THE VALUE OF ASSESSING UNCERTAINTY IN OIL AND GAS PORTFOLIO OPTIMIZATION

A Thesis

by

HOUDA HDADOU

Submitted to the Office of Graduate Studies of
Texas A&M University
in partial fulfillment of the requirements for the degree of

MASTER OF SCIENCE

Chair of Committee,	Duane A. McVay
Committee Members,	Maria Barrufet
	W. John Lee
	George Voneiff
Head of Department,	A.D. Hill

August 2013

Major Subject: Petroleum Engineering

ABSTRACT

It has been shown in the literature that the oil and gas industry deals with a substantial number of biases that impact project evaluation and portfolio performance. Previous studies concluded that properly estimating uncertainties will significantly impact the success of risk takers and their profits. Although a considerable number of publications investigated the impact of cognitive biases, few of these publications tackled the problem from a quantitative point of view.

The objective of this work is to demonstrate the value of quantifying uncertainty and evaluate its impact on the optimization of oil and gas portfolios, taking into consideration the risk of each project. A model has been developed to perform portfolio optimization using Markowitz theory. In this study, portfolio optimization has been performed in the presence of different levels of overconfidence and directional bias to determine the impact of these biases on portfolio performance.

The results show that disappointment in performance occurs not only because the realized portfolio net present value (NPV) is lower than estimated, but also because the realized portfolio risk is higher than estimated. This disappointment is due to both incorrect estimation of value and risk (estimation error) and incorrect project selection (decision error). The results of the cases analyzed show that, in a high-risk-tolerance environment, moderate overconfidence and moderate optimism result in an expected decision error of about 19% and an expected disappointment of about 50% of the estimated portfolio. In a low-risk-tolerance environment, the same amounts of moderate overconfidence and optimism result in an expected decision error up to 103% and an expected disappointment up to 78% of the estimated portfolio. Reliably quantifying uncertainty has the value of reducing the expected disappointment and the expected decision error. This can be achieved by eliminating overconfidence in the process of project evaluation and portfolio optimization. Consequently, overall industry performance can be improved because accurate estimates enable identification of superior portfolios, with optimum reward and risk levels, and increase the probability of meeting expectations.

DEDICATION

I dedicate this thesis to my parents, my sister and my fiancé.

ACKNOWLEDGEMENTS

I would like to express my gratitude to my committee chair, Dr. McVay, for his supervision, guidance and encouragement throughout the course of this research; it was a great learning experience. I would like to thank my committee members, Dr. Barrufet, Dr. Lee and Mr. Voneiff, for their help and support.

I would like to thank ExxonMobil Corporation and the Institute of International Education for offering me a scholarship and giving me the opportunity to pursue my education in my field of interest. Thanks to Texas A&M University and Harold Vance Department of Petroleum Engineering for providing the resources to pursue my Master's degree. I would like to also thank my friends and colleagues, and special thanks to Mr. Dossary for collaborating and sharing his results.

Last but not least, thanks to my father, Abdelkarim, my mother, Fawzia, my sister, Amal, and my fiancé, Youssef, for their continuous support and love that gives me the courage to carry on.

TABLE OF CONTENTS

Page

LIST OF FIGURES

Page

Page

1. INTRODUCTION AND BACKGROUND

1.1 Introduction

The oil and gas industry faces a lot of uncertainties related to reserves estimation, production forecasting, pricing fluctuation, and many other factors. Estimation of each of these factors might vary from one estimator to another. Often, these estimations can have more than one value, thus requiring a range of possible values, and that range should properly account for the level of uncertainty of the estimator. Brashear et al. (1999) distinguished two types of uncertainties in the field of petroleum—underground uncertainties and aboveground uncertainties. The underground uncertainties are related to the reservoir and geological characteristics of a project, while the aboveground uncertainties are related to the fluctuations in prices, changes in demand and supply, changes in regulations and variations in estimators' judgments.

Capen (1976) pointed out that project planning and budget process is heavily dependent on estimations. According to him, properly estimating uncertainties will not only significantly impact profit, but it will also impact the success of the project. He conducted a set of experiments that showed that petroleum engineers, managers, decision makers, and estimators in general provide narrow ranges for their estimations due to their overconfidence. They underestimate uncertainty and overestimate the precision of their knowledge. Rose (2004) stated that exploration departments of most petroleum companies delivered about only half of the estimated reserves over the last twenty years of the 20^{th} century. He also explained that the following issue is not specific to exploration only since most companies fail to meet their forecast rates. The work of Tversky and Kahneman (1974) showed that people tend to base their judgment and estimation on a limited number of heuristic principles instead of properly assessing probabilities, which leads to biases. Decisions are usually based on the belief of the likelihood of an event or a value, thus the decision making process is impacted by the cognitive biases that occur when assessing probabilities. Choosing the right set of projects that would return the highest Net Present Value (NPV) and meet the performance criteria and budget constraints of a company is a task that requires not only familiarity with technical and financial concepts, but also a considerable awareness of the impact of biases. Hence, in this work, I will demonstrate the value of quantifying uncertainty, and I will evaluate its impact.

1.2 Background

1.2.1 Uncertainty

In Capen's (1976) experiments that evaluate the difficulty of assessing uncertainty, it has been observed that estimators tend to use the same range of estimations no matter what kind of range they were asked for. In one of Capen's (1976) experiments, he asked 10 questions to a group of petroleum engineers and he required that they provide 90% confidence level ranges for their estimations; the average confidence level of the answers turned out to be 32%. The explanation for this behavior is that decision makers (and even experts) tend to build ranges that include expected events only rather than expected and possible events. The difference between including expected events versus including expected and possible events is considered as an important source of uncertainty. Consequently, the ranges are usually too narrow. Welsh et al. (2005) conducted a study that is an extension of Capen's (1976) experiment; they concluded that even if the industry personnel have familiarity and experience with the domain, their decision and estimations are still impacted by biases.

Tversky and Kahneman's (1974) study was on the same track. He explained the common practices of people when making judgments and decisions under uncertainty by a number of sets of heuristics that lead to biases. One of these heuristics is the representativeness that takes place when people are asked to estimate the probability that an object belongs to another. People tend to misjudge representativeness because they are insensitive to a set of factors such as the prior probability outcomes, the sample size, predictability, misconceived chance, validity, and regression. Availability is another heuristic that was studied by Tversky and Kahneman (1974) and it is explained as the ease of instances' occurrence to one's mind. Availability can be subject to biases due to the retrievability of instances, to the effectiveness of a search set, to imaginability, and to the illusory correlations. The stated factors are misleading in the decision-making process; the human brain does not have the most effective search tools especially when operating in the intuitive fast-response mode. Also, different starting points yield different estimates, which create adjustment and anchoring heuristics due to insufficient adjustment. Tversky and Kahneman (1974) concludes that these heuristics are economical and usually effective, but they also lead to systematic and predictable errors not only by "average" people, but also by experienced researchers when they think intuitively.

Begg and Bratvold (2008) reported that the work of Brown (1974) concludes that even if the estimates of project cash flow are not biased, the average of the actually realized cash flows will be lower than estimated, because the projects with higher NPV are preferably, and not randomly, selected. Harrison and March (1984) show that the standard decision-process is biased due to pre-decision expectations and post-decision evaluations, and the bias may lead to disappointment. Smith and Winkler (2006) and other authors reported a phenomenon called the "Optimizer's Curse" or inevitable disappointment; it is the systematic bias resulting from the decision process itself. They (2006) considered a hypothetical situation consisting of 3 alternatives, each with value estimates that are considered to be their expected values resulting from a decision analysis study and assumed to be conditionally unbiased. All these papers have shown that bias occurs in the process of project evaluation and portfolio optimization. Brashear et al. (2001) reported that in the 1990's the largest US based E&P companies, both integrated majors and large independents, realized an average return on projects of 7% while the minimum estimated internal rate of return "hurdle rate" was set at 15%. The authors explain that this is a result of using evaluation methods that do not account for full uncertainties and risk.

Begg and Bratvold (2008) investigated the uncertainty in estimates and the impact of prediction errors in the O&G industry. Their work started with a literature review that covered the different theories and observations mentioned above. They then studied three typical situations and their sensitivity analysis. The situations are an intra-project alternative selection, project "go/no go" decisions, and a constrained portfolio selection subject to a budget limit. They concluded that the expected disappointment is real and present but it does not appear to be large compared to the other prediction errors; the magnitude of the expected disappointment is in the order of 2% and 10% (2008). Also, as the development of a project starts, the level of uncertainty and the impact of bias lower because positive bias and real-time feedback are added and that counteracts the loss of value. Begg and Bratvold also observed that the larger the number of alternatives, the higher the expected disappointment.

The study of McVay and Dossary (2012) measured the value of reliably assessing uncertainty. They built a mathematical model that describes the relationship between the true project value distribution(s) and the estimated project value distribution(s) in terms of two primary biases that affect the decision making process: overconfidence and directional bias. They then simulated the portfolio optimization process for the case of true values, the estimated values, and a realized case in which the decision is made based on the estimated values, but the value realized is based on the true distributions. Based on the results of these three simulations, McVay and Dossary (2012) calculated the expected disappointment (the difference between the realized NPV and the estimated NPV), and the expected decision error (the portion of expected disappointment due to the selection of the wrong projects). Their results showed that for moderate amounts of overconfidence and optimism, the expected disappointment is 30-35% of the estimated NPV for the industry portfolios and optimization cases they analyzed. As the degrees of overconfidence and optimism are greater, the expected disappointment approached 100% of the estimated NPV. Both of Begg and Bratvold

(2008) and McVay and Dossary's paper (2012) agree on the fact that there are systematic prediction errors that affect the process of project evaluation. Begg and Bratvold are more moderate, with a disappointment in the order of 2% to 10%, than McVay and Dossary, with an average disappointment of 30%-35%, about the magnitude of the impact of these errors on the decision making process.

1.2.2 Portfolio Optimization

Another relevant area for this study is portfolio optimization. Merritt and Miguel (2000) investigated and concluded that Monte Carlo simulation combined with Markowitz' theory of efficient frontier provides a powerful integrated modeling environment to analyze the efficiency of assets in oil and gas companies. This technique ensures that value is maximized for a certain level of risk. It also identifies opportunities to decrease the level of risk while maintaining the current project value. Markowitz' portfolio theory suggests that for any level of risk, there will be only one portfolio that returns a maximum reward and inversely, for any level of reward, there will be only one portfolio that minimizes the risk (Markowitz, 1952). He names the portfolios that meet these conditions "the efficient frontier" and the plot named efficient frontier curve (EFC) has the return of the optimum portfolios plotted against their risk. Markowitz pointed out that a rational investor would seek a portfolio for which no other combination would have a higher return without increased risk or lower risk without loss of return. The choice of the portfolios along the efficient frontier depends on the decision-maker's tolerance for risk (Brashear et al. 2001).

The sequential approach and the systems approach are techniques used in portfolio optimization. On one hand, the sequential approach uses the whole distribution of a variable as input from one sub-module to the other. On the other hand, the systems approach uses as input the individual samples that iterate through all the sub-modules before going to other samples of the distribution. An example to illustrate these two techniques is using, as input variable, porosity that could be fitted into a distribution. The porosity can be used in the calculation of the original oil in place (OOIP). In the sequential approach, the whole distribution of porosity is input into the OOIP calculation process. While in the sequential approach, a single value of porosity is sampled from the distribution and input into the OOIP calculation process. Al-Harthy et al. (2006) compare the sequential and the systems approaches used to optimize portfolios. They conclude that the two methods complement each other in capturing inter-dependencies and intra-dependencies, which can add significant value to the decision-making process. According to their work, the systems approach captures the intra-dependence within a project. They also concluded that as the level of risk increases, the difference between the two approaches increases as well, and the impact of capturing dependency is greater using the systems approach.

Another portfolio optimization method, used for a long time by E&P decision-makers, is to rank projects by investment efficiency (IE) (the project NPV divided by the investment or CAPEX). Once all projects are ordered, the decision-makers select the projects in a descending order according to IE until the available budget is exhausted. Brashear et al. (1999) described this optimization method as conventional and labeled it the cherry-picking solution. They conducted an experiment where they evaluated 14 investment opportunities using both the conventional ranking method and using Markowitz portfolio theory. Their experiment showed that the conventional method of ranking does maximize return but it ignores risk. This optimization method maximizes risk as well, so if the decision-makers are unwilling to maximize risk, other optimal portfolios can be chosen from the efficient curve.

Besides the work of Begg and Bratvold (2008) and McVay and Dossary (2012), few other quantitative studies have been conducted so far in the field of uncertainty quantification in combination with portfolio optimization. Most studies approached this area from a qualitative perspective or through surveys and experiments. Although the study of McVay and Dossary (2012) was based on a model that quantifies the impact of biases on oil and gas investments and portfolios, they used a basic portfolio optimization approach that consists of ranking projects and selecting a handful, from best to worst, until the budget is reached. It has been explained in the literature that Monte Carlo Simulation combined with Markowitz' theory of efficient frontier provides good results because it accounts for the risk factor within the evaluation of projects, in addition to the fact that it is more representative of what is used in the industry. Also, the impact of underestimation of uncertainty has not been studied quantitatively in a Markowitz portfolio optimization context. Thus, this work investigates how uncertainty and cognitive biases impact portfolio optimization in the context of Monte Carlo simulation combined with Markowitz theory.

1.3 Objective

The objective of this research is to demonstrate the value of quantifying uncertainty and assess its impact on the optimization of oil and gas portfolios, taking into consideration the risk factor carried within each project. Impact will be assessed by determining expected disappointment and expected decision error as a function of the two main cognitive bias parameters—overconfidence and directional bias.

2. PRIOR MODELS

2.1 The Begg and Bratvold Model

Begg and Bratvold (2008) wrote a paper in which they investigated the uncertainty in the estimates and prediction errors in the oil and gas industry. They also evaluated the importance and relevance of these errors on the overall portfolio performance. The authors looked at three different cases, a case for intra-project alternative selection using NPV as a metric, a case for "go"/"no go" decisions using positive NPV as the decision criterion, and another case for selecting projects in a limited budget context using NPV/CapEx criterion. The last case is the one of interest as some of its input distributions will be used in this work. Begg and Bratvold model ranked the projects in the pool by IE (investment efficiency that is equal to NPV divided by CapEx) and successively selected projects until the budget limit was reached. They characterized the true distributions of IE and CapEx by lognormal distributions and the uncertainty by a pert distribution that models the variability of the SD. They sampled the true and expected values of NPV, CapEx and IE of the 100 projects from these distributions. Based on these variables they calculated the expected disappointment (ED) and the expected decision error (EDE). For the case they studied that had a $4298MM true return, the disappointment was 6.1% of this return and the decision error was 2.9% of the same return. Thus, the paper concludes that the bias is present and real, but it is not considerably large compared to other prediction errors. Begg and Bratvold also investigated the impact of the project pool size on the returns and on the decision error and disappointment. Their experiment showed that the larger the pool out of which the projects are selected, the higher the portfolio values. As the number of available projects increase, the ED and EDE increase too. Though, for pools of 40 projects or higher, ED and EDE became fairly constant. ED%T and EDE%T are the ED and EDE as a percentage of the true portfolio, meaning they represent the ED and EDE respectively divided by the true NPV. This explains the tendency of ED%T and EDE%T with the increase in the number of projects, because the true NPV is also increasing at a rate higher than the decrease in ED and EDE.

2.2 The McVay and Dossary Model

The McVay and Dossary (2012) paper has the objective of quantitatively determining the value or the cost of properly accounting for uncertainty. They provide a new framework to model uncertainty. The new framework is based on the premise that all biases that affect oil and gas project selection can be boiled down to two factors: directional bias and overconfidence. The model presented in their paper optimizes portfolios by ranking projects from high to low IE and selecting projects in a descending order until the budget is exhausted. This optimization was done using true values and estimated values to simulate the performance of different portfolios in different settings. This experiment was done for the cases of constrained and unconstrained budget. The paper shows that a moderate level of overconfidence and optimism results in a disappointment of 30-35% of estimated NPV and an expected decision error of 1-5% of estimated NPV. As the level of optimism and overconfidence increases, the disappointment approaches 100% of the estimated value. Their paper shows that reduction in overconfidence reduces the expected disappointment even if bias remains constant.

McVay and Dossary's paper shows that the EDE%E is relatively small compared to the ED%E. ED%E reaches 100% while EDE%E does not exceed 10%. This means that most of the disappointment results from the estimation error rather than the decision error. The paper concludes that the value of reliably quantifying uncertainty resides in reducing disappointment to improve industry performance and identify superior projects. Using my model, I reproduced the experiment of McVay and Dossary for the case of constrained budget and deterministic bias. Both models were run for the same budget limit ($805MM) and number of projects (8 projects). This model was set to a high-risk-tolerance limit to make the optimization insensitive to the risk factor. The risk factor was not used in the calculations of the expected disappointment and the expected decision error.

Table 1—Summary of comparison to previous work

DB	OC	ED%E		EDE%E	
		SPE160189	This work	SPE160189	This work
0.25	0.25	8.6%	8.8%	0.5%	0.6%
0.5	0.25	21.8%	21.9%	0.4%	0.4%
0.75	0.25	33.7%	33.7%	0.5%	0.6%
-0.25	0.25	-26.5%	-26.4%	0.8%	1.0%
0.5	0.75	46.6%	46.7%	0.9%	1.1%
0.75	0.75	63.5%	63.6%	0.9%	1.1%
-0.25	0.1	-13.3%	-13.2%	0.3%	0.4%
0.5	0.5	35.1%	35.3%	0.8%	1.0%
-0.5	0.5	-118.6%	-118.9%	9.0%	9.0%

Table 1 summarizes the results of the two models for a few cases of different overconfidence and directional bias levels. The results of the expected disappointment and the expected decision error as a percentage of the estimated portfolio NPV were matched within 0.1-0.3%.

3. THE MODEL

3.1 The Methodology

3.1.1 Uncertainty Parameters

Uncertainty is modeled by following McVay and Dossary's (2012) premise that all the cognitive biases that affect oil and gas project evaluation can be represented by overconfidence and directional bias. These two parameters can be explained as follows:

- Overconfidence: this parameter specifies the fraction of the true distribution that is not sampled by the estimators. As shown in **Fig. 1**, an overconfidence of 0.5 means that the estimator is only considering 50% of the true distribution. This parameter ranges from 0 to 1. A value of 1 means that no range is considered and that it is a deterministic estimate. A value of 0 means, the whole distribution is considered.
- Directional bias: this parameter specifies the location of the estimated distribution that is a subset of the true distribution. This subset could be located at the left end of the distribution, right end or anywhere in between (**Fig. 1**). This parameter ranges between -1 and 1. A value of -1 refers to extreme pessimism meaning that only the most pessimistic outcomes of the true distribution are considered by the estimator. Extreme pessimism means that there is no truncation from the low end of the true distribution, and all of the truncation is from the high end. A value of 1 refers to extreme optimism meaning that only the most optimistic outcomes of the true distribution are considered. In this case, there is

no truncation from the high end of the true distribution, and all of the truncation is from the low end. 0 directional bias means there is no bias, and the truncation of the true distribution is equal from both ends. For other values of directional bias, a linear interpolation is used to obtain the fractions of area truncated from each end. For a directional bias of 0.5, 75% of the area that is truncated, which depends on the value of overconfidence, is truncated from the left of the distribution, and 25% of the same area is truncated from the right end of the distribution. For cost-based parameters (E.g CapEx), the truncation will be in the opposite direction (**Fig. 1**).

3.1.2 The Model Description

The model simulates the impact of biases on the process of portfolio optimization; these biases are represented by overconfidence and directional bias. The portfolio optimization is done using Markowitz theory combined with Monte Carlo simulation. The model is built on @RISK by Palisade Corporation (2012).

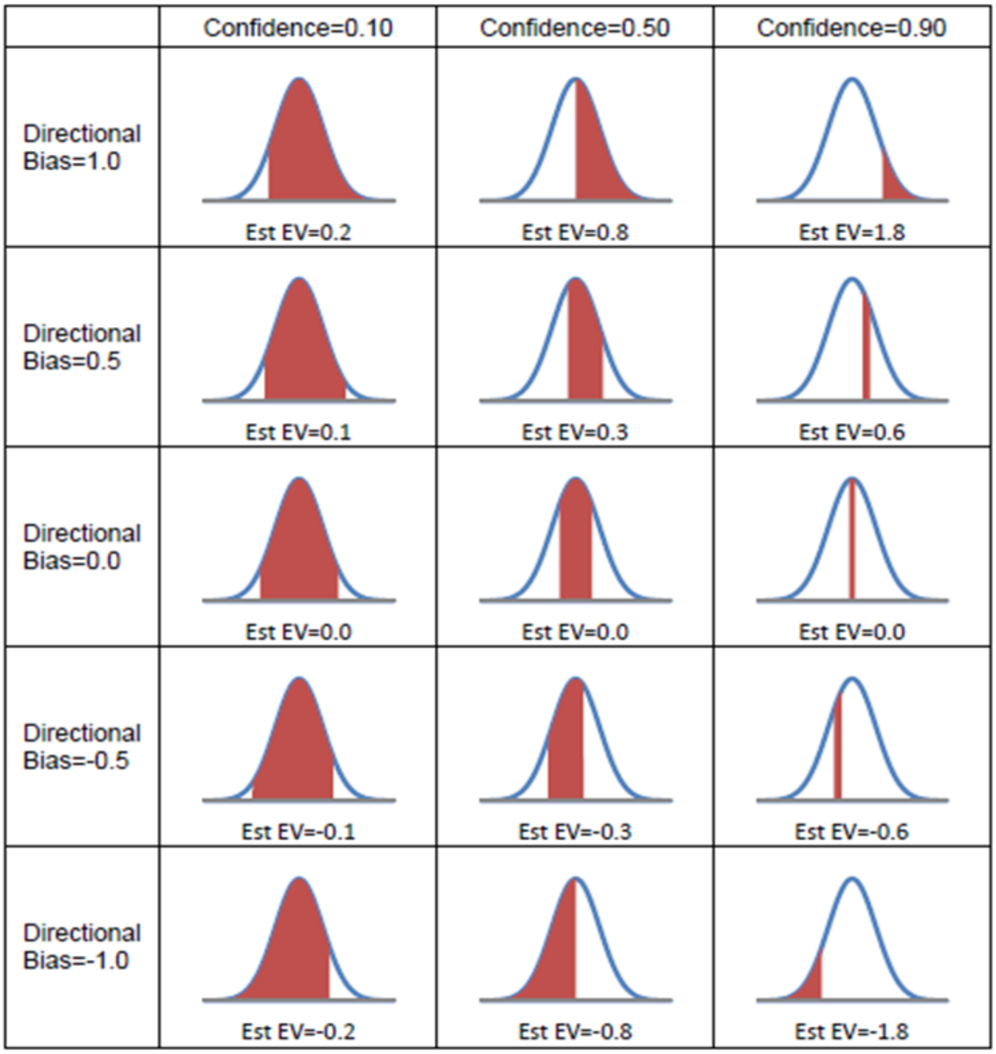

Fig. 1—Relationship between the estimated distribution (shaded) and the true distribution (unshaded) as a function of overconfidence and directional bias (McVay and Dossary, 2012)

Most of the parameters and distributions used in this work are similar to the ones used by McVay and Dossary (2012). Project economic performance is described using two random parameters—capital expenditure (CapEx) and the present value of the operating cash flow (PVOCF). PVOCF includes all cash outflows and inflows besides the CapEx. NPV is PVOCF minus CapEx. CapEx and PVOCF were modeled by McVay and Dossary (2012) as independent random variables even if there may be some correlations between these two variables in real life. This model also includes another factor—the portfolio risk. According to the Markowitz (1952) framework, the risk is the variance of the return. Other applications of the Markowitz framework defined risk as the standard deviation of the portfolio return (Guerard, 2009) or as the semi-standard deviation of the portfolio return (Brashear et al., 2001). In this work, risk is defined as the standard deviation of the portfolio NPV.

3.1.2.1 The True Project Value Distribution

The true project value distribution is the distribution that would be attained in what Begg and Bratvold (2008) call unlimited-resource environments. To get these true distributions, the company needs unlimited time, money and computational ability to assess the available data in the best and most accurate way. The data should already be available and the unlimited resources are only for assessment and not for acquiring more data or running more tests. The portfolios consist of projects that are sampled from global distributions designed to model the project alternatives available for a typical O&G company. For each project, CapEx and PVOCF distributions are determined by sampling means and standard deviations from global distributions. The standard deviations are specified relative to the true expected values, so that large projects have large uncertainty, and small projects have small uncertainty. The global distributions and the complete parameter set used are as follows:

- Mean true PVOCF: sampled from a lognormal distribution with mean equal to \$750MM, a SD equal to \$750MM and then shifted positively by \$300MM.
- Standard deviation of true PVOCF: specified relative to the true PVOCF. It is the true expected PVOCF multiplied by a value sampled from a Pert distribution with minimum 0.3, mode 0.8, and maximum 1.3.
- Mean true CapEx: sampled from a lognormal distribution with a mean equal to \$600MM, a SD equal to \$600MM and then shifted positively by \$100MM.
- Standard Deviation of true CapEx: calculated similarly to the SD of PVOCF, by multiplying the true expected CapEx by a value sampled from a Pert distribution with minimum 0.3, mode 0.8, and maximum 1.3.
- Individual-project true CapEx and PVOCF are lognormal distributions.
- True risk: the SD of the true NPV. Since NPV is the difference between the PVOCF and CapEx, and these two parameters are independent and uncorrelated, then the SD of the true NPV, which is true risk can be calculated as follows:

$$\text{SD of true NPV} = \sqrt{\text{variance of true CapEx} + \text{variance of true PVOCF}} \quad \text{...............} \quad (3.1)$$

3.1.2.2 The Estimated Project Value Distribution

The estimated project value distributions result from a typical probabilistic assessment done in industry in limited-resource environment, and it would include the biases present in typical O&G project evaluations. To get these estimated distributions, directional bias and overconfidence are applied to the true distributions. Similarly to McVay and Dossary's (2012) model, the same amount of directional bias and overconfidence are applied to both CapEx and PVOCF. The directional bias makes the estimated PVOCF and CapEx shift in opposite directions. Hence, for the cases of optimism (positive directional bias), the estimated distribution for the PVOCF will be shifted in the positive direction while the estimated distribution for the CapEx will be shifted in the negative direction. The estimated risk is the SD of the estimated NPV (knowing that CapEx and PVOCF are considered independent and uncorrelated) and is calculated as follows:

$$\text{SD of estimated NPV} = \sqrt{\text{variance of estimated CapEx} + \text{variance of estimated PVOCF}} \quad \text{.........................} \quad (3.2)$$

3.1.2.3 The Input Parameters

The number of projects available for the portfolio optimization is set to eight. This number was chosen because it is sufficient to capture the impact of biases on portfolio optimization without exceeding the computational capacity available in the visual basic for applications (VBA) coding language. The budget used in this model is set to $400MM. Begg and Bratvold (2008) and McVay and Dossary (2012) used a budget of $5 billion for a pool of 100 projects (the average budget per project is $50MM). This model is using 8 projects, so I used the same average budget per project ($50MM*8=$400MM). In all cases that I ran, this budget limit combined with the 8 projects pool yielded portfolios that never included more than 5 projects.

3.1.3 The Simulation

The portfolio optimization was performed for three different cases: the estimated case, the best-possible case and the realized case. These cases can be described as follows:

- The best-possible case: uses the true value distributions attained in unlimited-resource environments that were explained in the previous section. Projects are selected and reported using their true expected parameters.
- The estimated case: uses the estimated value distributions resulting from typical industry limited-resources probabilistic assessments. Projects are selected and reported using their estimated expected parameters.
- The realized case: selected based on estimated value distributions but it is assessed based on true value distributions. Each portfolio goes through the following cycle: selection, development and reporting. In the realized case, project selection is done using the estimated expected value distributions, but as the selected projects are being developed, the company realizes the true cost (CapEx) of each project. If the costs of the projects selected in the portfolio exceed the available budget, then the company has to give up the projects with the lowest investment efficiency (IE). IE is calculated by dividing the estimated expected NPV by the estimated expected CapEx. The way the model simulates this process is by first selecting the corresponding optimum estimated portfolio, and comparing the true CapEx of the selected projects to the available budget. If the total CapEx of the projects in the portfolio exceeds the available budget, then

the projects are ranked by IE and selected from high IE to low IE until the available budget is exhausted; an appropriate percentage of the last project is selected to fill the budget. This process assumes that projects with the highest IE are developed first and that the estimators know the true CapEx of the selected projects from day one of the development process. The reporting process is done using the true expected parameters because the reporting step occurs after the completion of the projects and by then the company is aware of the true costs, returns and risks.

3.1.3.1 The Disappointment and The Decision Error

In the previous models, the disappointment is defined as the estimated portfolio NPV minus the realized portfolio NPV. The decision error is the best-possible portfolio NPV minus the realized portfolio NPV. The estimation error is the estimated portfolio NPV minus the best-possible portfolio NPV. Thus, the estimation error is the disappointment minus the decision error. The decision error is always positive, and it is the part of disappointment that results from selecting the wrong projects.

This thesis provides a new framework that defines a portfolio using not only the NPV but also the risk. Consequently, disappointment and decision error can occur not only because of decrease in the NPV but also because of increase in risk. Thus, disappointment and decision error can be considered as vectors in a two-dimensional space with a return component and a risk component. This will be further explained in the results section.

3.1.3.2 Monte Carlo Simulation

The portfolio optimization process uses Monte Carlo simulation to get the expected values. In each Monte Carlo iteration, true distributions of PVOCF and CapEx are randomly generated for 8 projects. Based on these true distributions, the estimated distributions are calculated and input to the model. The model runs and performs the optimization of the three different portfolios. The output of the model is the portfolio NPV, risk, expected disappointment, expected decision error, expected disappointment as a percentage of the estimated portfolio value (ED%E) and expected decision error as a percentage of the estimated portfolio value (EDE%E).

The number of Monte Carlo iterations needed to reach a stable output was determined to be 1,000. **Fig. 2** shows two full Monte Carlo simulations that use the same input parameters; each is run for 1,000 iterations. The difference between the ED%E and the EDE%E between the two simulations is about 1%.

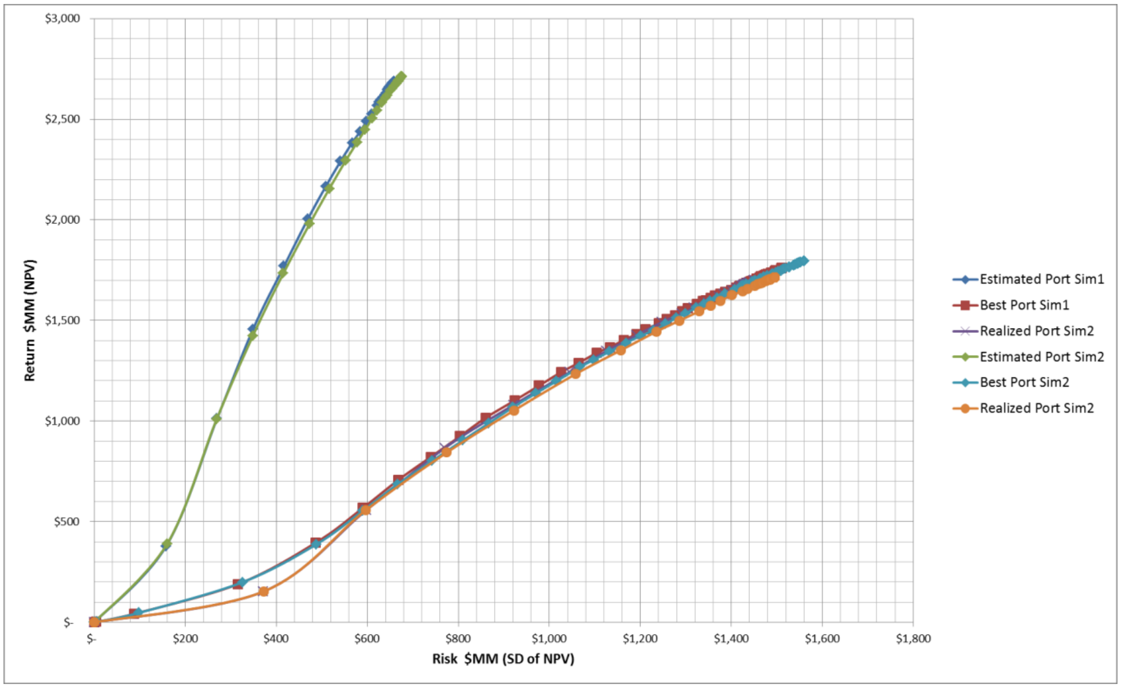

Fig. 2—Comparison of the efficient frontier curves generated from two different Monte Carlo simulations

In McVay and Dossary (2012), the ED%E is calculated by taking the expectation of the percent disappointments from the Monte Carlo iterations, as shown in the following equation:

$$\text{ED\%E}=\text{E}\left(\frac{\text{Estimated NPV - Real NPV}}{\text{Estimated NPV}}\right) \quad (3.3)$$

EDE%E is calculated using the same method. Begg and Bratvold (2008) calculated ED%E using the equation below:

$$\text{ED\%E}=\frac{\text{E(Estimated NPV) - E(Real NPV)}}{\text{E(Estimated NPV)}} \quad (3.4)$$

The difference between the two methods is relatively small; thus, I used the first method to be consistent with McVay and Dossary's (2012) work, which is the basis for my work. More detailed equations to calculate ED%E and EDE%E taking into consideration the risk factor are provided in the results section.

3.1.4 Portfolio Optimization

The portfolio optimization is done using Markowitz theory. As explained before, in each Monte Carlo iteration, true distributions of PVOCF and CapEx are randomly generated for 8 projects. Based on these true distributions, the estimated distributions are calculated and input to the model. The true and estimated expected project NPV and risk are also calculated and input to the model. **Fig. 3** is a snapshot of the table that has the global distributions stored and the expected values generated. The user inputs the available budget, the maximum risk tolerance, the number of Monte Carlo iterations, and the number of risk-tolerance increments that the user wants to see on the graph. The risk-tolerance limit is the risk limit that each portfolio should not exceed. Guerard (2009) states that the Markowitz procedure provides the optimal portfolio corresponding to the risk tolerance of any investor. The concept of risk tolerance will be explained further and illustrated with examples in the results section. The risk-tolerance limit starts from 0 and it increments until it reaches the maximum risk tolerance input by the user. The optimization is done every time the risk-tolerance limit is incremented. The risk-tolerance limit is incremented by a uniform amount; this amount is equal to the maximum risk tolerance divided by the number of risk-tolerance increments input by the user.

	TRUE EV $MM			SD of EV			Dir Bias		OverConf		Estimated EV $MM									
Name	PVCI	CapEx	NPV	PVCI	CapEx	NPV	PVCI	CapEx	PVCI	CapEx	PVCI				CapEx				NPV	SD of NPV
Description	Matched	Begg's	PVCI - CapEx	Relative to EV	Relative to EV	Relative to EV	Det	Det	Det	Det	a: Min of range	b: Max of range	Theoretical Mean	Theoretical Variance	a: Min of range	b: Max of range	Theoretical Mean	Theoretical Variance	PVCI - CapEx	
Distribution	LogNorm	LogNorm		Pert	Pert		Norm	Norm	Norm	Norm										
Multiplier																				
Mean	$ 750	$ 600					0.50	0.50	0.50	0.50										
Mode				0.8	0.8															
SD	$ 750	$ 600					0	0	0	0										
Min				0.3	0.3		-1	-1	0	0										
Max				1.3	1.3		1	1	1	1										
Shift	$ 300	$ 100																		
1	$ 384	$ 735	$ (351)	$ 279	$408	$495	0.5	-0.5	0.5	0.5	$ 252	$657	$ 403	$ 11,811,814,890	$ 354	$758	$ 549	$ 12,720,347,803	$ (146)	$ 157
2	$ 822	$ 1,097	$ (276)	$ 943	$902	$1,305	0.5	-0.5	0.5	0.5	$ 403	$1,550	$ 795	$ 91,934,125,641	$ 371	$1,066	$ 689	$ 37,694,844,137	$ 107	$ 360
3	$ 1,629	$ 568	$ 1,061	$ 964	$517	$1,094	0.5	-0.5	0.5	0.5	$1,177	$2,633	$ 1,738	$ 154,447,172,744	$ 172	$538	$ 336	$ 10,453,233,966	$ 1,401	$ 406
4	$ 743	$ 485	$ 257	$ 583	$472	$750	0.5	-0.5	0.5	0.5	$ 468	$1,296	$ 773	$ 49,262,733,503	$ 136	$451	$ 276	$ 7,750,694,562	$ 496	$ 239
5	$ 1,012	$ 673	$ 338	$ 1,008	$563	$809	0.5	-0.5	0.5	0.5	$ 740	$1,622	$ 1,081	$ 56,828,906,606	$ 224	$652	$ 419	$ 14,292,490,318	$ 662	$ 267
6	$ 1,391	$ 942	$ 449	$ 1,335	$933	$1,628	0.5	-0.5	0.5	0.5	$ 776	$2,543	$ 1,403	$ 221,116,125,401	$ 259	$871	$ 531	$ 29,288,045,815	$ 872	$ 500
7	$ 611	$ 362	$ 248	$ 700	$279	$754	0.5	-0.5	0.5	0.5	$ 300	$1,151	$ 591	$ 50,735,087,967	$ 131	$357	$ 235	$ 3,983,069,256	$ 356	$ 234
8	$ 1,062	$ 673	$ 390	$ 845	$426	$947	0.5	-0.5	0.5	0.5	$ 665	$1,860	$ 1,104	$ 102,542,033,498	$ 291	$684	$ 477	$ 12,011,003,418	$ 626	$ 338

Fig. 3—True and estimated expected values

Using all the parameters input and generated, the code examines the different combinations of projects that make portfolios with up to 8 projects. All the possible portfolios are generated once, and then an optimization function runs through them all to determine the portfolio that returns the highest NPV, fully uses the available budget, and does not exceed the risk-tolerance limit. To fully use the available budget, a fraction of the last project is added. Each portfolio can have up to one partial project. Projects with negative estimated expected NPV are automatically excluded by the model; projects with a negative true expected NPV are not excluded. These specifications are inherited from the Begg and Bratvold (2008) and McVay and Dossary (2012) models. This work uses a particular set of global portfolio parameters that was described earlier in this section. The results and conclusions of this model can be different if using another set of global portfolio parameters.

3.1.5 The Algorithm

The flowcharts included in this section explain the algorithm used to implement the model. The code is organized into four functions— Set_Matrix function, Get_Input function, Optimize function and Get_Mean_Values function. Each of these functions performs a different task that will be explained in the following sections. All of these functions are called from the main function. The symbols used in these flowcharts are explained in :

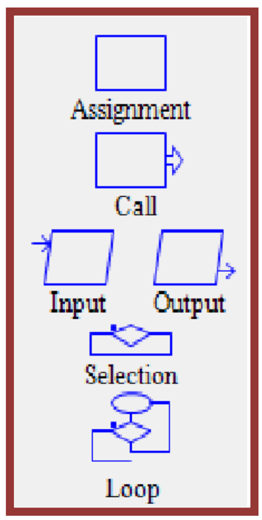

Fig. 4—Symbols used in the flowcharts

An assignment is when a value is assigned to a variable. A call is used when a function is needed and it is located outside the calling function. Selection is used for the case of an “If else statement.” The term “Put” used in these algorithms means that an output is returned.

3.1.5.1 The Main Function

The flowchart in **Fig. 5** is illustrating the algorithm of the main function. This function takes, as input, the number of Monte Carlo iterations, the maximum risk tolerance, the available budget and the desired number of risk-tolerance increments. A counter is a variable that counts the number of iterations. When the code starts the counter is assigned a value of 1; it is then incremented by 1 at the end of each iteration. The Set_Matrix function is called to set the different combinations of 0s and 1s needed for generating the initial set of possible portfolios; the process is explained in the next section. Before entering the loop, the code verifies that the counter has a value below the required number of Monte Carlo iterations. If this condition is satisfied, the Get_Input function is called and the optimization function for each of the three portfolios (estimated, best-possible and realized) is called. The counter is then incremented. If the counter exceeds the number of Monte Carlo iterations, then the loop is left and Get_Mean_Values function is called to return the mean, or expected values of the outputs. Each iteration of the loop represents a full Monte Carlo iteration, because each loop gets a new sample of projects’ parameters, performs the optimization of the three portfolio types for the different risk tolerance increments, and returns the parameters of the three efficient frontiers.

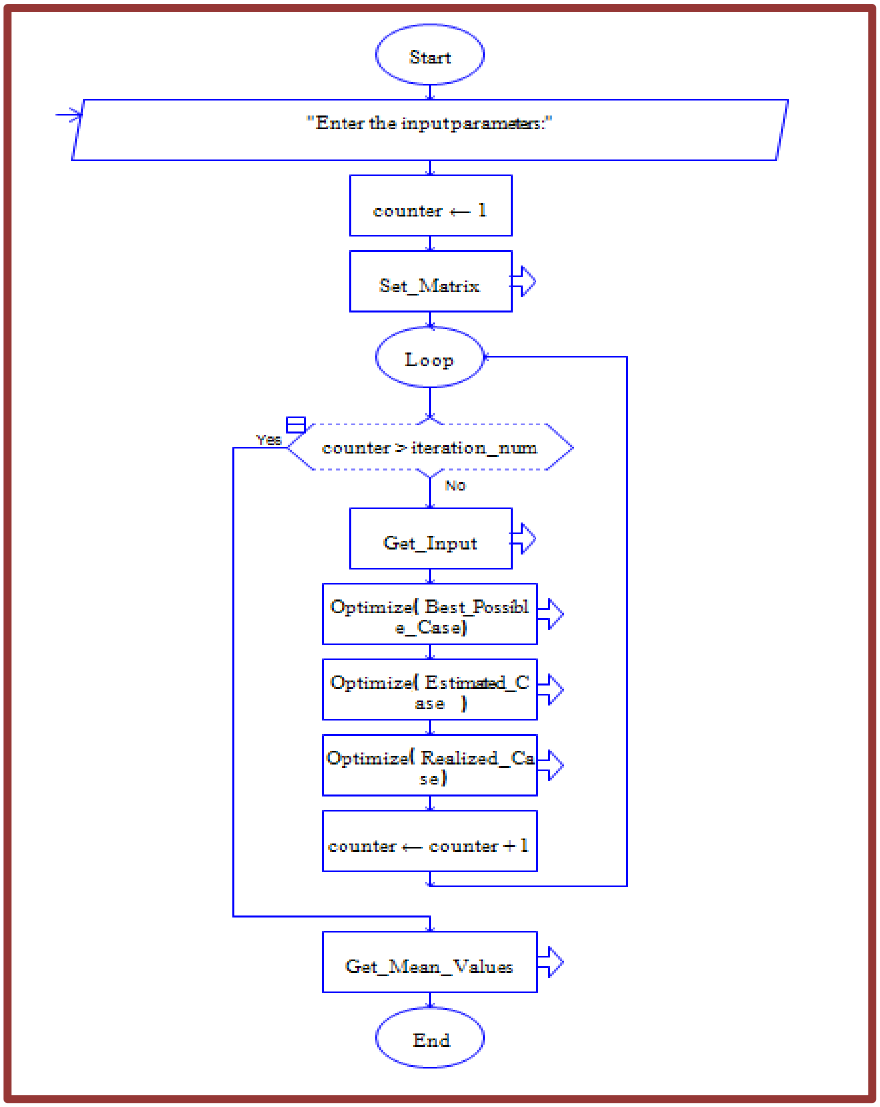

Fig. 5—The algorithm for the main function

3.1.5.2 Set_Matrix Function

The Set_Matrix function builds a two-dimensional array, with 8 columns and 256 rows, to set up the initial set of possible portfolios. A portfolio can include up to 8 projects. A project that is included in the portfolio is represented by a value of 1, a project that is not included in the portfolio is represented by a value of 0, and a project that is partially included is represented by a fraction greater than 0 and below 1. Initially, the matrix has only values of 0 and 1; the second step is to determine the projects fractions to include partial projects and fully use the available budget.

The Set_Matrix function generates the matrix shown in **Fig. 6**. This matrix is stored in a two-dimensional array with 8 columns and 256 rows. The number 256 is the result of 2^8 because a portfolio can have up to 8 projects and each project can have 2 states; it can be either included in the portfolio and take a value of 1, or not included in the portfolio and take a value of 0. Hence, the matrix is populated by the participation levels of projects (the values of 1 and 0).

0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	1
0	0	0	0	0	0	1	0
0	0	0	0	0	1	0	0
:	:	:	:	:	:	:	:
1	1	1	1	1	1	0	1
1	1	1	1	1	1	1	0
1	1	1	1	1	1	1	1

Fig. 6—The matrix generated by the "Set_Matrix" function

The different combinations of portfolios generated at this level provide an initial set of possible portfolios. These portfolios are labeled as initial because an extra step will be performed on these portfolios to include partial projects and make sure all the portfolios fully use the available budget. **Fig. 7** shows the algorithm for the Set_Matrix function. It is a nested loop, with all of the 8 variables starting with a value of 0. Then each of these values is incremented to have a value of 1. The increment occurs for one variable at a time. After each variable change, a different combination of variables is generated and stored as a row in the two-dimensional array. This process ensures 256 combinations with no redundancies.

3.1.5.3 Get_Input Function

The Get_Input function shown in **Fig. 8** reads the true and estimated expected values that are stored in the table shown on **Fig. 3**. For each Monte Carlo iteration, the expected values in the table **(Fig. 3)** are updated as they are randomly sampled from the global distributions described in the previous section. Thus, this function is called at the beginning of each Monte Carlo iteration to provide data for the optimization.

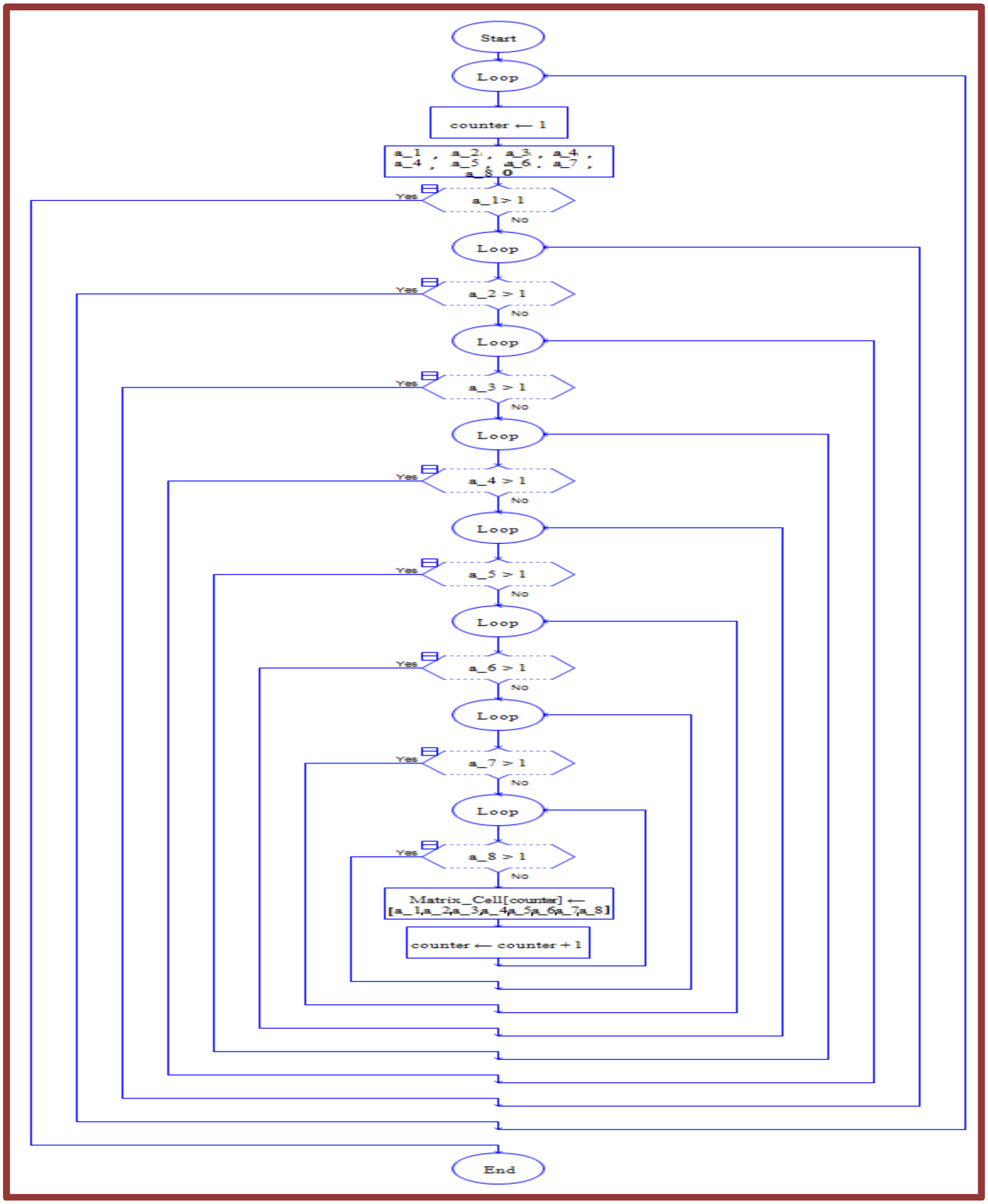

Fig. 7—The “Set_Matrix” algorithm

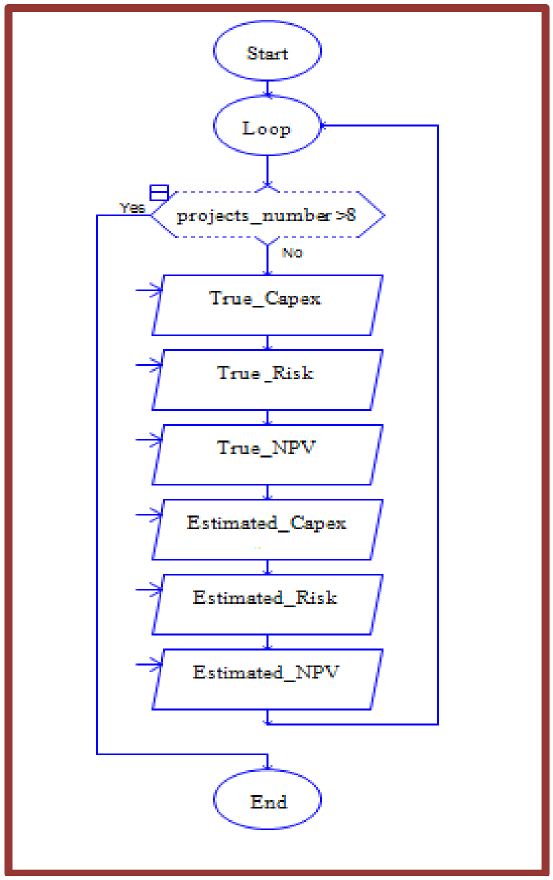

Fig. 8—The "Get_Input" algorithm

3.1.5.4 Optimize Function

The Optimize function performs the optimization of the portfolios at each particular risk-tolerance limit. It returns the portfolio that has the highest NPV, fully uses the available budget, and does not exceed the risk-tolerance limit. This function is designed in a generic way so that it can perform the optimization for any of the three types of portfolios (estimated, best-possible and realized). If true values are passed to the Optimize function, then it will optimize the best-possible portfolios; if estimated values are passed to this function, then it will optimize the estimated portfolios. For the realized portfolios, optimization is not needed; this function will use the corresponding optimum estimated portfolios, impose the budget constraint on their true CapEx and return the true expected NPV, risk and CapEx. Since this function is generic, the rest of this section will use the parameter names (NPV, risk, CapEx) without specifying if they are true or estimated.

The Optimize function generates the initial set of possible portfolios (provides the portfolios NPV, CapEx and risk). It then adds partial projects to these portfolios to fully use the available budget. The last step is to select the optimum portfolios that are within the risk-tolerance limit. **Fig. 9** illustrates the major steps in the Optimize function.

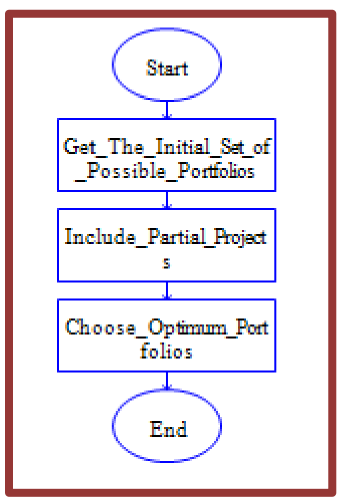

Fig. 9—The three main steps in the algorithm of the "Optimize" function

NPV of project 1	NPV of project 2	NPV of project 3	NPV of project 4	NPV of project 5	NPV of project 6	NPV of project 7	NPV of project 8
1,200	853	1598	204	2,015	122	1,000	3,500

Multiplied by

0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	1
0	0	0	0	0	0	1	0
0	0	0	0	0	1	0	0
:	:	:	:	:	:	:	:
1	1	1	1	1	1	0	1
1	1	1	1	1	1	1	0
1	1	1	1	1	1	1	1

Is equal to

Portfolio 1	0	0	0	0	0	0	0	0
Portfolio 2	0	0	0	0	0	0	0	3,500
Portfolio 3	0	0	0	0	0	0	1,000	0
Portfolio 4	0	0	0	0	0	122	0	0
:	:	:	:	:	:	:	:	:
Portfolio 254	1,200	853	1598	204	2,015	122	0	3,500
Portfolio 255	1,200	853	1598	204	2,015	122	1,000	0
Portfolio 256	1,200	853	1598	204	2,015	122	1,000	3,500

Fig. 10—The multiplication of the two-dimensional array by the input NPV data to generate the initial set of possible portfolio NPVs

The first step in the algorithm of the Optimize function is to generate the initial set of possible portfolios using the output of both the Set_Matrix and Get_Input functions. The Set_Matrix function provides a matrix that stores the participations levels of the projects (**Fig. 6**). The Get_Input function provides the input data of the projects. The Optimize function multiplies the matrix by the input data stored in arrays (an array of the projects' NPV, an array of the projects' CapEx and an array of the projects' risk).

This multiplication is done inside the Optimize function and it has the purpose of generating the NPV, CapEx and risk of the different possible portfolios. **Fig. 10** is an example of the multiplication process of the matrix by the NPV data.

The portfolio NPV is the sum of the projects' NPVs multiplied by their level of participation in the portfolio (0 for projects not included, 1 for projects fully included, and a fraction between 0 and 1 for projects partially included). The portfolio CapEx is similarly calculated by summing up the projects' CapEx values multiplied by their level of participation. According to Markowitz theory, the portfolio risk is the standard deviation of the portfolio NPV (Guerard, 2009); for the case of a portfolio with two projects, it is calculated as follows:

$$\text{Portfolio Risk}=\sqrt{(x_1\sigma_1)^2+(x_2\sigma_2)^2+2(x_1\sigma_1)(x_2\sigma_2)\rho} \quad (3.5)$$

Where x_i and x_2 are the participation levels of project 1 and 2, σ_1 and σ_2 are the risk (standard deviation of the NPV) of each project and ρ is the correlation between the two projects. The projects used in this model are randomly sampled from a distribution and assumed independent and uncorrelated, meaning $\rho=0$. The multiplication process shown in **Fig. 10** is repeated for the CapEx and the risk variables, and the result is three matrices of NPV, CapEx and risk for the same initial set of portfolios. To keep track of these portfolios a consistent array index is used for all arrays. For example, index 254 refers to the same portfolio in the NPV matrix, the CapEx matrix and the risk matrix. An extra column is added to the matrix in **Fig. 10** to store the portfolio NPV. The CapEx and risk matrices also have a 9th column added to store the portfolio CapEx and Risk.

Hence the portfolio parameters as well as the details of the projects are stored in the matrices. This process is described in the first part of the Optimize function as illustrated in the flowchart in **Fig. 11**.

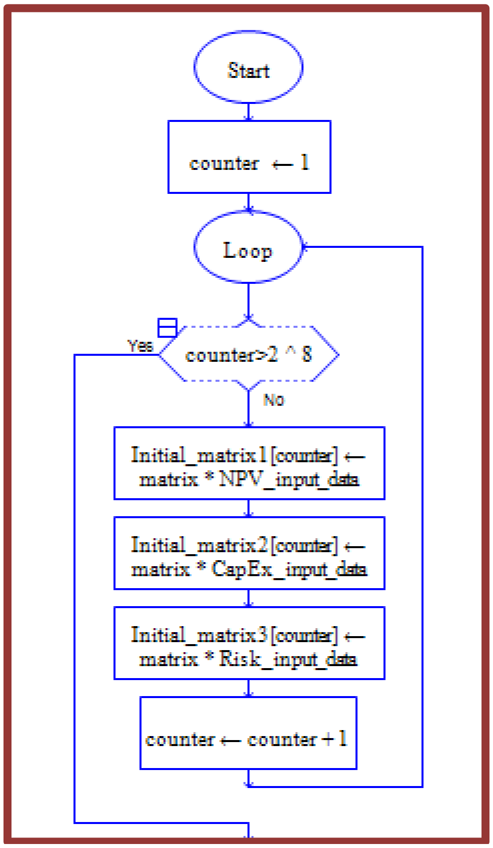

Fig. 11—The first part of the "Optimize" function that generates the initial set of possible portfolios

The next step is to create a new matrix that only contains the portfolios that fully use the budget. This new matrix excludes the portfolios that require a budget higher than the available budget. Portfolios that require a budget less than the available budget are subject to another step before they are added to this new matrix. This step consists of adding a partial project (a fraction of an unselected project) to these portfolios. This fraction is calculated by dividing the unused part of the budget by the CapEx of the projects not included in that specific portfolio, one at a time, because a portfolio can include only one partial project. Thus, if a portfolio includes 5 projects and is not fully using the available budget, the code first calculates the unused budget (available budget - portfolio CapEx) and divides it by the first unselected project to get the fraction of that project that is added to the portfolio to get one possible portfolio. It then divides by the CapEx of the second unselected project to get another possible portfolio and so on.

The fraction that the code calculates (unused budget divided by the project CapEx) gives us the fraction of the partial project that can be added to the portfolio to fully use the budget. This fraction is then multiplied by the corresponding project NPV and risk to calculate how much NPV and risk this partial project adds to the portfolio. Once this process is completed, all the possible portfolios are ready and stored in the new matrix. This process of generating the possible portfolios that fully use the available budget is illustrated in **Fig. 12**.

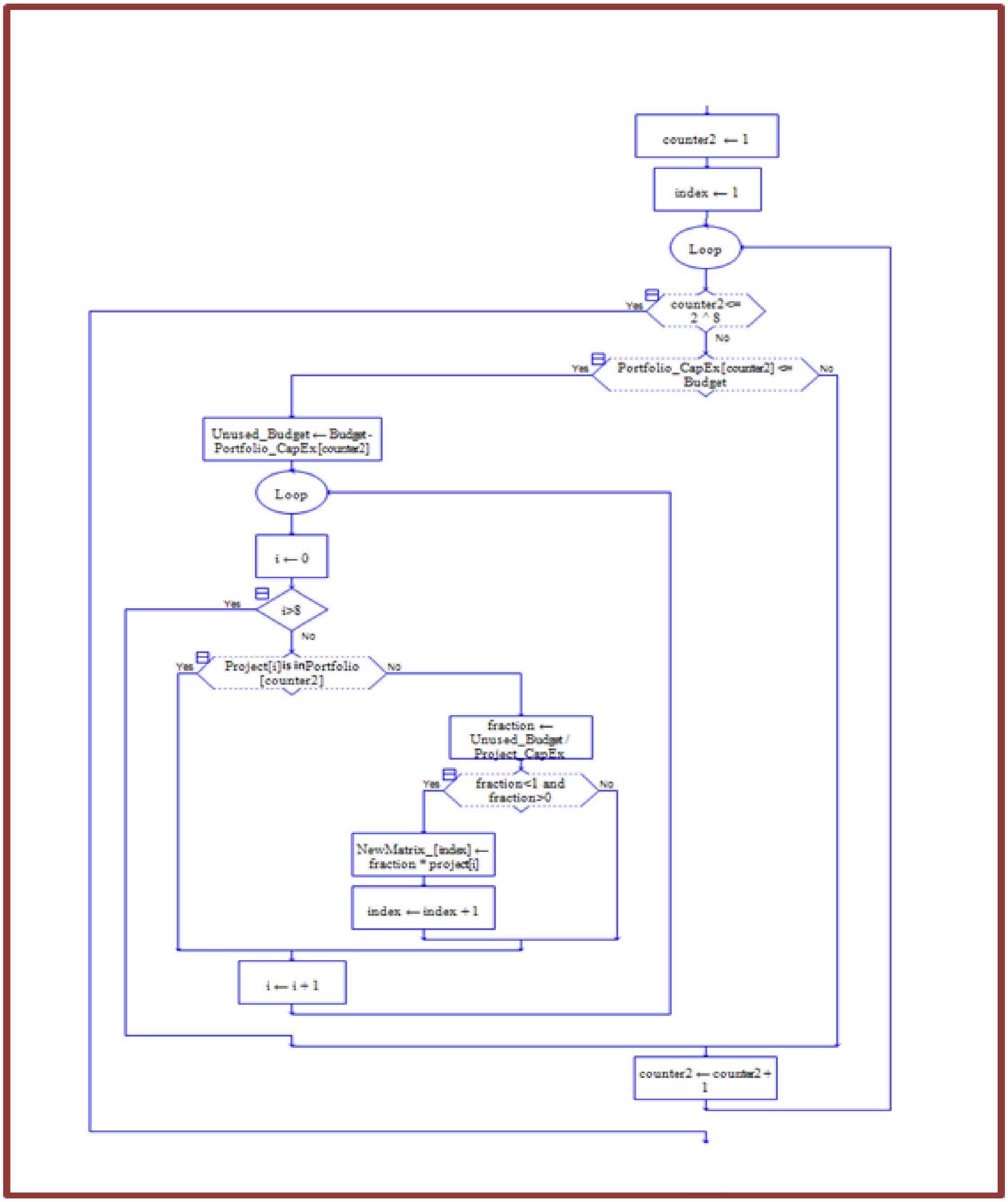

Fig. 12—The second part of the "Optimize" function that generates all the possible portfolios that fully use the budget

The last step in the Optimize function is to search for the portfolio with the highest NPV within the risk-tolerance limit after each risk tolerance increment. The risk-tolerance limit starts from 0 and it is incremented by a uniform amount until it reaches the maximum risk-tolerance limit input by the user. This uniform amount is equal to the maximum risk-tolerance limit divided by the number of risk-tolerance increments that the user inputs. For example, if the user inputs a maximum risk limit of $2,500MM and inputs a number of risk-tolerance increments of 100, then the increment will be $25MM. After each risk-tolerance increment, the Optimize function will search for the portfolio with the highest NPV (from the set of possible portfolios that are already generated and stored in the previous steps) within the incremented risk-tolerance limit. Hence all the possible portfolios are generated only once per Monte Carlo iteration and the optimization is done from this pool of possible portfolios after each increment of the risk-tolerance limit. **Fig. 13** is a flowchart that illustrates this last step in the optimize function algorithm.

As mentioned before, the Optimize function is designed in a generic way, so that it can perform the optimization for any of the three types of portfolios (estimated, best-possible and realized). If true values are passed to this function, then it will determine optimal best-possible portfolios; if estimated values are passed, then it will determine optimal estimated portfolios.

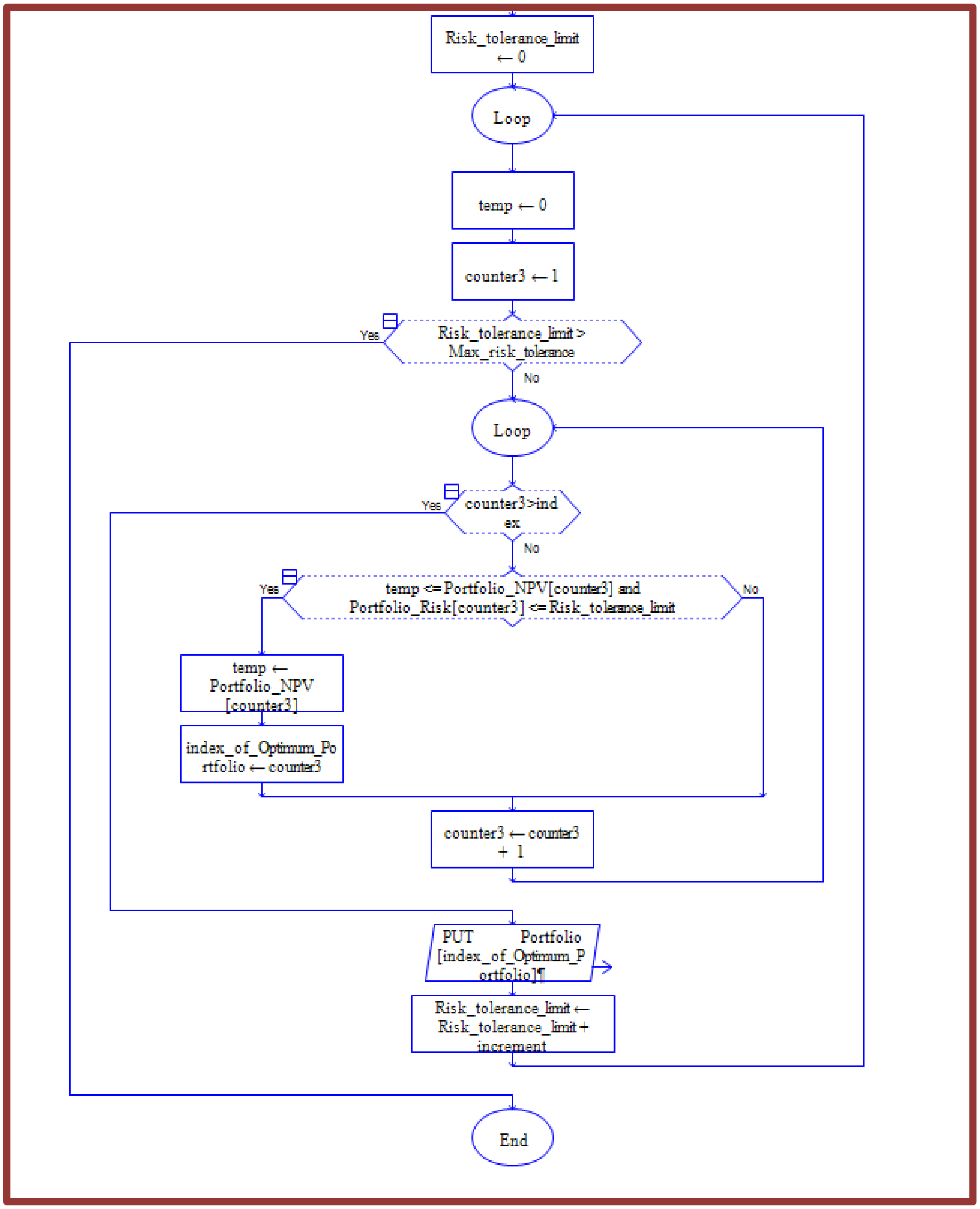

Fig. 13—The third part of the "Optimize" function that selects the optimum portfolios

For the realized portfolios, optimization is not needed. The function will use the corresponding optimum estimated portfolios, then calculate realized portfolio results using true project CapEx values (as budget constraints) and true NPV and risk values (to calculate portfolio performance) (**Fig. 14**). Similarly to McVay and Dossary's (2012) approach for the realized portfolio, this function checks if the optimum estimated portfolio has a true CapEx that is higher than the available budget. Checking this condition is a way of simulating what happens in reality. In companies, the decision to develop a particular portfolio is made based on estimated values, but if the true CapEx of the selected projects turns out to be higher than the available budget (which is usually revealed early in the development), then the company either will not develop some projects or will reduce participation in some projects to stay within their capital budget. To determine which projects to forego or reduce participation, the projects are ranked by investment efficiency (IE),

$$IE=\frac{\text{Estimated NPV}}{\text{Estimated CapEx}} \quad (3.6)$$

Then projects are selected from highest to lowest investment efficiency until the true CapEx of the selected projects is equal to the available budget. A fraction of the last project is included to fully use the budget. **Fig. 14** illustrates the steps of this selection of the realized portfolio.

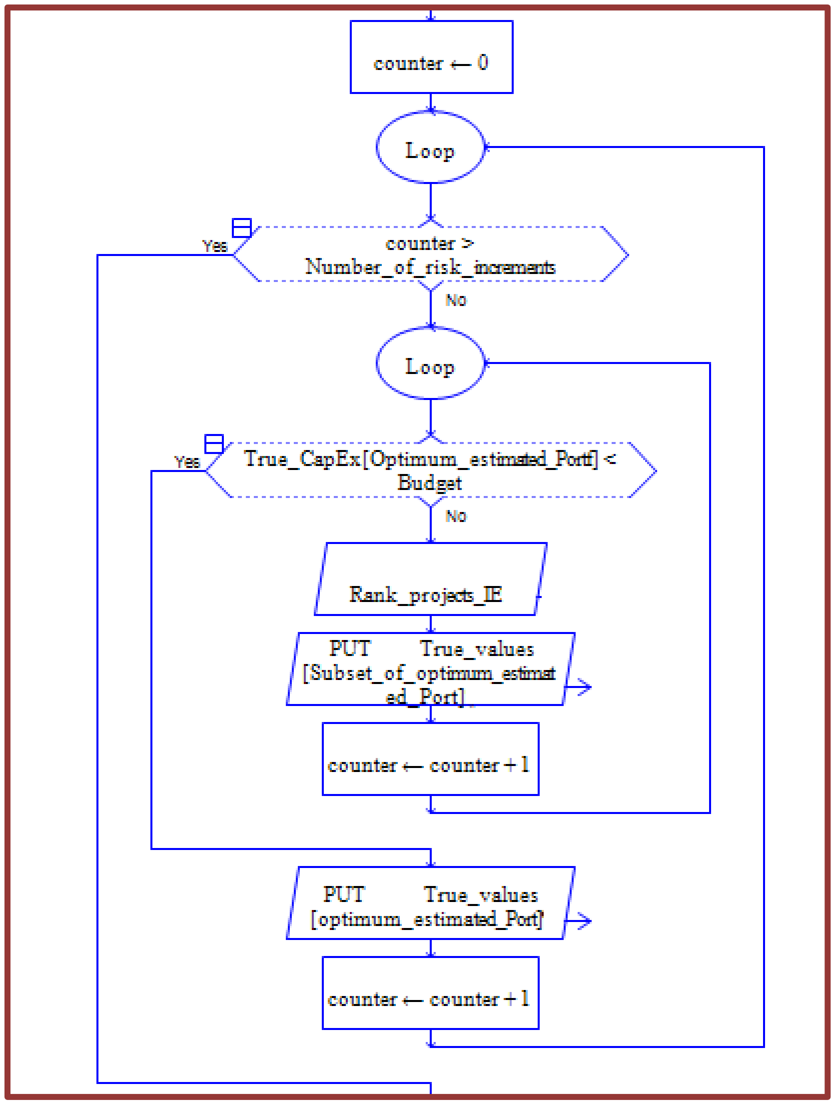

Fig. 14—The last part of the "Optimize" function that selects the realized portfolios

3.1.5.5 Get_Mean_Values Function

After each Monte Carlo iteration, the code generates the NPV and risk of the optimum portfolios. All these values are stored in arrays. The Get_Mean_Values function is called at the end of the simulation. This function goes through the stored values and returns the mean or expected values that are then printed to tables and plotted on graphs.

4. RESULTS

4.1 The Case of Moderate Overconfidence and Moderate Optimism

I ran the model for the case of moderate Overconfidence (0.5) and moderate Optimism (positive directional bias of 0.5). The other input variables used for this run were as follows: a budget of $400 MM, a maximum risk-tolerance limit of $50,000 MM, 1,000 Monte Carlo iterations and 500 risk-tolerance increments. This means at each of these 500 increments, the model will return the portfolio with the highest NPV and with a risk below or equal to the risk-tolerance limit. The risk-tolerance limit starts from 0 and is incremented until it reaches the maximum risk-tolerance limit ($50,000MM in this case) input by the user. The risk-increment amount is uniform, and it is equal to the input maximum risk-tolerance limit divided by the total number of risk-tolerance increments; in this case, the risk-increment amount is equal to $100MM ($50,000MM divided by 500).

4.1.1 Expected Efficient Frontier Curves

Fig. 15 shows the expected efficient frontier curves for the estimated portfolio, the realized portfolio and the best-possible portfolio. The curves were run as far out as they can go by setting a high maximum risk-tolerance limit; this case can be referred to as the unlimited-risk case. The curves are expected curves because they have been generated using Monte Carlo Simulation; they represent the mean of 1,000 runs. At different values of risk on the x-axis, the curves show the highest NPV for each of the three portfolios. These curves are plotted against the calculated expected portfolio risk as opposed to the risk-tolerance level. The estimated portfolio NPV is plotted against the estimated portfolio risk, while the realized and the best-possible portfolio NPVs are plotted against the true portfolio risk.

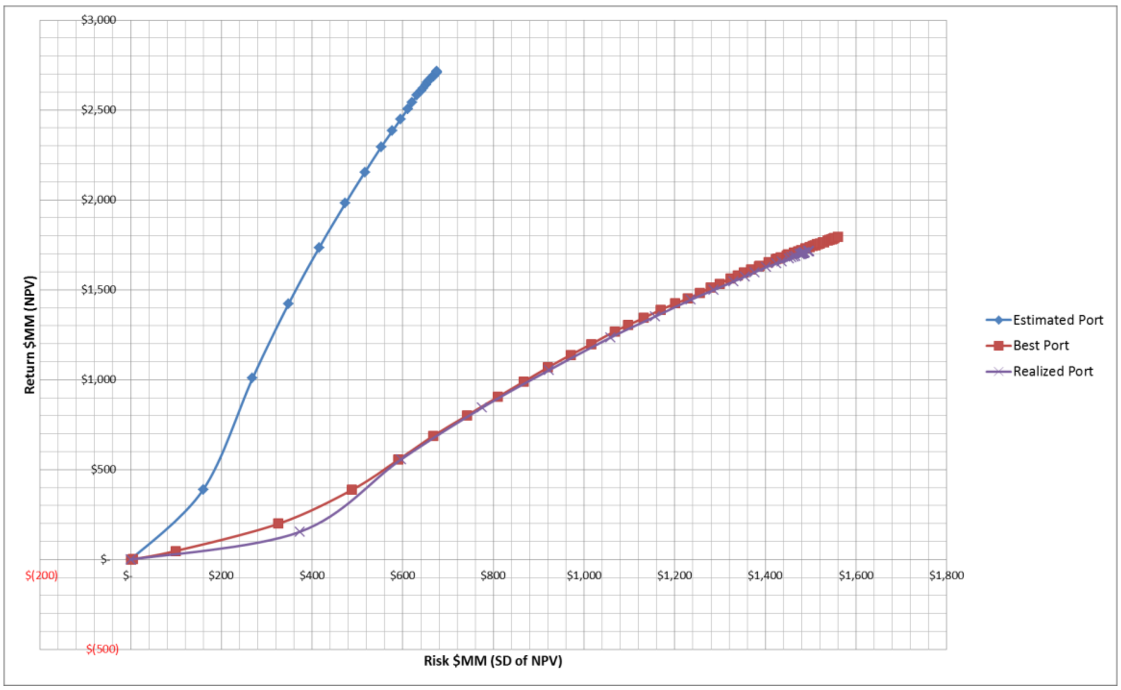

Fig. 15—Expected efficient frontier curves for 0.5 overconfidence & 0.5 directional bias

The concept of risk tolerance was documented in the literature by a number of authors; they stated that the risk tolerance of the decision-maker is the main criterion for the choice of the optimum portfolio (Brashear et al., 2001; Guerard, 2009). Hence, the risk tolerance is the common factor between the three different portfolios. For each specific risk-tolerance limit, the model returns the optimum best-possible expected portfolio and the optimum estimated expected portfolio along with its corresponding realized expected portfolio.

The curve for the estimated portfolio is far above the best-possible and the realized portfolios' curves; this is explained by the impact of optimism and overconfidence on the estimations of CapEx, NPV and risk. Following McVay and Dossary's (2012) model, the same amount of directional bias and overconfidence is applied to both CapEx and PVOCF. As noted earlier, directional bias shifts PVOCF and CapEx in opposite directions. For the optimistic case, the estimated distribution for the PVOCF is shifted in the positive direction and the CapEx distribution is shifted in the negative direction. Thus, portfolio CapEx is underestimated and portfolio NPV (i.e., PVOCF minus CapEx) is overestimated. The estimated risk (the standard deviation of the estimated NPV) is calculated using the standard deviations of the estimated distributions of CapEx and PVOCF; these estimated distributions are obtained by truncating true distributions. A truncated distribution has a lower standard deviation than the original distribution. Thus, the portfolio risk is underestimated. Compounding the situation, more projects can be fit in the estimated portfolio because the portfolio risk and CapEx are underestimated.

Conversely, the best-possible portfolio has a lower number of projects, a lower NPV and higher risk. This is because true distributions are used and, thus, the portfolio CapEx and risk are not underestimated, and the portfolio NPV is not overestimated.

The projects for the realized portfolio are selected based on the estimated CapEx and PVOCF distributions, but the portfolio is developed and assessed using the true distributions. The realized portfolio has the same projects as the estimated portfolio, or a subset of these projects in the event the true CapEx of the selected projects exceeds the available budget. The realized portfolio curve is either equal to or below (to the right of) the best-possible portfolio curve because (1) the best-possible portfolio curve is the highest portfolio NPV that could be realized using the true distributions of CapEx and PVOCF, and (2) the realized portfolio is selected based on the estimated risk (which is underestimated) and is assessed based on the true risk (which is higher).

There is a difference between the risk tolerance and the portfolio risk. The portfolio risk is the risk that is calculated using the risks of the projects included in the portfolio. The risk tolerance is the risk criterion used to choose the optimum portfolio (Brashear et al., 2001; Guerard, 2009). In the cases studied, the portfolio risk (whether it is the risk of the estimated portfolio or the risk of best-possible portfolio) is usually lower than the risk-tolerance limit, and they are rarely equal. The difference between the risk tolerance and the portfolio risk (the estimated and best-possible portfolio risk) can be explained with the following example. A company might tolerate a risk that goes up to $100MM, but when optimizing a portfolio from a pool that contains a discrete set of projects, the optimum estimated portfolio might have a risk level that is equal to $55MM, which is lower than the risk-tolerance limit by $45MM. With the set of conditions and assumptions that I used, it is rare when the portfolio risk is exactly equal to the risk-tolerance limit, because (1) the optimization fills up the budget limit and not the risk limit; (2) the optimization uses a discrete set of projects (8 in this case) and can only include one partial project. The example shown in **Fig. 16** illustrates the difference between the portfolio risk and the risk tolerance, explains how this difference is magnified when considering the expected values, and sets the stage for **Fig. 17**. **Fig. 16** shows the risk of optimum portfolios selected in Monte Carlo iterations (to simplify the example, I assume that the Monte Carlo simulation has only two iterations, three risk tolerance limits, and the values were arbitrary selected for this example). The last row of the table gives the expected portfolio risk (the mean).

Risk tolerance	$25MM	$50MM	$100MM
Portfolio risk (Monte Carlo iteration 1)	$0MM	$0MM	$55MM
Portfolio risk (Monte Carlo iteration 2)	$10MM	$40MM	$40MM
Expected Portfolio risk	$5MM	$20MM	$47.5MM

C B A

Fig. 16—Example to illustrate the difference between risk tolerance and portfolio risk

Fig. 16 shows that for a risk tolerance of $100MM, the expected value of the optimum portfolios is referred to as A, and it has an expected risk of $47.5MM. For a risk tolerance of $50MM, the expected value of the optimum portfolios is referred to as B, and it has an expected risk of $20MM. For a risk tolerance of $25MM, the expected value of the optimum portfolios is referred to as C, and it has an expected risk of $5MM. The expected value A is optimum for the risk tolerance of $100MM and not $47.5MM; if the risk tolerance is set to $47.5 then the expected value of the optimum portfolio will be different, and if the risk tolerance is set to $50 then the expected value of the optimum portfolio will be B. The table shows that in some Monte Carlo iterations, even if the risk tolerance increases the same optimum portfolio is returned; this happens often because (1) the models uses a discrete set of 8 projects and only one partial project can be added; (2) the optimization fills up the budget limit and not the risk-tolerance limit. Hence, the difference between risk tolerance and the portfolio risk can be large. The expected value of the optimum portfolios A has an expected risk of $47.5MM for a risk tolerance of $100MM, which can create confusion because $47.5 is lower than the risk tolerance of $50MM. Even if the expected value A has an expected risk lower than the risk tolerance of $50MM, the accurate risk tolerance of A is $100MM, because $100MM is the limit that was imposed on the optimization (having a risk tolerance of $100MM allows the selection of the portfolio with the $55MM while a risk tolerance of $50MM would not have allowed it). Similarly, even if the risk level of B (that is $20MM) is below the $25MM risk tolerance level, the accurate risk tolerance of B is $50MM. This is due to (1) the use of a low number of projects; (2) the discrete nature of the results, and (3) the constraint of using one partial project.

Fig. 17 shows the expected efficient frontier curves for the estimated, the best-possible and the realized portfolios for the case of 0.5 overconfidence and 0.5 directional bias. For the purpose of illustrating expected disappointment and expected decision error calculations, the curves in **Fig. 17** were stopped at the risk-tolerance level after which further increase in the risk tolerance will not increase the expected NPV and risk of the estimated expected portfolio by much (further increase in the risk tolerance will increase the expected NPV of the estimated expected portfolio by less than 0.5%). Increasing the risk-tolerance further will not significantly increase the estimated expected NPV and risk, while it will significantly increase the expected NPV and risk of the best-possible expected portfolio, which will make the best-possible expected curve go further from the estimated expected curve. For the case of the example shown in **Fig. 17**; the risk tolerance at which the curves stop is equal to $1,325MM, and the expected disappointment and expected decision error are calculated for a risk tolerance of $600MM. As explained before, the risk tolerance is the main criterion for the selection of optimum portfolios, so the risk tolerance is the common factor between the three different portfolios (estimated, best-possible and realized). For each specific risk-tolerance limit and at each Monte Carlo iteration, the model returns the best-possible portfolio and the optimum estimated portfolio along with its corresponding realized portfolio. Expected disappointment and expected decision error are calculated using the expected values of NPV and risk. The example in **Fig. 17** shows the expected value of the optimum portfolios for the three types of portfolios—the estimated, best-possible and realized portfolios (points E, B and R) for a risk tolerance limit of $600MM. This means, as explained in the algorithms, that $600MM is the risk tolerance limit capping the optimization of the estimated and best-possible portfolios at each of the Monte Carlo iterations (as opposed to capping the estimated expected and best-possible expected portfolios). Point E in **Fig. 17** refers to the estimated expected NPV and estimated expected risk for a risk tolerance of $600MM. Hence, at point E, the estimated expected NPV is $1,945MM and the estimated expected risk is $465MM. As explained in the example of **Fig. 16**, \$465MM is the expected risk and it is different from the risk tolerance. If the risk tolerance is set to \$465MM, the expected value of the optimum portfolio will not be point E; it will be another point on the estimated curve with a lower estimated expected NPV and risk (this is similar to the example in **Fig. 16** where A is the expected value of the optimum portfolio for \$100MM risk tolerance and not for \$47.5MM risk tolerance)**.** Point R refers to the realized expected NPV and risk that corresponds to point E. The realized expected NPV at point R is \$1,194MM and the realized expected risk is \$1,034MM. For the same risk-tolerance level of \$600MM, point B refers to the best-possible expected NPV and risk. At point B, best-possible expected NPV is \$397MM and the best-possible expected risk is \$488MM. The difference between the expected risk at B (\$488MM) and the risk tolerance (\$600MM), which can create confusion, has been explained earlier and in the example of **Fig. 16** (the case of the expected value A that although it has an expected risk lower than the risk tolerance of \$50MM, the accurate risk tolerance of A is \$100MM). It is because (1) the optimization uses a low number of projects; (2) the optimization completely fills up the budget limit and not the risk limit, and (3) the constraint of using one partial project. Plotting the expected curves versus the risk tolerance will not clarify this confusion, and it will deviate from the principle of efficient frontier curves. The best-possible curve can go further to the right as the risk tolerance increases, but as was stated in the beginning of this paragraph, for the purpose of illustrating expected disappointment and expected decision error calculations, all the curves in **Fig. 17** were stopped at the risk-tolerance level at which the estimated curve stops, which is \$1,325MM in this case.

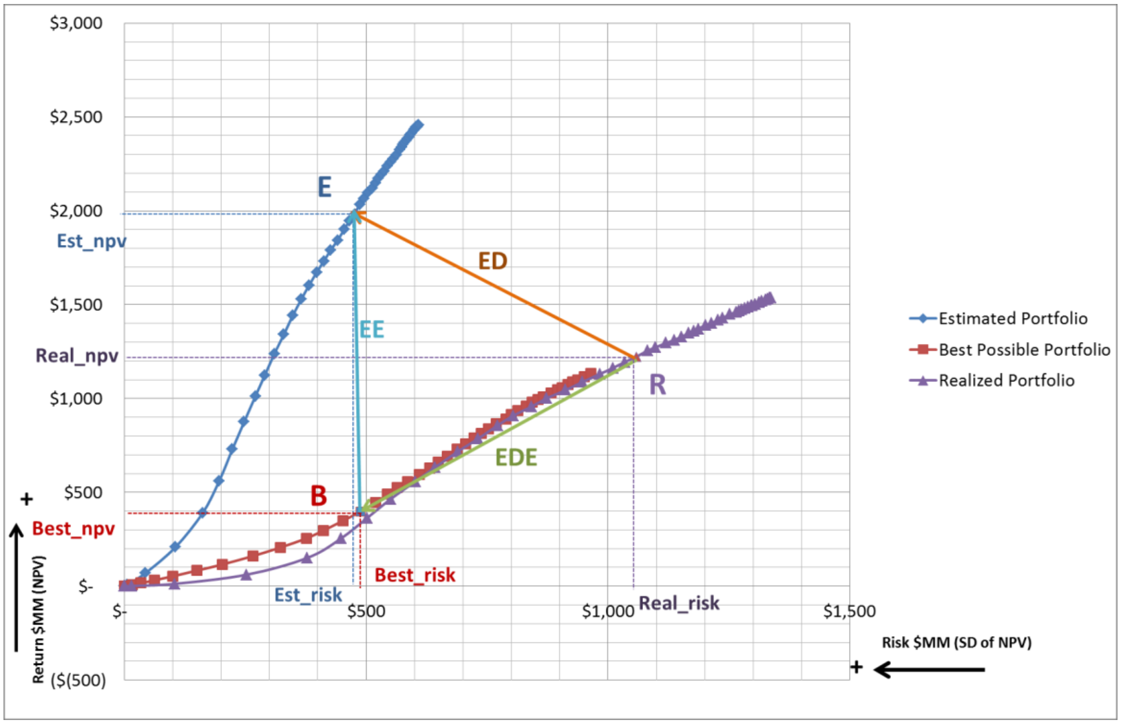

Fig. 17—Expected efficient frontier curves for 0.5 overconfidence & 0.5 directional bias and a maximum risk tolerance of $1,325MM; ED, EDE and EE are calculated for a risk tolerance of $600MM

The estimated expected risk and the best-possible expected risk are lower than the risk-tolerance limit, while the realized expected risk exceeds this risk-tolerance limit. This is because the optimization of the estimated portfolio and the best-possible portfolio is constrained by a risk-tolerance limit and a budget limit, while the realized portfolio is assessed based on the true risk (which is higher than the estimated risk and which was not considered in the project selection). As explained before, the projects for the realized portfolio are selected based on the estimated CapEx, PVOCF and risk, but the portfolio is developed and assessed using the true CapEx, PVOCF and risk. The true CapEx used to assess the realized portfolio is constrained by the budget limit but the true risk is not constrained by the risk-tolerance limit. Hence, the realized portfolio is selected based on the estimated risk (which is underestimated), and it is assessed based on the true risk (which is higher and not capped by the risk-tolerance limit).

4.1.2 Expected Disappointment

In previous work, the expected disappointment (ED) has been defined as the NPV of the estimated portfolio minus the NPV of the realized portfolio. In this work, Markowitz theory is used, so portfolios are optimized based on not only the NPV but also the risk. Consequently, disappointment can occur not only because of decrease in NPV but also because of increase in risk. Thus, in this work disappointment is considered a vector in a two-dimensional space with a NPV component and a risk component. NPV and risk are two different quantities with the same unit ($). This approach assumes that $1 of NPV is equivalent to $1 of risk, but based on how much a company values $1 of risk, the unit vectors of the two-dimensional Cartesian coordinate system can be adjusted. Expected disappointment is calculated using the expected values of risk and NPV. Using the example illustrated in **Fig. 17**, the ED for a $600MM risk-tolerance limit is the vector $\overrightarrow{RE}$. The components of $\overrightarrow{RE}$ can be determined as follows:

$$\overrightarrow{\mathrm{RE}} = \begin{pmatrix} \mathrm{risk_E\text{-}risk_R} \\ \mathrm{NPV_E\text{-}NPV_R} \end{pmatrix} \qquad (4.1)$$

These vectors can be expressed using unit vectors. The unit vectors in a standard 2D Cartesian coordinate system are $\vec{\imath}$ and $\vec{\jmath}$ and their magnitude is 1 unit length. In this framework, the unit vectors are denoted as $\overrightarrow{\mathrm{risk}}$ on the x-axis and $\overrightarrow{\mathrm{NPV}}$ on the y-axis, and their magnitude is \$1MM. The direction of the $\overrightarrow{\text{risk}}$ is from right to left, because the lower the risk, the better. The direction of $\overrightarrow{\text{NPV}}$ is to toward greater NPV. So the expected disappointment can be expressed as follows:

$$\overrightarrow{\text{ED}}=(\text{Real Risk-Estimated Risk})\overrightarrow{\text{risk}}+(\text{Estimated NPV-Real NPV})\overrightarrow{\text{NPV}} \quad \text{............} (4.2)$$

The disappointment is reported as the magnitude of the vector, and it is calculated as follows:

$$\|\overrightarrow{\text{ED}}\|=\sqrt{(\text{Real Risk-Estimated Risk})^2+(\text{Estimated NPV-Real NPV})^2} \text{...............} (4.3)$$

The magnitude of the expected disappointment in the example shown in **Fig. 17** :

$$\|\overrightarrow{\text{RE}}\|=\sqrt{(\$1{,}034\text{MM-}\$465\text{MM})^2+(\$1{,}946\text{MM-}\$1{,}194\text{MM})^2}= \$943\text{MM} \text{..........} (4.4)$$

4.1.3 Expected Decision Error

In previous work, the expected decision error (EDE) has been defined as the NPV of the best-possible portfolio minus the NPV of the realized portfolio; it is that portion of disappointment due to selecting the wrong projects. Similarly to ED, EDE is calculated using not only the NPV component but also the risk component. Expected decision error is also calculated using the expected values of risk and NPV. Using the same example in **Fig. 17**, EDE for a \$600MM risk-tolerance limit is the vector $\overrightarrow{\mathbf{RB}}$. The components of $\overrightarrow{\mathbf{RB}}$ can be determined as follows:

$$\overrightarrow{RB} = \begin{pmatrix} risk_B\text{-}risk_R \\ NPV_B\text{-}NPV_R \end{pmatrix} \quad (4.5)$$

The EDE can be expressed in terms of the unit $\overrightarrow{risk}$ and $\overrightarrow{NPV}$ as follows:

$$\overrightarrow{EDE} = (\text{Real Risk-Best Risk})\overrightarrow{risk} + (\text{Best NPV-Real NPV})\overrightarrow{NPV} \quad (4.6)$$

The decision error is reported as the magnitude of the vector, and it is calculated as follows:

$$\|\overrightarrow{EDE}\| = \sqrt{(\text{Real Risk-Best Risk})^2 + (\text{Best NPV-Real NPV})^2} \quad (4.7)$$

The magnitude of the expected decision error in the example shown in **Fig**. **17** is:

$$\|\overrightarrow{RB}\| = \sqrt{(\$1{,}034MM\text{-}\$488MM)^2 + (\$397MM\text{-}\$1{,}194MM)^2} = \$966MM \quad (4.8)$$

4.1.4 Estimation Error

In previous work, estimation error has been defined as the difference between the expected disappointment and the expected decision error; it is that portion of disappointment due to estimation errors. It can also be defined as the NPV of the estimated portfolio minus the NPV of the best-possible portfolio. Estimation error is also calculated using the expected values of NPV and risk. In **Fig. 17**, estimation error for a risk tolerance of $600MM is the vector $\overrightarrow{BE}$. Thus, estimation error can be calculated using the equations below:

$$\overrightarrow{BE} = \overrightarrow{RE} - \overrightarrow{RB} \quad (4.9)$$

$$\overrightarrow{BE} = \begin{pmatrix} risk_E - risk_B \\ NPV_E - NPV_B \end{pmatrix} \quad (4.10)$$

The magnitude of the expected estimation error in the example shown in **Fig. 17** :

$$\|\overrightarrow{BE}\| = \sqrt{(\$488MM - \$465MM)^2 + (\$1,946MM - \$397MM)^2} = \$1,549MM \quad (4.11)$$

4.2 Impact of Biases on the Estimated Portfolio

In this section, a sensitivity study is performed to assess the impact of varying overconfidence and directional bias on the expected efficient frontier curve of the estimated portfolio. The reason the estimated portfolio was specifically chosen for this sensitivity study is that the best-possible portfolio is not impacted by biases and the realized portfolio is dependent on the estimated portfolio.

4.2.1 Variable Overconfidence and Fixed Directional Bias

Fig. 18 shows the efficient frontier curves for the estimated expected portfolio at a fixed value of directional bias (0.5) and different values of overconfidence (high 0.75, medium 0.5 and low 0.25). The graph also includes the efficient frontier curve for the best-possible expected portfolio to visualize how far the estimates deviate from the truth case. By fixing the directional-bias level and looking at different levels of overconfidence, one can see how overconfidence impacts the estimates, or more specifically, the optimization of the estimated portfolio.

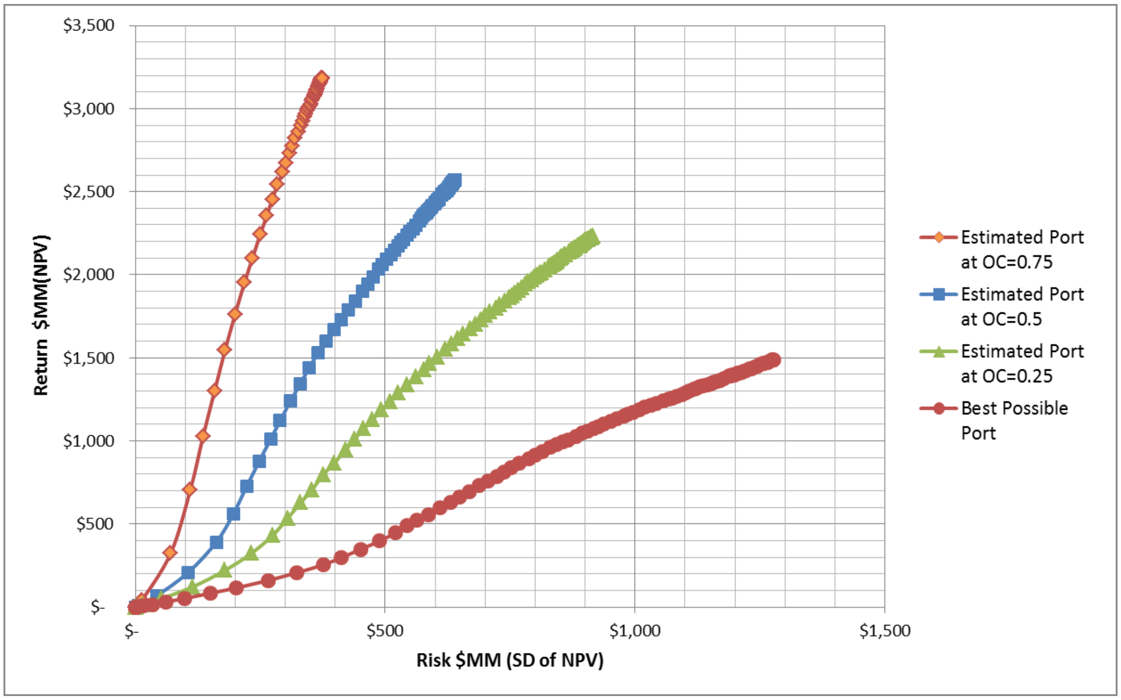

Fig. 18—Efficient frontier curves for the estimated expected portfolios at 0.5 directional bias and variable overconfidence for a maximum risk tolerance of $2,500MM

Fig. 18 shows that the higher the overconfidence, the higher the estimated NPV and the lower the estimated risk. This is because the impact of high overconfidence combined with optimism on the estimated distribution that is a subset of the true distributions. In this case, the estimated distributions are located at the right side of the PVOCF distribution and the left side of the CapEx distribution, which results in greater overestimation of the NPV and greater underestimation of the CapEx. Underestimation of CapEx also means more projects are selected, and overestimation of the NPV means a higher NPV is expected. Concerning the estimated risk, increasing overconfidence makes the estimated distribution narrower and the standard deviation lower. Hence, the risk (the standard deviation of the estimated NPV) is more underestimated for high values of overconfidence.

To conclude, **Fig. 18** shows that increasing the level of overconfidence, with moderate optimism, increases underestimation of CapEx and risk, and increases overestimation of NPV. More projects are included in the estimated portfolio and a higher NPV and a lower risk are expected.

4.2.2 Variable Directional Bias and Fixed Overconfidence

Fig. 19 shows the efficient frontier curves for the estimated expected portfolio at a fixed value of overconfidence (0.5) and different values of directional bias (high 0.75, medium 0.5 and low 0.1) to evaluate the impact of directional bias on the estimates. The graph also includes the efficient frontier curve for the best-possible expected portfolio to visualize how far the estimates deviate from the truth case.

Fig. 19 shows that as directional bias increases, at moderate overconfidence, the efficient frontier curve of the estimated expected portfolio goes further, and is slightly lower. This is an artifact of the way directional bias and risk are defined. A directional bias of zero means that the truncation of the true distribution occurs equally from both ends. **Fig. 20** illustrates that for the case of PVOCF, as directional bias increases (increased optimism), an area fraction is truncated from the left side and added to the right side. This makes the estimated distribution more spread out because most of the area (probability) in a lognormal distributions is concentrated to the left. Thus, removing a small area fraction from the left side of the distribution and adding it to the right side will make the estimated distribution more spread out. For the case of CapEx, the increase in directional bias (increased optimism) results in removing an area fraction from the right and adding it to the left side. The difference in the spread for the case of CapEx is not significant because of the shape of the lognormal distribution, which is skewed to the right. Hence, because of the shape of the lognormal distribution, increasing directional bias (increased optimism), increases the spread out of the estimated distribution of PVOCF, which increases the estimated risk.

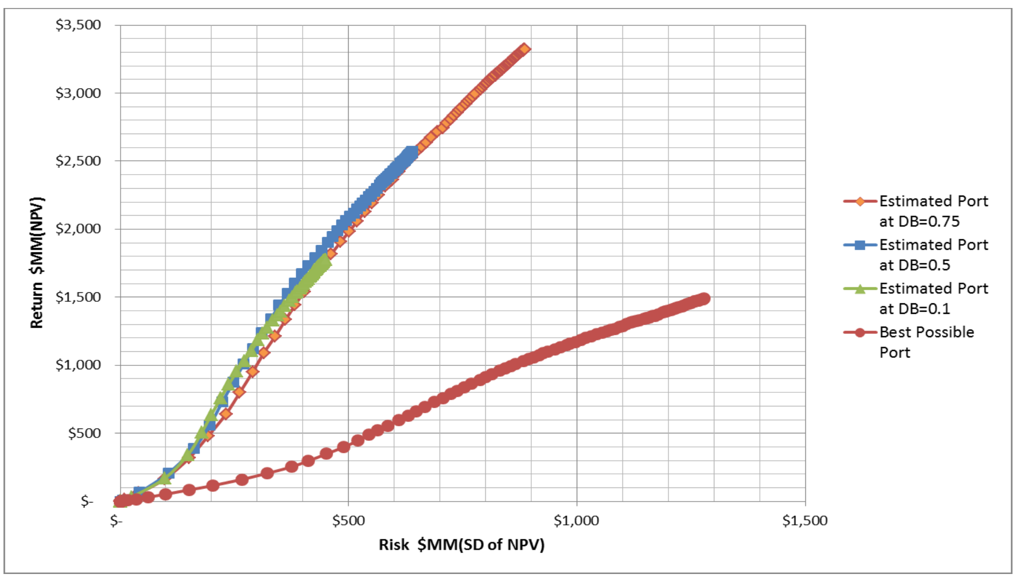

Fig. 19—Efficient frontier curves for the estimated expected portfolios for a maximum risk tolerance of $2,500MM at 0.5 overconfidence and variable directional bias

The higher the directional bias, the more spread out is the estimated distribution of PVOCF. A large spread of the estimated distribution makes the standard deviation higher. The estimated risk is defined as the standard deviation of the estimated NPV, and is calculated using the standard deviation of the estimated PVOCF and estimated CapEx.

Thus, the higher the directional bias, the larger is the estimated risk, especially at high values of directional bias. On one hand, the increase in directional bias (increased optimism) makes overestimation of PVOCF and underestimation of CapEx greater, which should increase the performance of the estimated portfolio. On the other hand, the increase in directional bias (increased optimism) makes estimated risk greater, which reduces the number of projects selected and decreases the portfolio NPV. These two factors offset each other, and the results in **Fig. 19** show that for high directional bias, the increase in the estimated risk seems to have the biggest impact and lowers the performance of the estimated portfolio. In addition, the higher the directional bias level, the further the curves go on both the NPV axis and the risk axis. This can be explained by the fact that the increase in directional bias (increased optimism) increases both the estimated NPV and the estimated risk. This artifact is the result of the way directional bias and risk have been defined. **Fig. 20** shows an example of how the spread in the estimated distributions (the shaded area) increases with the increase in directional bias (increased optimism) at a fixed overconfidence level for CapEx and PVOCF.

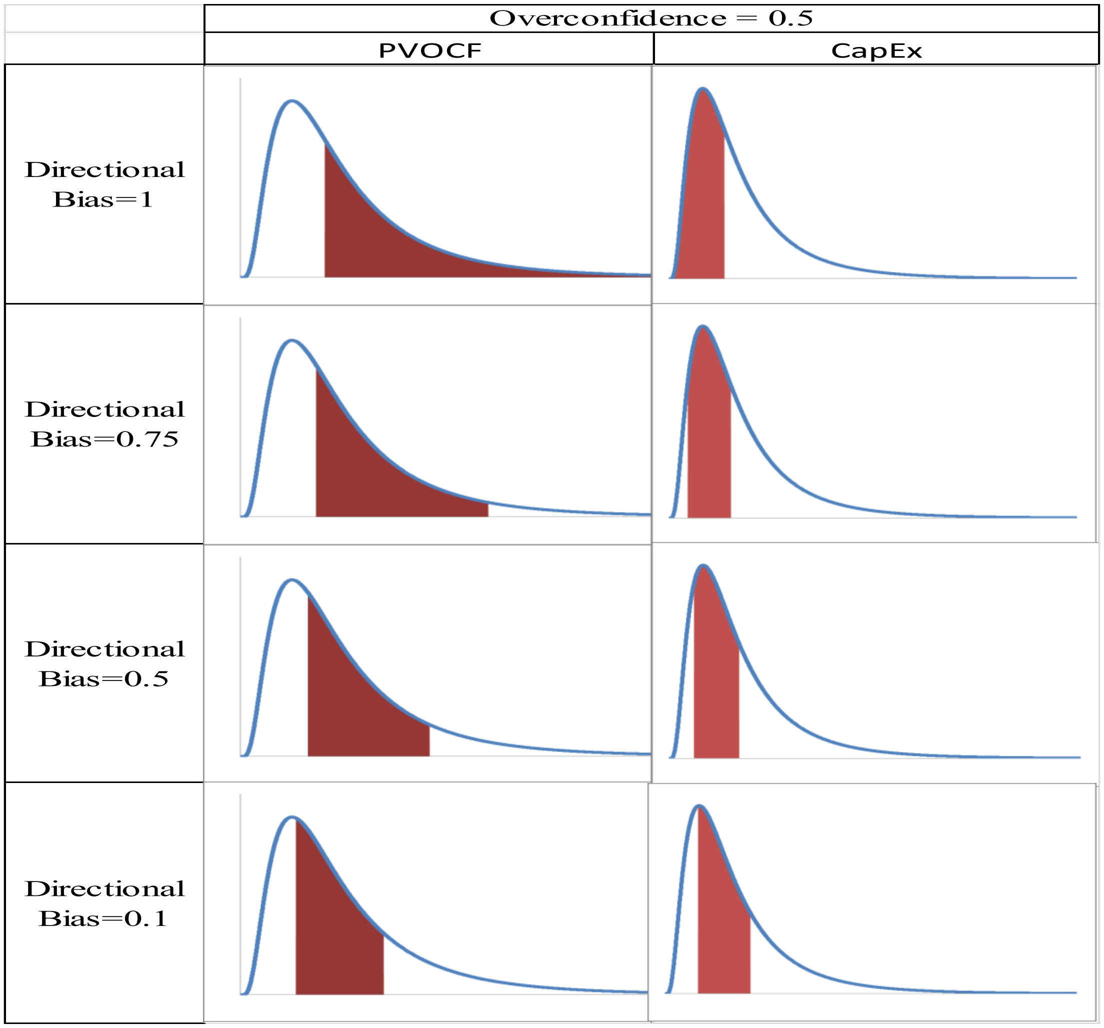

Fig. 20—Relationship between the estimated distribution (shaded) and the true distribution (unshaded) at a fixed overconfidence level and different directional bias levels

4.3 Impact of Overconfidence on Expected Disappointment and Expected Decision Error

4.3.1 Expected Disappointment

Fig. 21 is a plot of the expected disappointment at different values of risk tolerance for a fixed value of directional bias (0.5) and different values of overconfidence (high 0.75, medium 0.5 and low 0.25), which corresponds to the case shown in **Fig. 18**. The expected disappointment curves shown in **Fig. 21** stop at the risk-tolerance level at which the corresponding estimated portfolio (with similar overconfidence and directional bias levels) stops. Hence, these expected disappointment curves in **Fig. 21** go further to the right for low values of overconfidence, as shown and explained in **Fig. 18.**

The highest values of expected disappointment occur at 0.75 overconfidence, the second highest at 0.5 and the lowest at 0.25. Hence it can be concluded that the higher the overconfidence, the higher the expected disappointment. As explained in the previous section, overconfidence combined with optimism results in underestimation of CapEx and risk, and overestimation of NPV. At a fixed value of directional bias, the more overconfidence is increased, the more the estimated values deviate from the true. Hence, increasing overconfidence increases the difference between the realized portfolio performance (based on true values) and the estimated portfolio performance (based on estimated values), which increases the expected disappointment. The curves show that the higher the risk-tolerance level, the higher the expected disappointment. This is because at high-risk-tolerance values, the difference between the realized curve and the estimated curve grows larger as the realized curve shifts further to the right because of underestimation of risk (**Fig. 17**). The first values of expected disappointment in **Fig. 21** are equal to zero because the three portfolios (estimated, realized, best-possible) cannot fit in any project within that risk-tolerance limit knowing that, as explained in the previous section, each portfolio should (1) not exceed the risk-tolerance limit (2) fully use the available budget.

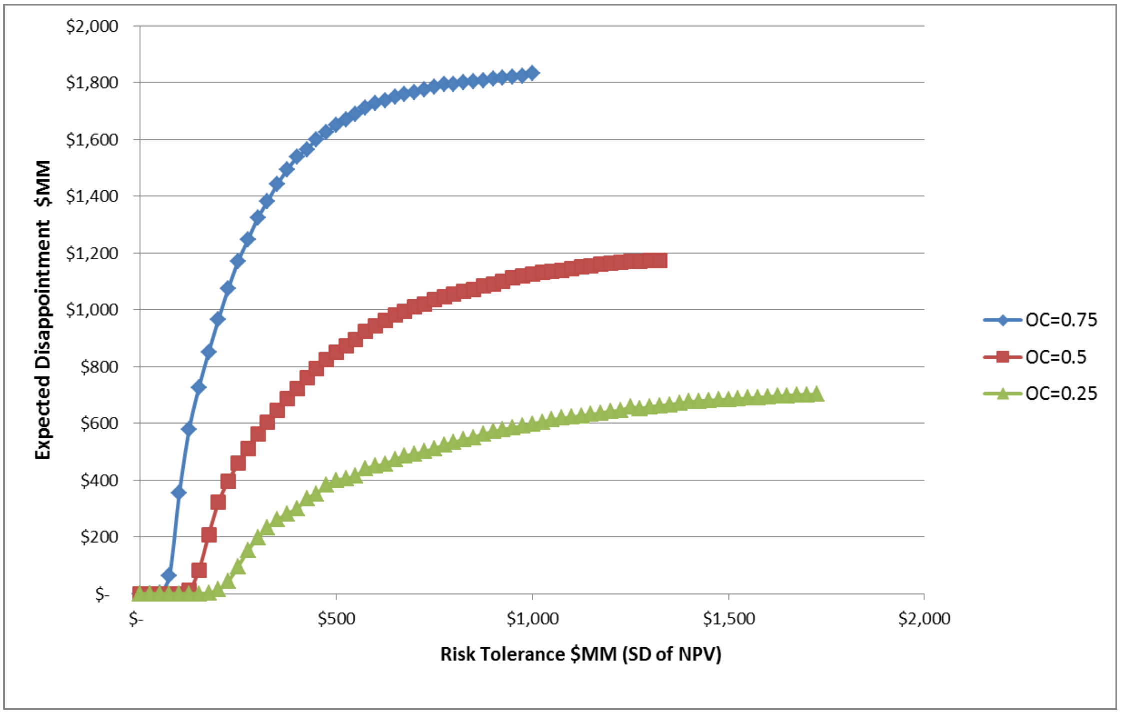

Fig. 21—Expected disappointment at 0.5 directional bias

4.3.2 Expected Decision Error

Fig. 22 is a plot of the expected decision error for different values of risk tolerance at a fixed level of directional bias (0.5) and different values of overconfidence, which corresponds to the case shown in **Fig. 18**. The curves stop at the risk-tolerance level at which the corresponding estimated portfolio (with similar overconfidence and directional bias levels) stops. Hence, the expected decision error curves in **Fig. 22** go further to the right for low values of overconfidence, as shown and explained in **Fig. 18.**

These expected decision error curves in **Fig. 22** show that higher levels of overconfidence have higher associated expected decision error. The decision error is caused by choosing the wrong portfolios. At a fixed value of directional bias, the higher the overconfidence, the more the estimated expected values deviate from the true expected values. Moreover, increasing overconfidence makes the estimated distribution even narrower, which decreases the standard deviation and makes the estimated risk much smaller than the true risk. Hence, increasing overconfidence makes the estimated values deviate more from the true values, which increases the expected decision error.

The expected decision error increases at lower levels of risk tolerance, and becomes constant or decreases at high levels of risk tolerance. This is because having high-risk tolerance makes the risk limit (one of the two limiting factors of the optimization) less constraining. Having one of the limiting factors less constraining, makes the pool of portfolios (out of which the optimum portfolio is selected) larger, which reduces the probability of choosing a suboptimal portfolio and consequently reduces the decision error. Also, the best-possible portfolio is more constrained by the risk-tolerance limit than the realized portfolio, and the lower the risk-tolerance limit the more constrained is the best-possible portfolio (resulting in a high decision error at low-risk-tolerance). As explained before, the projects of the realized portfolio are selected based on the estimated values (in which the risk is underestimated), and then the true CapEx of the realized portfolio is constrained by the budget limit while the true risk of the realized portfolio (which is higher) is not constrained. Hence, decision error is large at low-risk-tolerance limits, because the risk-tolerance limit is too small for the best-possible portfolio (selects projects based on true risk, which is higher) to pick up projects, while the realized portfolio (selects projects based on estimated risk, which is lower) has projects included in it. As the risk tolerance increases, the best-possible portfolio starts picking up projects, which reduces the decision error.

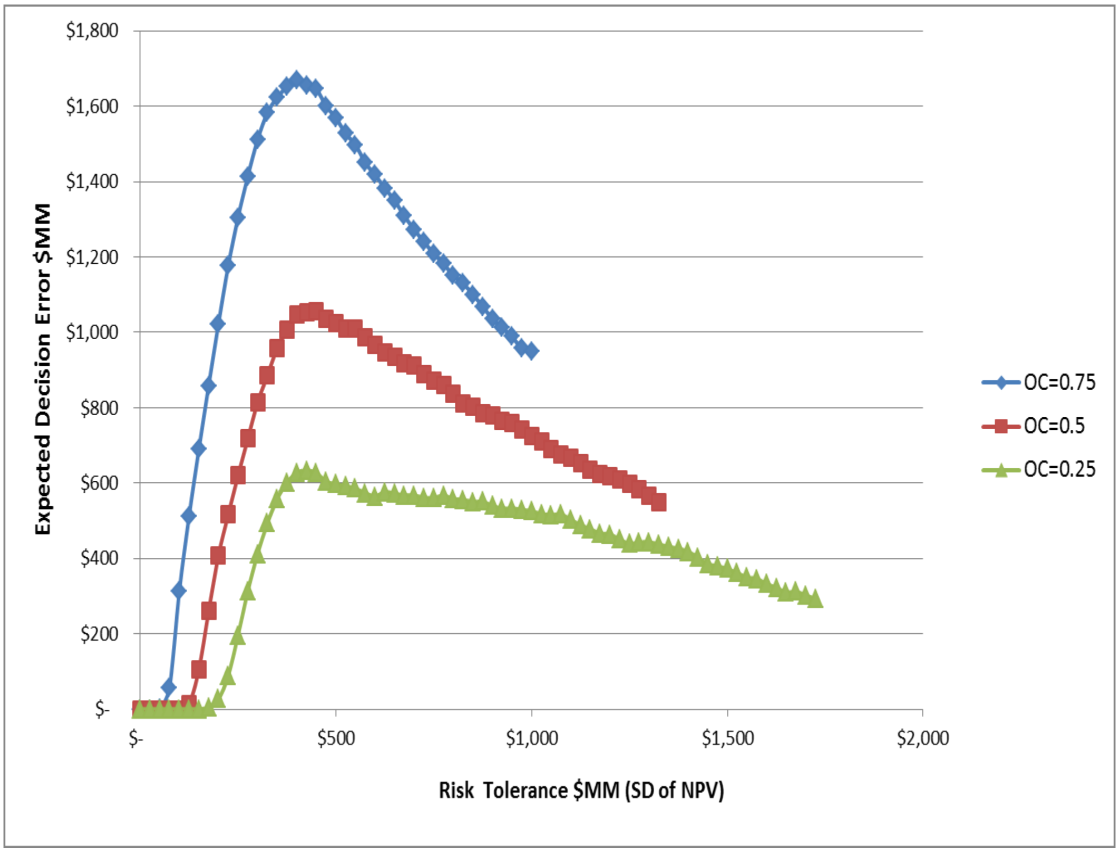

Fig. 22—Expected decision error at 0.5 directional bias

The first values of expected decision error in **Fig. 22** are equal to zero because the risk-tolerance limit is so small that the estimated portfolio (and consequently the realized portfolio) as well as the best-possible portfolio cannot fit in any project within that risk-tolerance limit, knowing that each portfolio should fully use the available budget. A portfolio can have one partial project as long as this partial project is fully using the available budget, but sometimes the risk-tolerance limit is so small that none of the partial projects within this limit can fully use the available budget.

4.4 The Case of Pessimism

In the case of pessimism, or negative directional bias, the estimations will be conservative to reflect the pessimistic attitude of estimators. Thus, with negative directional bias, the estimated distribution of PVOCF is shifted in the negative direction, and the estimated distribution of CapEx is shifted in the positive direction. Consequently, the portfolio NPV is underestimated and the portfolio CapEx is overestimated. Pessimism combined with risk should result in overestimated risk. Overestimated risk means an estimated distribution with a greater spread than the true distribution, which is referring to the concept of underconfidence. This work models biases using the work of McVay and Dossary (2012) that choose to not model underconfidence because it is uncommon in industry. Pessimism in this work refers to the direction of the overconfidence, which results in an estimated risk always lower than the true risk. In the presence of overconfidence, negative directional bias combined with the lognormal distributions of value-based parameter (e.g., PVOCF) results in an underestimated risk (**Fig. 23**). **Fig. 24** is an expected efficient frontier curve for the case of pessimism (-0.5 directional bias) and overconfidence (0.5), with the same $400MM budget limit and the same maximum risk-tolerance limit of $2,500MM (a higher maximum risk-tolerance limit will let the best-possible curve go further out until it exceeds the realized portfolio curve, but it will tremendously increase the running time). Unexpectedly, running the model for this case provides an estimated expected portfolio curve that is shifted to the left (**Fig. 24**).

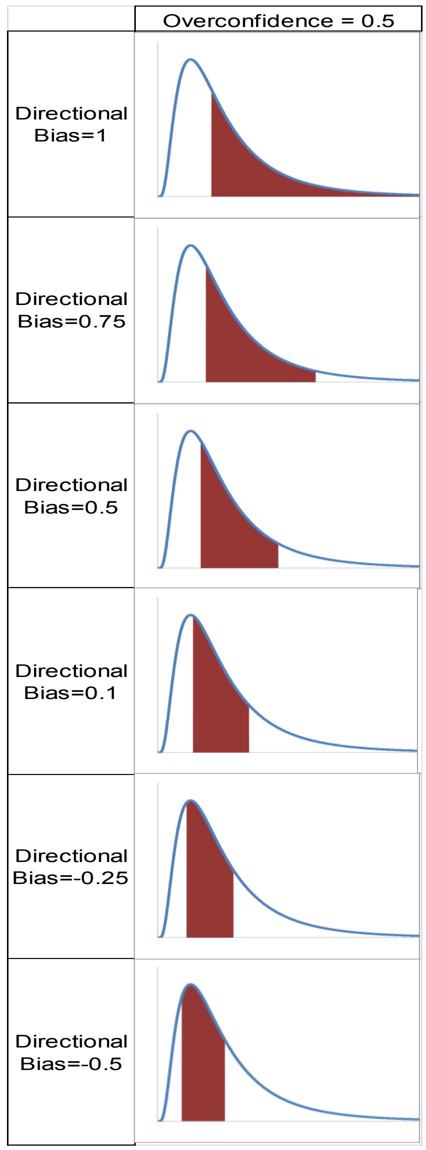

Fig. 23—Relationship between the estimated distribution (shaded) and the true distribution (unshaded) for a value-based parameter (e.g., PVOCF) at a fixed level of overconfidence and different levels of directional bias

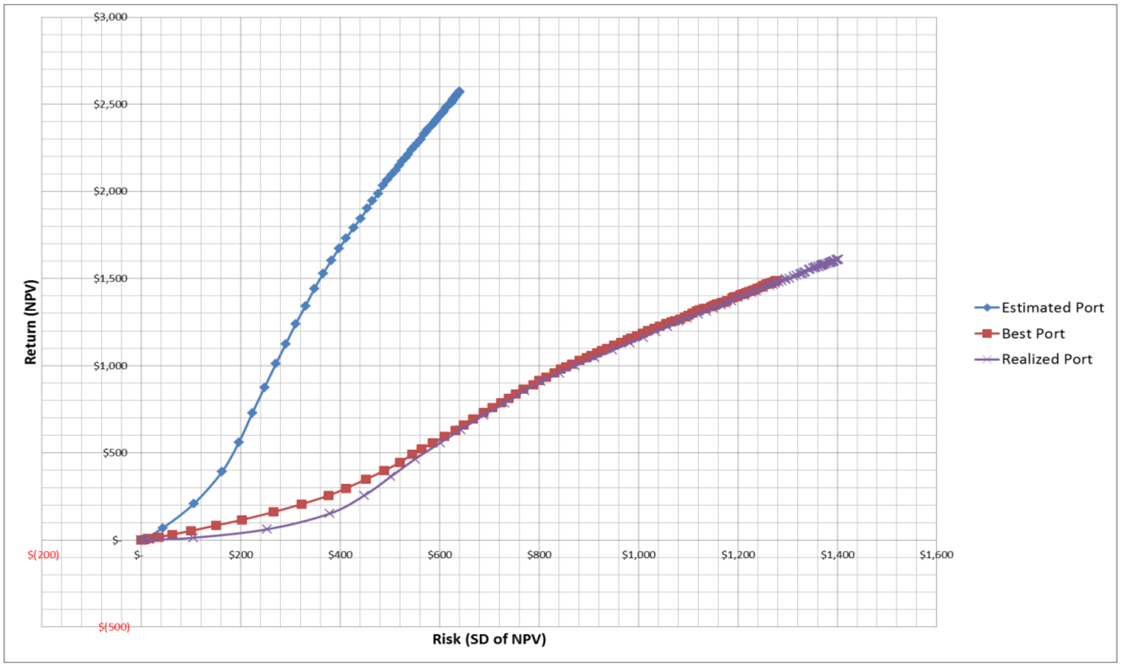

Fig. 24—Expected efficient frontier curves for the case of -0.5 directional bias and 0.5 overconfidence

Fig. 25 shows that for the same overconfidence level (0.5), the estimated expected portfolio curve in the case of pessimism (-0.5 DB) is underperforming (lower NPV, higher CapEx) compared to the estimated expected curve of the corresponding optimistic case (0.5 DB), as expected. This is because pessimism results in underestimation of NPV and overestimation of CapEx. In the presence of overconfidence and negative directional bias, the estimated expected portfolio curve is expected to be located to the right and underneath the realized expected portfolio curve, but this not the case due to the way risk and biases are defined.

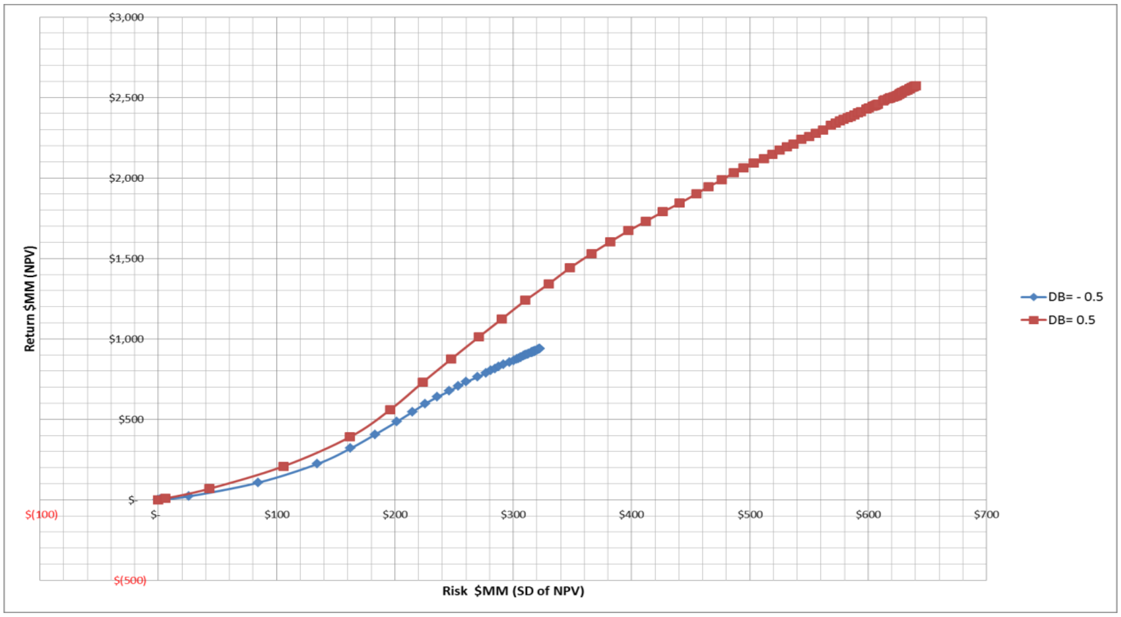

Fig. 25—Comparison of the efficient frontier curves for the estimated expected portfolios for the case of optimism and pessimism at 0.5 overconfidence

4.5 Summary of Results

4.5.1 Expected Disappointment as a Percentage of the Estimated Portfolio

In previous work, ED%E is defined as the expected disappointment as a percentage of the estimated NPV. In this work, ED%E is defined as the expected disappointment as a percentage of not only the estimated NPV but also the estimated risk. This definition is used because disappointment is expressed in terms of NPV and risk, so it is more consistent to consider both NPV and risk of the estimated portfolio; especially that this work uses Markowitz theory that considers that each portfolio has a reward (NPV) and risk level. ED%E is calculated at each Monte Carlo iteration by dividing the magnitude of the disappointment by the magnitude of the estimated portfolio using the following equation:

$$ED\%E = \frac{\sqrt{(Real\ Risk - Estimated\ Risk)^2 + (Estimated\ NPV - Real\ NPV)^2}}{\sqrt{(Estimated\ Risk)^2 + (Estimated\ NPV)^2}} * 100 \quad \text{.................... (4.12)}$$

Fig. 26 is a plot of the ED%E at different values of overconfidence and directional bias for high-risk and low-risk tolerance. The dotted curves refer to high-risk tolerance and the solid line curves refer to low-risk-tolerance. The high-risk-tolerance level is determined as the risk-tolerance level after which further increase in the risk tolerance will not increase the expected NPV and risk of the estimated expected portfolio by much. This risk level is different from a case to another, because the extent to which the estimated expected curve goes to the right is dependent on the level of overconfidence and directional bias. The low-risk-tolerance level is when the EDE%E peaks (as shown in **Fig. 22**), which occurs at around 15% of the high-risk-tolerance level. Three different levels of overconfidence (high at 0.75, medium at 0.5 and low at 0.25) and three different levels of directional bias (high at 0.75, medium at 0.5 and low at 0.1) are examined. As pointed out in the literature review section, the petroleum industry is typically optimistic in its estimates; thus, only positive directional bias (optimism) is considered in this plot.

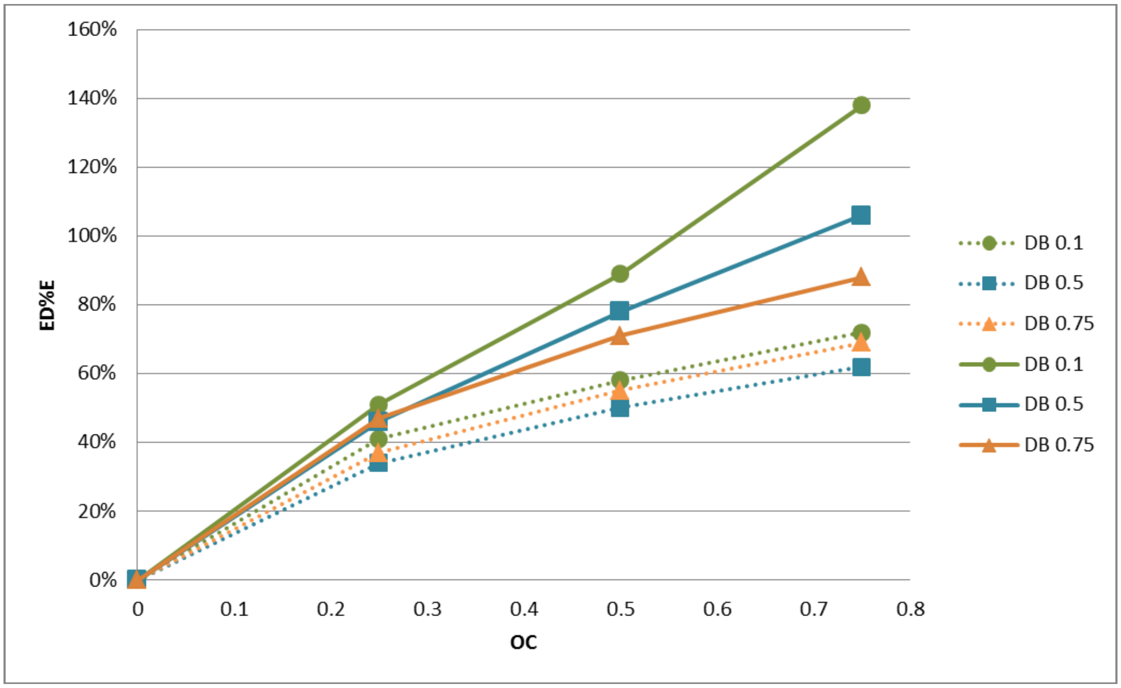

Fig. 26—ED%E for high-risk tolerance (dotted lines) and low-risk tolerance (solid lines)

Fig. 26 shows that for both high-risk tolerance and low-risk tolerance, the higher the overconfidence level, the greater the ED%E, which is similar to the results of McVay and Dossary (2012). As directional bias decreases, ED%E increases, especially at low-risk-tolerance. This is an artifact of the impact of directional bias on the spread of the estimated distribution and the estimated risk. This result is not observed in the cases investigated by McVay and Dossary (2012), because they do not use the risk factor. As explained in the previous sections, decreasing directional bias makes the estimated risk smaller, which further underestimates the risk. Consequently, the difference between the true and the estimated risk is even greater resulting in higher disappointment. For high-risk-tolerance environments, the optimization is less sensitive to the underestimation of risk and more sensitive to the overestimation of NPV and underestimation of CapEx that results from increasing directional bias (increased optimism); this is caused by what was explained in earlier sections as the two opposing impacts of increasing directional bias. Similarly to McVay and Dossary's (2012) results, at 0 overconfidence, ED%E is always equal to 0 even at nonzero directional bias.

Moderate overconfidence (0.5) and moderate optimism (0.5) result in an expected disappointment of 50% of the estimated NPV and estimated risk in a high-risk-tolerance environment. The same amount of moderate overconfidence and directional bias result in an expected disappointment that goes up to 78% of the estimated NPV and estimated risk in a low-risk-tolerance environment. ED%E is lower in high-risk-tolerance environments because the estimated NPV and risk (by which the disappointment is divided) are high. For the same amount of moderate overconfidence (0.5) and moderate optimism (0.5), McVay and Dossary (2012) model results in ED%E of 30-35%, which are lower than the ED%E this model generates. This difference is due to considering the risk factor in this model, while McVay and Dossary (2012) model can be considered as having an unlimited risk tolerance. Hence, it can be deduced that as risk tolerance increases, ED%E becomes less significant because the estimated NPV and risk increase at a rate higher than the rate of increase of disappointment and decision error.

4.5.2 Expected Decision Error as a Percentage of the Estimated Portfolio

Fig. 27 is a plot of the EDE%E at different values of overconfidence and directional bias for high-risk and low-risk tolerance (similar **Fig. 26** but for the case of EDE%E). In previous work, EDE%E is defined as the expected decision error as a percentage of the estimated NPV. This work defines EDE%E as the expected decision error as a percentage of the estimated NPV and risk. This is because decision error is expressed in terms of NPV and risk, so it is more consistent to consider both NPV and risk of the estimated portfolio. EDE%E is calculated at each Monte Carlo iteration by dividing the magnitude of the decision error by the magnitude of the estimated portfolio using the following equation:

$$EDE\%E = \frac{\sqrt{(Real\ Risk - Best\ Risk)^2 + (Best\ NPV - Real\ NPV)^2}}{\sqrt{(Estimated\ Risk)^2 + (Estimated\ NPV)^2}} * 100 \quad \text{(4.13)}$$

The high-risk tolerance (dotted line curves in **Fig. 27**) and the low-risk tolerance (the solid line curves in **Fig. 27**) were determined in the same way they were determined for the plots of **Fig. 26**. Similarly to McVay and Dossary's (2012) results, as the level of overconfidence increases, the EDE%E increases as well for both high and low-risk-tolerance. The increase in EDE%E due to the increase in overconfidence is less significant for higher values of directional bias. This is because at high values of directional bias, the estimated risk is higher and closer to the true risk, which reduces the decision error. In addition, the EDE%E is the expected decision error divided by the estimated risk and the estimated NPV, and as shown in **Fig. 18** at high values of directional bias, the estimated portfolio goes further to the right reaching high values of estimate risk and NPV. Hence, although the expected decision error increases, but dividing it by a large estimated NPV and risk makes the EDE%E smaller.

The curves show that the lower the directional bias, the higher the decision error. This is because of the artifact resulting from the way directional bias and risk are defined. This result is not observed in the cases investigated by McVay and Dossary (2012), because they do not use the risk factor. As explained previously, high values of directional bias result in a greater spread out of the estimated distribution, which overestimates the risk. The higher the estimated risk, the closer it is to the true risk, which makes the decision error smaller.

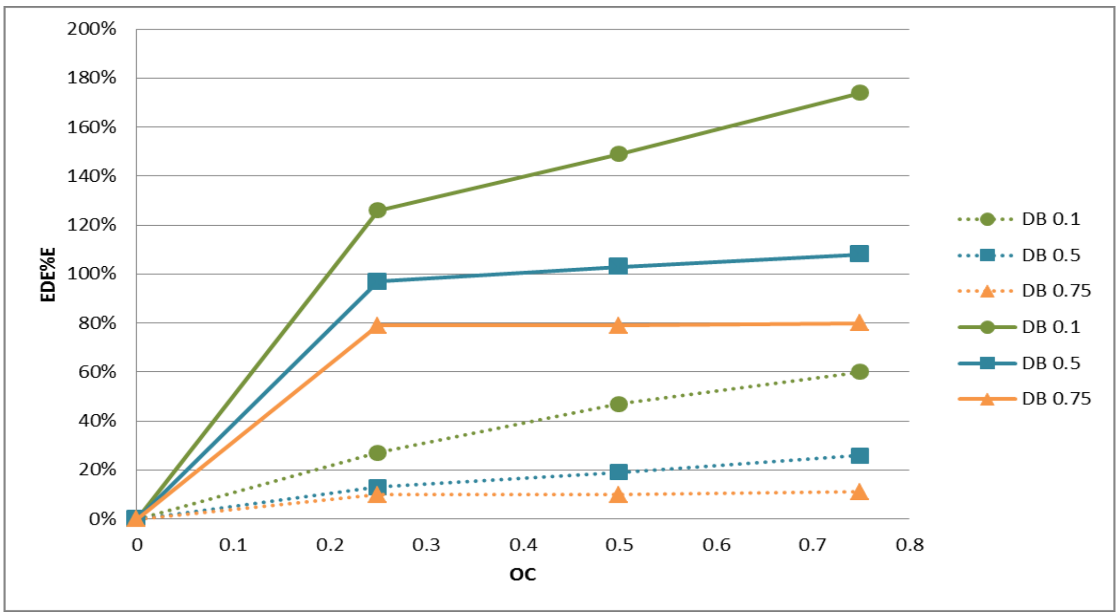

Fig. 27—EDE%E for high-risk tolerance (dotted lines) and low-risk tolerance (solid lines)

Similarly to McVay and Dossary's (2012) results, at 0 overconfidence, EDE%E is always equal to 0 even at nonzero directional bias. **Fig. 27** shows that for moderate overconfidence (0.5) and moderate optimism (0.5) result in an expected decision error of 19% of the estimated NPV and estimated risk for a high-risk-tolerance environment. The same amount of moderate overconfidence and directional bias results in an expected decision error that goes up to 103% of the estimated NPV and estimated risk for a low-risk-tolerance environment. As explained earlier, the EDE%E is high in low-risk-tolerance environments, because the realized portfolio is not subject to the risk-tolerance limit while the best-possible portfolio and the estimated portfolio are constrained by a low-risk-tolerance limit. Hence the decision error (the difference between the risk of the best-possible portfolio and the realized portfolio) is large and it is divided by a low estimated NPV and estimated risk. Hence, EDE%E reaches very high values, because the risk tolerance is very low and the true risk is very high that, in some cases, the best-possible portfolio does not include any projects, while the realized portfolio has projects that usually return a low NPV and high risk.

EDE%E is less significant in high-risk-tolerance environments because the estimated NPV and risk (by which the decision error is divided) are high. For the same amount of moderate overconfidence (0.5) and moderate optimism (0.5), McVay and Dossary (2012) model results in EDE%E of 1-5%, which are lower than the EDE%E this model generates. This difference is due to considering the risk factor in this model, while McVay and Dossary (2012) model can be considered as having an unlimited risk tolerance. Hence it can be deduced that as risk tolerance increases, EDE%E becomes less significant because the risk limit is less constraining, which makes the pool of portfolios (out of which the optimum portfolio is chosen) larger, this reduces the probability of choosing a suboptimal portfolio.

From the plots in this section, it can be concluded that even if uncertainty will be always present, reliably quantifying uncertainty can be achieved by focusing on reducing overconfidence, while other biases are reduced in the process. Reliably quantifying uncertainty can be achieved by providing ranges that are wide enough to include not only expected outcomes but also possible outcome. This will have the value of significantly reducing the expected disappointment and the expected decision error. Consequently, overall industry performance can be improved because accurate estimates enable identification of superior portfolios that have optimum reward and risk levels; it also increases the probability of meeting expectations.

5. CONCLUSIONS AND FUTURE WORK

5.1 Conclusions

This work shows that, in the presence of overconfidence and directional bias (optimism and pessimism), disappointment in portfolio performance occurs not only because the realized portfolio NPV is lower than estimated, but also because the realized portfolio risk is higher than estimated. This disappointment is due to both incorrect estimation of value and risk (estimation error) and incorrect project selection (decision error). More conclusions, based on portfolio modeling results using a particular set of global portfolio parameters, are listed below:

- Increasing overconfidence, even if directional bias remains fixed, increases expected disappointment and expected decision error.
- When risk tolerance is high relative to potential portfolio values, moderate overconfidence (0.5) and moderate optimism (0.5) result in an expected disappointment of about 50% and an expected decision error of about 19% and of the estimated portfolio value.
- When risk tolerance is low relative to potential portfolio values, the same amounts of moderate overconfidence and directional bias result in an expected disappointment up to 78% and an expected decision error up to 103% of the estimated portfolio value. The EDE%E reaches very high values at low risk tolerance.

- For the same amount of moderate overconfidence (0.5) and moderate optimism (0.5), McVay and Dossary's (2012) model results in ED%E and EDE%E that are lower than the ones generated by this model. ED%E and EDE%E are larger in this work because they include additional disappointment and decision error from the risk component, which was not considered in earlier work.
- For the case of pessimism, in the presence of overconfidence, the model generates estimated expected portfolios that have higher NPV and lower risk than the realized expected portfolios. While unexpected, this results because, in the model employed, pessimism relates only to the direction of bias and does not mean pessimism in risk, i.e., underconfidence.

Reliably quantifying uncertainty can be achieved by focusing on reducing overconfidence, while other biases are reduced in the process. Reliably quantifying uncertainty has the value of significantly reducing the expected disappointment and the expected decision error, which will improve the overall industry performance.

5.2 Future Work

- Increase the number of projects and consider the opting of adding more than one partial project per portfolio to overcome the discretization issue noticed in the results.
- In this work, disappointment and decision error are considered vectors in a two-dimensional space with a NPV component and a risk component. This approach assumes that $1 of NPV is equivalent to $1 of risk, which might not be always valid. Investigate other measures of disappointment and decision error.

- In this work, ED%E is defined as the expected disappointment as a percentage of the estimated NPV and risk; and EDE%E is defined as the expected decision error as a percentage of the estimated NPV and risk. The equations to calculate ED%E and EDE%E are explained in the results section. Further investigation of other methods to calculate the ED%E and EDE%E can be informative.
- The way directional bias and risk have been defined results in an artifact—increasing directional bias increases the estimated risk, which impacts the results as explained in the thesis. Further investigation could be helpful.
- Implement the model using a more powerful coding language that has dynamic structures and more capacity to be able to use a higher number of projects.
- Although PVOCF and CapEx implicitly cover aboveground (price fluctuations) and underground (reservoir properties) parameters, extending the project's scope to deal with these parameters directly will enable modeling them with a higher resolution.
- The results and conclusions of this work are based on a particular set of global portfolio parameters. It would be useful to investigate other global portfolios to check if these conclusions are general.

NOMENCLATURE

Best NPV	The net present value of the best-possible portfolio
Best risk	The risk of the best-possible portfolio
CapEx	Capital expenditure
Counter	Variable used to count the number of iterations
DB	Directional bias
ED	Expected disappointment
EDE	Expected decision error
ED%E	Expected disappointment as a percentage of the estimated portfolio value
EDE%E	Expected decision error as a percentage of the estimated portfolio value
ED%T	Expected disappointment as a percentage of the true portfolio value
EDE%T	Expected decision error as a percentage of the true portfolio value
EFC	Efficient frontier curve
Estimated NPV	The net present value of the estimated portfolio
Estimated risk	The risk of the estimated portfolio
E&P	Exploration and production
IE	Investment efficiency
NPV	Net present value

NPV_B	The best-possible expected net present value at point B
NPV_E	The estimated expected net present value at point E
NPV_R	The realized expected net present value at point R
OC	Overconfidence
O&G	Oil and gas
OOIP	Original oil in place
Port	Portfolio
PVOCF	Present value of the operating cash flow
Real NPV	The net present value of the realized portfolio
Real risk	The risk of the realized portfolio
$\|\overrightarrow{RE}\|$	The magnitude of the vector $\overrightarrow{RE}$
$risk_B$	The best-possible expected risk at point B
$risk_E$	The estimated expected risk at point E
$risk_R$	The realized expected risk at point R
SD	Standard deviation
VBA	Visual basic for applications
x_i	The participation level of project i
σ_i	The standard deviation of the NPV (risk) of project i
ρ	The correlation factor
$MM	Million dollars

REFERENCES

Al-Harthy, M.H. and Khurana, A.K. 2006. Portfolio Optimization: Systems Vs. Sequential Approach. Paper presented at the SPE Asia Pacific Oil & Gas Conference and Exhibition, Adelaide, Australia. Society of Petroleum Engineers 101152-MS. DOI: 10.2118/101152-MS.

@RISK, version 6.1.1. 2012. Palisade Corporation, http://www.palisade.com/risk/.

Begg, S.H. and Bratvold, R.B. 2008. Systematic Prediction Errors in Oil and Gas Project and Portfolio Selection. Paper presented at the SPE Annual Technical Conference and Exhibition, Denver, Colorado, USA. Society of Petroleum Engineers 116525-MS. DOI: 10.2118/116525-MS.

Brashear, J.P., Becker, A.B., and Faulder, D.D. 2001. Where Have All the Profits Gone? *SPE Journal of Petroleum Technology* **53** (6): 73141.

Brashear, J.P., Becker, A.B., and Gabriel, S.A. 1999. Interdependencies among E&P Projects and Portfolio Risk Management. Paper presented at the SPE Annual Technical Conference and Exhibition, Houston, Texas. Society of Petroleum Engineers 56574-MS. DOI: 10.2118/56574-MS.

Capen, E.C. 1976. The Difficulty of Assessing Uncertainty (Includes Associated Papers 6422 and 6423 and 6424 and 6425). *SPE Journal of Petroleum Technology* **28** (8): 843-850. 5579-PA.

Guerard, J.B. Jr. ed. 2009. *Handbook of Portfolio Construction: Contemporary Applications of Markowitz Techniques*, 31-60. Springer-Verlag New York.

Harrison, J.R. & March J.G. 1984. Decision Making and Post-decision Surprises. *Admin. Sci. Quarterly.* **29**, 26-42.

Markowitz, H. 1952. Portfolio Selection. *The Journal of Finance* **7** (1): 77-91.

McVay, D.A. and Dossary, M. 2012. The Value of Assessing Uncertainty. Paper presented at the SPE Annual Technical Conference and Exhibition, San Antonio, Texas, USA. Society of Petroleum Engineers 160189-MS. DOI: 10.2118/160189-MS.

Merritt, D. and Miguel, A.d.S. 2000. Portfolio Optimization Using Efficient Frontier Theory. Paper presented at the SPE Asia Pacific Conference on Integrated Modelling for Asset Management, Yokohama, Japan. Society of Petroleum Engineers 59457-MS. DOI: 10.2118/59457-MS.

Rose, P.R. 2004. Delivering on Our E&P Promises. *Leading Edge* **23** (2): 165.

Smith, J.E. and Winkler, R.L. 2006. The Optimizer's Curse: Skepticism and Postdecision Surprise in Decision Analysis. *Management Science* **52** (3): 311-322. 20335587.

Tversky, A. and Kahneman, D. 1974. Judgment under Uncertainty: Heuristics and Biases. *Science* **185** (4157): 1124-1131.

Welsh, M.B., Bratvold, R.B., and Begg, S.H. 2005. Cognitive Biases in the Petroleum Industry: Impact and Remediation. Paper presented at the SPE Annual Technical Conference and Exhibition, Dallas, Texas. Society of Petroleum Engineers 96423-MS. DOI: 10.2118/96423-MS.