on the idea that an investment with real options should be valued as if it was a traded asset in markets even though it would not be publicly listed. According to Copeland and Antikarov, the present value of the cash flows of the project without flexibility is the best unbiased estimate of the market value of the project were it a traded asset. They call this assumption *marketed asset disclaimer*.

After the simulation, real option valuation is done applying some algebraic numerical method, i.e. either with a risk-adjusted decision tree or a lattice method. Most common of these lattice methods is using binomial tree pricing, which is based on 1) a simple representation of the evolution of the continuous underlying asset value process with a discrete random walk and 2) the assumption of risk-neutral pricing. Lattice methods are popular because they are accurate, robust, and intuitively appealing tools for valuing financial and real options (Hahn 2005, p. 8), and they are much more easily explained to and accepted by management (Mun 2002, pp. 100). They also allow valuation of multiple interacting options within the same project (Trigeorgis 1993), which is essential for practical valuation purposes.

**Figure 2: Five-step real option valuation framework according to Copeland & Antikarov (2005).**

There are also other Monte Carlo simulation based volatility estimation methods along with the logarithmic present value approach of Copeland & Antikarov (2001). Herath & Park (2002) also use a modified version of Copeland & Antikarov (2001) for volatility estimation, Brandão, Dyer and Hahn (2005) present a conditional logarithmic present value approach, and Godinho (2006) presents two-level simulation and least-squares regression method. All these methods use simulation on cash flows to consolidate a high-dimensional stochastic process of several correlated variables into a low-dimension (univariate) geometric Brownian motion summary process. The volatility parameter $\sigma$ of the underlying asset is
estimated by calculating the standard deviation of the simulated probability distribution for the rate of return.

A good question is whether the consolidated cash flow methods are any better than applying directly the Longstaff & Schwartz approach for similar cases with several input parameters and interacting options. Longstaff & Schwartz (2001) approach for valuing American options uses ordinary least squares regression for estimating forward-looking estimator. Then, at each time period and for each underlying asset price level, given the combination of regression estimator input parameter values, a decision is made of the optimal continuing strategy. This approach is sufficiently flexible and straightforward if the number of options in a case is small, and the method is also computationally fast. The method can also be applied as a sensitivity analysis method by changing the regression estimator input parameter values, which in turn change the expected value according to the estimator.

The shortcoming of the approach is that it is difficult to apply the same method when there are several interacting sequential and parallel options. Determining a good forward-looking estimator for the whole project value, especially in the early stages, is difficult if not impossible, because the estimator should be able to capture the future value including the forthcoming interacting options. This problem can be mitigated by applying Bellman’s principle of optimality with backward induction logic. Starting from the end of the project, regression equations are determined for each stage, and the optimal decision forward-looking decisions are actually worked backwards. This is very similar to the common logic of binomial trees. However, even if the previously presented idea is not impossible, it is also very burdensome. Forward-looking estimators need to be defined for many different points in option state space of time and underlying asset value, and therefore, the number of required regression equations becomes large.

Most interesting of the previously explained consolidating cash flow simulation approaches – inspired by Longstaff & Schwartz (2001)\(^1\) – is the least-squares regression method of Godinho (2006). Godinho (2006) uses OLS regression to estimate the future value of the underlying asset, conditional on earlier time periods’ realizations of cash flow simulation input parameter values. This approach can be extended to allow simultaneously changing volatility and displaced diffusion process of Rubinstein (1983). The resulting valuation approach has advantage of being as straightforward to use as the original Copeland & Antikarov (2001) approach, while taking into account the forward-looking nature similarly to Godinho (2006) with Longstaff & Schwartz (2001) regression estimator. Thus the approach mostly shares benefits of both methods.

\(^1\) Longstaff & Schwartz (2001) on the other hand was inspired by Carriere (1996)
However, similarly to all the other cash flows simulation based methods, there are two shortcomings in this approach. Firstly, the approach often uses also subjective estimates in the cash flow calculation. The reliability of the results is strongly dependent on the assumptions used in cash flow modeling. Therefore, justified use of this approach also requires as much market based information as possible. This is naturally not a method specific issue but lack of information is a common obstacle in using nearly any valuation method. The second problem in these consolidated approaches is that their valuation logic hides the actual decisions that should be made. The procedures show the optimal strategy given the value of the underlying asset value, but it does not show what are the actual input parameters and their values to determine the underlying asset value in each stage. Therefore, the decision maker does not necessarily know what input parameter values should be monitored and what their combined effect on the underlying asset value is.

Because of both of these reasons, there is a need to do careful sensitivity analysis on the valuation case. Firstly, decision maker needs to know what are actually the most significant primary cash flow calculation simulation parameters driving the underlying asset value. Knowing this information helps in gathering more information about the most uncertain parameters. It also helps in understanding what of part of this uncertainty is because of volatility –like uncertainty, and what part of uncertainty can be explained by ambiguity or lack of own knowledge. This information is required to improve the quality of the valuation model as well as in interpretation of the calculated result. Secondly, decision-maker requires some method that helps in understanding what the combined effect of several input parameters is and how this information affects the estimated future value.

Two adjustments were made to the original valuation approach of Copeland & Antikarov. First extension to the process is using a three-parameter shifted log-normal binomial tree (Haahtela 2006) following a displaced diffusion process of Rubinstein (1983) instead of ordinary Cox, Ross & Rubinstein (1979). It allows modeling of the terminal value distribution shapes that can vary between normal distribution and a lognormal distribution and it also allows negative underlying asset values. The second extension is to use stochastic time varying volatility. The solution procedure is mostly based on the combination of using least-squares regression method similarly to Godinho (2006) in combination with generalized risk-neutral volatility estimation (Hull 1997) and ideas of conditional volatility estimation (Brandão et al. 2005, Smith 2005). Also ideas of volatility estimation method presented in Haahtela (2010) and Haahtela (2011) are used.

3. Sensitivity analysis

The purpose in sensitivity analysis is to determine which input parameters are significant and contribute most to the output variability. This is measured as the magnitude of a change in the solution resulting from a change in independent variables and parameters. Sensitivity analysis
is essential in order to identify the critical aspects of the investment model that affect model output uncertainty. Based on this relative importance, sensitivity analysis provides guidance for further data collection and modeling (Manache & Melching 2008). Sensitivity analysis reveals under which conditions the model or with which parameter values the solution is robust, i.e. under what circumstances the model is insensitive to the changes in input parameters. It helps in defining under what circumstances the change in operational activities is required. In case of real options, this is required in order to be able to decide on option exercise timely. Sensitivity analysis may also reveal what is the accuracy of the final model and the degree of uncertainty. This may be related to data accuracy, model uncertainty, and this uncertainty is reflected both as ambiguity and volatility uncertainty.

Four desirable properties of sensitivity analysis are 1) the ability to cope with the influence of scale and shape, 2) grouping of factors, 3) model independency, and 4) multidimensional averaging (Saltelli et al., 2004). The first requirement, ability to cope with the influence of scale and shape, means that the influence of the input should incorporate the effect of the range function. The second desirable property of grouping factors means being able to treat grouped factors as if they were single factors. This property of synthesis is essential for the agility of the interpretation of the results. Third desirable property is model independency, which means that the method should work regardless of the additivity or linearity of the model. Fourth property, multidimensional averaging, means capability for global sensitivity analysis instead of local sensitivity analysis. Local sensitivity analysis is concerned with output variability in the neighborhood of a nominal value in the input space while global sensitivity analysis is concerned with the variability of model output over the entire input space (Yeh and Tung, 1993). Option sensitivity analysis with computation of the Greek’s is very typical and similar approach. Most academic models use local sensitivity analysis.

The problem setting is different for practitioners involved in the analysis of risk. For these the degree of variation of the input factors is material, as one of the outputs being sought from the analysis is a quantitative assessment of the uncertainty around some best estimate value for Y (uncertainty analysis). A global method evaluates this effect of a factor while all others are also varying. This is also essential for real options based on cash flow simulations where several interacting parameters affect the underlying asset simultaneously. A global sensitivity measure must be able to handle interaction effect, which is especially important for non-linear, non-additive models. These arise when the effect of changing two factors is different from the sum of their individual effects.

When real options valuation is based on simulated cash flows, it is natural to consider the sampling (random or pseudo-random simulation) based methods as viable alternative for the sensitivity analysis. Monte Carlo analysis is based on performing multiple evaluations with randomly selected model input, and then using the results of these evaluations to determine
both uncertainty in model predictions and apportioning to the input factors their contribution to this uncertainty.

Sampling based methods for sensitivity analysis have a number of desirable properties. They have the advantage of conceptual simplicity with ease and flexibility in adaptation to specific analysis situations. Sampling based uncertainty analyses are global by nature, and they also allow stratification over the range of each uncertain variable. Sampling based methods also allow estimation of the distribution functions to characterize the uncertainty in model predictions. There are also many analysis techniques available (Helton 2000, 2006).

Most common sampling based sensitivity analysis methods are computations of linear correlation coefficients (LCC), partial correlations coefficients (PCC), semi-partial correlation coefficients (SPCC) and regression coefficients (RC). Also the rank-correlated versions of the four methods are suggested, and they are especially useful when the relationship between the inputs and outputs is not linear but still monotonic. Rank versions of the previous methods calculate the relationship between two data sets by comparing the rank (order number) of each value in a data set. Properties and differences of these models have been discussed in Campolongo (2000), Helton & Davis (2000), Helton & Davis (2002), Coyle et al. (2003), and Manache & Melching (2008). The main findings in the previous studies show that, without strong non-linearity on non-monotonicity, all the methods recognize and rank different uncertainties nearly identically.

Of the above mentioned alternatives, regression coefficient approach was selected because of its popularity and qualities. The analysis is performed by feeding the regression algorithm (such as ordinary least squares) with model input and output values. The regression algorithm returns a regression meta-model, whereby the output \( Y \) is described in terms of a linear combination of the input factors (Equation 1).

\[
y_i = b_0 + \sum_j b_j x_{ij} + \epsilon_j
\]  

(1)

The regression coefficient for a given factor plays the role of a sensitivity measure for that factor. To avoid the problem of units with ordinary regression coefficients, a common standardization method is normalization, in which each single data point is divided by their average values. With advanced stratified sampling techniques, of which especially the Latin hypercube is most common, the extraction of a large amount of uncertainty and sensitivity information is possible with a relatively small sample size. (Saltelli et al., 1995).

Regression coefficients quantify the effect of varying each input variable away from its mean by a fixed fraction of its variance while maintaining other variables at their expected values. The measure, \( R^2 \), represents the fraction of the variance that the regression model can explain. Statistical significance of the overall fit can be checked by an F-test and the significance of
individual coefficients with a t-test. Other regression diagnostics are the check of model linearity, homoscedasticity, non-multicollinearity, and normality of error terms without autocorrelation.

Regression coefficient method can handle globally the properties of scale and shape as well as multidimensional averaging. It explores the entire interval of definition for each factor. Another is that each ‘effect’ for a factor is in fact an average over the possible values of the other factors. As a global sensitivity method, regression coefficient analysis allows evaluation of individual model inputs while taking into account the simultaneous impact of other model inputs (Cullen & Frey 1999, 563).

On the other hand, regression based methods are not model independent and as such they do not have grouping of factors. Regression based methods can be totally misleading for non-linear and non-monotonic models. However, neither of these deficiencies causes actually problems in the underlying asset forecasting and sensitivity analysis. First, the grouping of variables can be conducted manually or with a little help of computer software. Second, even if the pattern of the cash flows with different time series and correlations may fluctuate strongly, their overall effect on the underlying asset gradually levels out during several time periods. As a result, the entire underlying asset stochastic process behaves sufficiently smoothly as well.

The first alternative is to recognize which are actually the most important primary variables driving the underlying asset value. These can be recognized by applying a standardized sensitivity regression analysis. The regression coefficients are computed and presented in table form with other statistical information. This information can be used for estimating how well the regression model can estimate the final output (most notably R squared). Also a graphical illustration in a form of tornado graph may be used to visualize the results, as it also shows the sign (+/-) of the coefficient on the output.

Ranking input parameters by their contribution to the variance of the outcome is a commonly suggested measure of parameter importance (Birnbaum 1969, Iman & Helton 1988). This can be done by squaring the calculated correlation coefficients and normalizing the results to 100%. Another suggested modification is grouping of the factors. In cash flow calculation, most parameters are discretized over the time either on a monthly, quarterly or annual basis. Regression sensitivity analysis calculates coefficients separately for each of these parameters on each time period. As a result, variables spanning over several time periods do not come up in comparison with those variables that are modeled as a single variable (yet probably affecting also other time periods).

A suggested method is to construct and present the contribution to variance chart with a matrix table of primary parameters and time periods (Figure 3). The table shows:
1) what are the most significant uncertainties and how much of the total variance can be explained by each variable or variable group
2) how much the model can explain uncertainty on the whole
3) how much of the total variability is related to each time period
4) how the order of most significant uncertainties changes between time periods
5) how the uncertainty changes inside a grouped variable between the periods

All these are relevant information for decision making and guide in finding new information and consideration of possible actions. Decision maker also wants to know how much of the uncertainty is revealed after each time period. At the same time it also reveals how much of the future uncertainty cannot be explained even if we had information of some input values at certain time points. This all provides information of how much the expected value is likely to change in other stages as a result of uncertainty, and when the uncertainty has resolved, how much there is still uncertainty left in form of volatility.

The results of this regression sensitivity analysis are also of significant use in constructing a forward-looking regression-based value estimator, or a response surface method estimator, for the underlying asset and its stochastic process. Because different uncertainties are of different importance in each period, this graph helps in recognizing the parameters that should be considered during each time period for regression based estimator of present value. The present value estimator \( PV_x \) is usually of the following functional form where \( X_{k,t} \) are cash flow calculation state variables (Equation 2):

\[
\bar{PV}_x = a_x + b_{1,x}X_{1,x} + b_{2,x}X_{2,x} + \cdots + b_{k,x}X_{k,x} + \text{possible higher order terms of } X_t
\] (2)

What the common ranking based methods do not take into account are the higher order effects between the parameters. Also, they assume that the relationship is monotonic and without non-random patterns, and when these assumptions do not hold, the results may not be accurate. Samuelsson’s proof (1965) and central limit’s theorem suggest that the expected value is a random walk regardless of the non-monotonic (e.g. cyclical) pattern of the variable. Scatterplots of dependent variable versus independent variables often display if there are non-monotonic or non-random patterns, thresholds and even rather complex variable interactions (Iman & Helton 1988), (Kleijnen & Helton, 1999). Also statistical testing can be conducted. Non-monotonic patterns can be recognized with F-test for equal means, \( \chi^2 \) test for equal medians, and the Kruskal-Wallis for common locations with rank-transformed data, are means of determining if measures of central tendency for dependent variable change as a function of the values of individual independent variables. The \( \chi^2 \) statistic may also be used for identifying non-random patterns by superimposing reasonable grids on the scatterplots. The purpose of these statistical calculations and their test setups are for the purpose of sensitivity ranking and not for accepting or rejecting hypotheses (Hamby 1994). Discussion
and details about these additional tests are described in more detail for the use in sensitivity analysis in Hamby (1994), Helton (2000), and Helton et al. (2006).

4. The case example

The method is illustrated here with a stylized example adapted from actual industrial systemic case. In the case example, unit selling price is modeled as a decreasing time series with normally distributed deviation during each time period. Variable unit costs are modeled similarly. Sales quantity is difficult to estimate for the first year. After that, the sales are much more predictable, because the same customers are likely to buy the same devices. As a result, the first year is modeled with Pert distribution (minimum, median, maximum) as Pert(15 000, 20 000, 40 000). The later years 2011…2015 are modeled according the historical knowledge of customer behavior, given that each time the sales of the previous years are known. Both selling price and sales quantity are modeled so that their initial value is presented by a distribution, and the following year parameter values are modeled in a time series manner as a change from the previous years’ values. Also some other variables have small variation around their mean values, but their effect on the overall uncertainty is not significant. The following Table 1 illustrates this simplified cash flow calculation.
Table 1: Cash flow calculation of the case example.

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>unit sales price</td>
<td>2.00</td>
<td>1.80</td>
<td>1.62</td>
<td>1.46</td>
<td>1.31</td>
<td>1.18</td>
<td></td>
</tr>
<tr>
<td>sales quantity</td>
<td>25 000</td>
<td>37 500</td>
<td>37 500</td>
<td>28 125</td>
<td>21 094</td>
<td>15 820</td>
<td></td>
</tr>
<tr>
<td>Variable unit costs</td>
<td>1.70</td>
<td>1.36</td>
<td>1.16</td>
<td>0.98</td>
<td>0.88</td>
<td>0.80</td>
<td></td>
</tr>
</tbody>
</table>

**Cash flow calculation**

<table>
<thead>
<tr>
<th></th>
<th>2010</th>
<th>2011</th>
<th>2012</th>
<th>2013</th>
<th>2014</th>
<th>2015</th>
</tr>
</thead>
<tbody>
<tr>
<td>Revenue</td>
<td>50 000</td>
<td>67 500</td>
<td>60 750</td>
<td>41 006</td>
<td>27 679</td>
<td>18 683</td>
</tr>
<tr>
<td>Variable costs</td>
<td>-42 500</td>
<td>-51 000</td>
<td>-43 350</td>
<td>-27 636</td>
<td>-18 654</td>
<td>591</td>
</tr>
<tr>
<td>fixed costs</td>
<td>5 000</td>
<td>5 000</td>
<td>5 000</td>
<td>5 000</td>
<td>5 000</td>
<td>5 000</td>
</tr>
<tr>
<td>depreciation</td>
<td>-833</td>
<td>-833</td>
<td>-833</td>
<td>-833</td>
<td>-833</td>
<td>-833</td>
</tr>
<tr>
<td>EBIT</td>
<td>11 667</td>
<td>20 667</td>
<td>21 567</td>
<td>17 538</td>
<td>13 192</td>
<td>10 259</td>
</tr>
<tr>
<td>Taxes</td>
<td>-4 667</td>
<td>-8 267</td>
<td>-8 627</td>
<td>-7 015</td>
<td>-5 277</td>
<td>-4 104</td>
</tr>
<tr>
<td>depreciation</td>
<td>833</td>
<td>833</td>
<td>833</td>
<td>833</td>
<td>833</td>
<td>833</td>
</tr>
</tbody>
</table>

| initial investment| -10 000|
| inc in NWC        |        |
| cash flows (CFx)  | 7 833  | 13 233 | 13 773 | 11 356 | 8 748  | 6 988  |
| WACC              | 0.110  | 0.110  | 0.110  | 0.110  | 0.110  | 0.110  |
| DF                | 0.901  | 0.812  | 0.731  | 0.659  | 0.593  | 0.535  |
| discounted cash flow (11%) | 7 057 | 10 740 | 10 071 | 7 480  | 5 192  | 3 736  |
| PV_{i}(FCF)       | 41 314 | 11 922 | 11 179 | 8 303  | 5 763  | 4 147  |
| Present value      | 44 276 |
| PV(Proj)          | PV\_0  | 44 276 | 46 546 | 48 933 | 51 442 | 54 079 |
| (r_f = 5.0 %)     |        | 56 852 | 59 767 |        |        |        |

The project is such, that the initial investment requirement is 9 000. Then, after one year, investment of 30 000 is required to complete the manufacturing construction. Two years from that, if the market outlook is bright, the company may further invest 20 000 and increase their expected profits by 30%.

Figure 3 presents the contribution to variance chart with a matrix table of primary parameters and time periods. It shows in columns how much of the total variance can be explained by each primary variable when grouped together, and the rows tell how much of the variability is related to each time period. As the results indicate, the order of significance for different uncertainties is unit sales price, sales quantity and variable unit costs. The results also show that uncertainty related to most parameters declines significantly over time. If on the whole a large proportion of variance is explained by the first year, it usually means that there is a lot of ambiguity in the underlying asset value. This is mostly due to lack of not having knowledge or sufficiently precise estimate of the primary variables and therefore also about the underlying asset value. In the example, 68.6 % of the total variance can be explained by
the first year, which is more than the remaining five years contribute on the whole (see Figure 3).

Also of interest is to see how single grouped variable changes during the time span. Results of the Figure 3 illustrate this very well. Second important aspect is to recognize the pattern inside a certain group of variables. Sales quantity is a good example of this. Sales quantity is a good example of this. First year demand is hard to predict, but after that it becomes more reliable. This is because the illustration in graph is adapted from an industrial case in a systemic environment, where the first year determines very much the demand for products and service in the next following years. Variable unit costs between the years, on the other hand, cause mostly the same amount of variation regardless of the realization of the earlier years.

Figure 3: Contribution to variance chart with corresponding table of information.

<table>
<thead>
<tr>
<th></th>
<th>2011</th>
<th>2012</th>
<th>2013</th>
<th>2014</th>
<th>2015</th>
<th>Total variance by variable</th>
</tr>
</thead>
<tbody>
<tr>
<td>Unit sales price</td>
<td>32,0 %</td>
<td>13,1 %</td>
<td>3,0 %</td>
<td>1,2 %</td>
<td>0,4 %</td>
<td>49,7 %</td>
</tr>
<tr>
<td>Sales quantity</td>
<td>24,9 %</td>
<td>0,5 %</td>
<td>0,2 %</td>
<td>0,1 %</td>
<td>0,1 %</td>
<td>25,8 %</td>
</tr>
<tr>
<td>Variable unit costs</td>
<td>6,3 %</td>
<td>5,1 %</td>
<td>2,8 %</td>
<td>1,0 %</td>
<td>0,3 %</td>
<td>15,5 %</td>
</tr>
<tr>
<td>Total explained annual variance</td>
<td>63,2 %</td>
<td>18,8 %</td>
<td>6,1 %</td>
<td>2,3 %</td>
<td>0,7 %</td>
<td>91,0 % explained</td>
</tr>
<tr>
<td>Cumulative annual explained variance</td>
<td>63,2 %</td>
<td>81,9 %</td>
<td>88,0 %</td>
<td>90,3 %</td>
<td>91,0 %</td>
<td>91,0 % unexplained</td>
</tr>
</tbody>
</table>
Table 2: Summary output of the regression analysis and the calculation of explained variation, volatility and shifted volatility.

<table>
<thead>
<tr>
<th>Regression Statistics</th>
<th>2011</th>
<th>2012</th>
<th>2013</th>
<th>2014</th>
</tr>
</thead>
<tbody>
<tr>
<td>Multiple R</td>
<td>0.823</td>
<td>0.945</td>
<td>0.981</td>
<td>0.995</td>
</tr>
<tr>
<td>R Square</td>
<td>0.677</td>
<td>0.892</td>
<td>0.963</td>
<td>0.990</td>
</tr>
<tr>
<td>Adjusted R Square</td>
<td>0.677</td>
<td>0.892</td>
<td>0.963</td>
<td>0.990</td>
</tr>
<tr>
<td>Standard Error</td>
<td>12420</td>
<td>7540</td>
<td>4643</td>
<td>2563</td>
</tr>
<tr>
<td>Observations</td>
<td>10000</td>
<td>10000</td>
<td>10000</td>
<td>10000</td>
</tr>
<tr>
<td>Explained variation</td>
<td>17977</td>
<td>21698</td>
<td>23700</td>
<td>25260</td>
</tr>
<tr>
<td>Volatility $\sigma_1$</td>
<td>36.40 %</td>
<td>22.35 %</td>
<td>11.49 %</td>
<td>7.40 %</td>
</tr>
<tr>
<td>Shifted volatility $\sigma_{dd,1}$</td>
<td>16.83 %</td>
<td>10.71 %</td>
<td>5.57 %</td>
<td>3.60 %</td>
</tr>
</tbody>
</table>

Figure 3 and table information can also be used for estimating how the volatility changes over time periods. If we compare the results of Figure 3 (Cumulative annual explained variable) to the Table 2 Summary output Regression statistics (R Square) of the same case example using the best found response surface for each time period, we can see that the results are very similar. This is not surprising, because often very basic regression models without any higher order or interacting terms are capable to forecast future cash flows nearly as well as more advanced response surface equations. As such, the contribution to variance chart and table also reveals what is the minimum proportion of the cumulative uncertainty that can be explained by a simple linear regression model. If there were more interaction between the cash flow calculation parameters, this difference would be more significant. The difference between the two is mostly because of the structure of cash flow calculation model. The case example cash flow calculation model has auto-correlation, but no cross correlation or any significant higher order or interaction effects beyond those interactions presented in the equations (see footnote of this page)

5. Ordinary least squares regression based sensitivity estimator

---

3 The regression estimators for each $PV_i$ of which the volatilities in Table 2 are calculated, are the following:

$$PV_1 = \alpha_1 + \beta_{11}(SQ_1 \times SUP_1) + \beta_{12}(SQ_1 \times UC_2) + \beta_{3,1}FC_1$$

$$PV_2 = \alpha_2 + \beta_{1,2}(SQ_2 \times SUP_2) + \beta_{2,2}(SQ_2 \times UC_2) + \beta_{3,2}FC_1 + \beta_{4,3}FC_2$$

$$PV_3 = \alpha_3 + \beta_{1,3}(SQ_3 \times SUP_3) + \beta_{2,3}(SQ_3 \times UC_2) + \beta_{3,3}FC_3 + \beta_{4,4}FC_2 + \beta_{5,3}FC_3$$

$$PV_4 = \alpha_4 + \beta_{1,4}(SQ_4 \times SUP_4) + \beta_{2,4}(SQ_4 \times UC_2) + \beta_{3,4}FC_4 + \beta_{4,4}FC_1 + \beta_{5,4}FC_2 + \beta_{6,4}FC_3$$
First consideration is how much the expected underlying asset value changes as a result of a change in a primary cash flow calculation input parameter. However, the problem is that the change in the input parameter is likely to change the uncertainty as well. It is also more likely that larger values have larger absolute deviations from the mean than smaller values, while the relative proportion of standard deviation in comparison with expected value is likely to decrease. Therefore, sensitivity analysis has to take into account the effect of both the expected value and the deviation change on the underlying asset value.

It would have been possible to use the exact form of (see the footnotes on the previous page). However, to illustrate the method use, only the simple form of linear regression is applied instead of the more advanced (but only slightly) more accurate response surface equations. The use of approach is illustrated for the first year of cash flow calculation.

After the cash flow simulation, the regression model with most significant primary variables is made to predict the underlying asset values. Independent variables in this case are the Demand, Sales price and Variable costs in time period one. The dependent variable is present value of the cash flows. Linear regression estimator results into the following equation for the expected mean value of the underlying asset value:

\[
\hat{PV}_1 = -9327 + 1.282 \times Demand + 55393 \times Sales \text{ prices} - 51123 \times Variable \text{ costs}
\] (3)

Because of the nature of the risk-neutral assumption in valuation, the expected underlying asset value increases by risk-free interest rate between the time periods.

After that, simulation data is sorted according to the dependent variable (PV) estimator from smallest to largest. Then the difference between the regression estimator values \(\hat{PV}_1\) and the simulation based realized values \(PV_i\) is calculated. The standard deviation of these differences is 13 419, the same result which is also confirmed by regression software as standard error. However, this average does not tell whether the standard deviation is the same for large and small values expected values of the underlying asset. Therefore, the standard deviation of the differences between simulated values and regression estimator values are computed as a moving average consisting of 200 data points (-100…100) around each estimated value. The results in the case example illustrate that the standard deviation increases as a result of increase in expected value both for the standard deviation in time period \(T_1\) and for the terminal value underlying asset distribution at the maturity \(T_e\). Linear regression estimators result into the following equations for the standard deviations in \(T_1\) and in \(T_e\)

\[
std(\hat{PV}_1) = 7535 + 0.0972 \times \hat{PV}_1
\] (4)
\[
\text{std}(\overline{PV}_p) = 9.675 + 0.0972 \times \overline{PV}_p
\]  
(5)

It is also possible to explain the standard deviations using the primary variables as explanatory parameters instead of using combined regression estimator based expected mean value of the underlying asset. This is suggested if such model has better explanatory power or if it reveals some other property in the behavior of the underlying asset cash flow model, for example as a result of significant non-linearity or heteroscedasticity caused only by a single or few of the explanatory variables.

If the standard deviations of the underlying asset process at certain time points are known, it is possible to compute the average volatility for each time period. Starting from the beginning of the process, volatility \( \sigma_i \) for each time period can be calculated according to Equation (6):

\[
\sigma_i = \sqrt{\ln \left( \frac{\text{std}(S)_i}{S_0 e^{rT}} \right)^2 + 1} - \sum_{i=0}^{i-1} (\sigma_i^2 \cdot t_i)
\]  
(6)

When the regression estimators are done, and they can be used for parameterizing common option valuation methods, and especially their stochastic process parameters. In case of several interacting options, this means usually choosing a lattice valuation approach. As a result, we can finally check how real option value changes as a result of the change in its primary cash flow calculation input parameter.

6. Results

Figure 4 shows how the project value with real options changes as a result of change in unit price (on the left) and Demand (on the right). The results and their linearity in this particular case example are not very surprising. Increase in both cases is mostly due to the increase in the underlying asset value as a result of the increase in the primary variable\(^4\). Also the options contribute to the total value, but because they are of opposite types (option to abandon like a put option and the extension option like a call option), they are valuable at opposite situations. Therefore their combined added value does not change the rather linear nature of the change in project value as a result of a change in a primary variable.

---

\(^4\) Next revision of this paper will include more detailed analysis of how much this change is caused by a change in 1) expected value, 2) volatility, and 3) by their combined interaction.
One strength of the regression based sensitivity analysis in combination with cash flow simulation based real options valuation is that it allows straightforward analysis of changes in multiple variables at the same time. The decision maker can change one or several primary parameter values at the same time, and the procedure calculates the total value of the project instantly, and it also shows how this happens as a result of changes in the underlying asset. Another approach to do this would require changing directly the cash flow simulation model, simulating the whole model again, and doing reconstruction on some parts of the model. However, using this approach for several scenarios as a double-check is suggested. If the results between re-simulation and regression-based estimators differ significantly, it suggests that the regression model is not capable of detecting some non-linearity or combined effect of multiple variables.

No research and constructed methods are without their limitations. Firstly, the underlying asset shape is assumed to remain the same regardless of the level of the underlying asset value. Secondly, the estimations and regression analyses are based on the cash flow simulation, which may have highly subjective estimates. Thirdly, any method based on regression estimations and averaging of several variables is likely to provide better approximations closer to the expected values of the primary parameters, while the decision-maker involved with risk analysis might want to know what happens in the extreme events. Fourthly, the results of the sensitivity analysis only illustrate what is the expected value of the project, given the information available, but there is still more or less uncertainty around this mean value.

7. Conclusions

This paper presented a regression based sensitivity analysis method for cash flow simulation based real option valuation. The method combines the uncertainties in the underlying asset together by Monte Carlo simulation, and then uses regression analysis approach for
estimating how a change in a primary cash flow variable – selling price, unit cost etc. - affects simultaneously to the underlying asset properties - expected value, volatility, and ambiguity - and as a result, via these indirect variables, changes the value of the project with real options. The method allows changing of several primary variables at the same time, and it calculates their combined effect on the project value instantly without need to re-simulations. However, critical interpretation of the sensitivity analysis results is required because of the limitations related to the use of regression model and partly subjective estimates in cash flow simulation.

References


Haahtela, T. (2006). Extended Binomial Tree Valuation when the Underlying Asset Distribution is Shifted Lognormal with Higher Moments. 10th Annual International Conference on Real Options, 14-17 June, New York, USA


