# IT SOURCING PORTFOLIO MANAGEMENT FOR IT SERVICES PROVIDERS – AN APPROACH FOR USING MODERN PORTFOLIO THEORY TO ALLOCATE SOFTWARE DEVELOPMENT PROJECTS TO AVAILABLE SITES

**Steffen Zimmermann**
University of Innsbruck, Universitaetsstr. 15, A-6020 Innsbruck
phone: +43 512 507 - 9809, fax: +43 512 507 - 9809, email: steffen.zimmermann@uibk.ac.at

Steffen Zimmermann holds a graduate degree in Business Administration and a doctoral degree in Information Systems Engineering & Financial Management from the University of Augsburg, Germany. In addition, he was consultant and project manager in the Financial Services Industry. Since 2009 he is Assistant Professor of Information Systems at the School of Management at the University of Innsbruck (Austria). In 2011 he had an assignment as Visiting Professor at the University of Calgary, Canada. His research interests are Business Value of IT, Economics in IS, Global Software Management, and IT Service Management.

**Arne Katzmarzik**
Senacor Technologies AG, Erika-Mann-Str. 55, D-80636 München
phone: +49 89 588 089 - 0, fax: +49 588 089 - 700, email: arne.katzmarzik@senacor.com

Arne Katzmarzik holds a graduate degree in Civil Engineering from Ruhr-University Bochum and in Business Administration from the University of Hagen. He received his doctoral degree in Information Systems Engineering & Financial Management from the University of Augsburg. From 2003 until 2007 he was consultant at Accenture and responsible for a software development team. From 2007 until 2011 he was research assistant at FIM Research Center Finance & Information Management. Since 2011 he is senior consultant at Senacor Technologies AG. His research interests are Business Value of IT, IT Governance, IT Sourcing, and Cloud Computing.

**Dennis Kundisch**
University of Paderborn, Warburger Str. 100, D-33100 Paderborn
phone: +49 5251 60 - 5533, fax: +49 5251 60 - 5540, email: dennis.kundisch@wiwi.uni-paderborn.de

Dennis Kundisch holds graduate degrees in Business Administration from the University of Augsburg, Germany and the University of Dayton, USA. He received his doctoral degree and his habilitation from the University of Augsburg in 2002 and 2006, respectively, for publications in the area of E-Finance. From 2006 until 2008 he was the representing director of the Department of Information Systems at the University of Freiburg. In 2008 and 2009 he had assignments as Visiting Professor at the University of Calgary, Canada, the University of New South Wales, Australia, and the Technical University of Brandenburg, Germany. Since 2009 he holds the Chair of Business Information Systems, Information Management & E-Finance at the University of Paderborn, Germany. His research interests include Economics of IS, E-Finance, IT Business Value, IT Governance, and Social Networks.

## Abstract

Global sourcing of Information Technology (IT) work has become a widely accepted practice among transnational corporations. Most of the big IT Services Providers (ITSPs) maintain a portfolio of globally distributed delivery centers and have to decide on the assignment of specific software development projects to their available sites. ITSPs have to consider *expected costs*, *risks*, and *interdependencies* between projects and sites when making value-based sourcing decisions. However, value-based decision approaches that are both well founded in theory and applicable in practice have until now been missing in the Information Systems literature. As decision making with respect to the construction of portfolios of risky financial assets exhibits similar characteristics compared to value-based sourcing decision making, we base our approach on the Modern Portfolio Theory. This paper makes two contributions in this context: (1) It provides a conceptual foundation for the application of Modern Portfolio Theory within the scope of global sourcing of software development projects by ITSPs. Therefore, we adapt the Modern Portfolio Theory to ensure an optimal and full allocation of given software development projects to available sites. Our newly formed model considers site/project combinations as risky assets, assumes discrete portfolio shares, and factors in transaction costs as well as dependencies between both projects and sites. (2) It is the first to actually apply Modern Portfolio Theory using a real world business case. Thereby, we illustrate that using our model leads to considerably different project allocations to the available delivery centers of our case company as well as to substantially lower costs of the sourcing portfolio.

**Keywords:** Global Software Management, Modern Portfolio Theory, Site Selection, IT Services Providers, Software Development Projects, Decision Model

**Categories and subject descriptors:** K.6.1 [Management of Computing and Information Systems]: Project and People Management---Strategic information systems planning; I.6.5 [Simulation and Modeling]: Model development---Modeling methodologies

## Introduction

Global sourcing of IT work has become a widely accepted practice among transnational corporations. This development has been facilitated by (1) the increasingly modular design of software production, (2) the increasing availability and maturation of technologies for managing and coordinating work across geographic distances, and (3) the drastically improved capabilities in software development and project management in low-wage offshore regions, such as China or India (Carmel and Agarwal 2002). Today, big Information Technology Services Providers (ITSPs) like IBM, Accenture, or Wipro maintain a portfolio of globally distributed delivery centers and constantly have to decide on the allocation of software development projects to their available sites. As the principle of Shareholder Value Management has been established as the leading management philosophy in theory and practice for all activities of a firm – including sourcing decisions of an ITSP – these have to be aligned with a value-based management (Rappaport 1998). This implies that ITSPs have to consider *expected costs*, *risks*, and *interdependencies* among projects and sites when making sourcing decisions. Varying *expected costs* among sites result from differing productivities (Cusumano et al. 2003) as well as differing transaction costs (Dibbern et al. 2008). Differing *risks* among sites are due to geographical, legal and cultural differences (Winkler et al. 2006, Apte et al. 1997, Kliem 2004). Finally, *interdependencies* influence the project costs since ITSPs usually conduct several software development projects simultaneously. Interdependencies among projects at an ITSP are usually due to resource conflicts, when projects compete for the same scarce competencies in specific technical or functional domains (Santhanam and Kyparisis 1996). To complicate things, the magnitude of interdependencies changes based on whether two or more projects are conducted at the same or at different sites.

In the Information Systems (IS) literature only very few approaches have been described that address single aspects of value-based sourcing decision making for software development projects. To estimate the expected development costs, cost estimation models like COCOMO (Boehm 1981) have been enhanced to COCOMO II (Boehm et al. 2000) and COSYSMO (Valerdi 2008) to reflect – among other things – additional effort resulting from globally distributed development teams. To estimate project risks for different sites, Aubert et al. (1998) identify undesirable outcomes that may result from sourcing risks. They use probabilities of undesirable outcomes and their expected losses to calculate a monetary risk exposure. Approaches to allocate software development projects to the available sites of an ITSP taking into account expected costs, risks and interdependencies have thus far been lacking in the IS literature. But such value-based approaches that are both conceptually founded and practically implementable are required to make better sourcing decisions on a managerial level.

Decision making with respect to the construction of portfolios of financial assets exhibits similar characteristics compared to value-based sourcing decision making. The field of Finance has developed a variety of decision models to construct an optimal portfolio of risky financial assets (Elton et al. 2007), with the seminal one being the Modern Portfolio Theory (MPT) by Harry Markowitz (Markowitz 1952). The basic preconditions of MPT are that the two critical target figures of a portfolio are its expected return and risk – measured by the mean and its variance – and that efficient portfolios either minimize risk for a given expected return or maximize return for a given level of risk. A single financial asset is not only judged on its individual merits but in terms of its interdependencies to all other investments included in a portfolio. Thus, a simple sum of the weighted risks of the single assets to determine the portfolio risk is not correct because of disregarding interdependencies. Accordingly, the percentage risk of a portfolio might even be lower than the percentage risk of the least risky asset in the portfolio. This effect is called "risk diversification" and is due to the fact that the risks of different assets generally do not materialize all at the same time.

In the literature on global manufacturing planning MPT has been transferred to the problem of selecting an optimal portfolio of manufacturing sites (Hanink 1985, Hodder and Dincer 1986, Hodder and Jucker 1985). International price and exchange rate risks are considered as a penalty to be minimized; multinationality, via diversification effects, is valuable insofar that the variance of the portfolio of manufacturing sites decreases with geographical dispersion.

These site selection approaches of manufacturing planning allow us to expect that MPT can also be transferred to the site selection for software development projects. But in the application of MPT in the area of IT project portfolio planning, some challenges exist that are emphasized in the literature (Asundi and Kazman 2001, Verhoef 2002, Kersten and Verhoef 2003). The critique is directed particularly against the nature of the considered assets (i.e., IT projects) and their marketability. Furthermore, the manufacturing planning approaches are mostly concerned with the selection of the sites to form the optimal portfolio. This means, it is assumed that risk and return of the single sites have already been determined. But to apply these approaches in a real world setting the well-founded estimation of the input parameters is a necessary prerequisite for the site selection process.

This paper acknowledges the unique characteristics of the site selection challenge for IT projects and makes two contributions in this context: (1) it provides a conceptual foundation for the application of MPT within the scope of global sourcing of software development projects by ITSPs and (2) it is the first to actually apply MPT using a real world business case. The business case reflects the actual decision situation of an international ITSP that had to decide on the allocation of three software development projects to three available sites. This study combines current theoretical and empirical understanding of the IT sourcing decision with decision making theory. Thereby, it adds to the stream of normative decision making in the IT outsourcing literature (Dibbern et al. 2004). It is the first, to provide an actual methodology that helps internationally operating ITSP to make more informed decisions for allocating customer projects to global sites.

The remainder of this paper is organized as follows. In the next section we describe the global sourcing decision problem of our case company, discuss the validity and applicability of MPT to solve the site selection problem and deduce modeling issues. We then present our model to allocate software development projects to available sites (Site Selection Model) based on MPT. Afterwards, we illustrate the application of the model based on the actual data from our case company. We then discuss the limitations and the generalizability of the model before we finally conclude with the implications of our findings for management and research and the identification of perspectives for further research.

## Global Sourcing Decision Problem: Modeling Issues

Based on discussions with executives of our globally active case company ACME (actual name is undisclosed for reasons of confidentiality), we were inspired to develop a global sourcing decision model for software development projects. We were asked by ACME to enhance their current approach to allocate software development projects to their available delivery centers. To illustrate the decision problem, we describe how ACME made its decision. Subsequently, we discuss the validity of MPT for the described decision situation and deduce modeling issues that have to be addressed in our sourcing decision model.

### *ACME's Decision Problem*[1]

ACME is a worldwide operating ITSP and runs software development projects for its clients at its globally distributed software delivery centers. The German branch of ACME has a specialized division for clients in the financial sector. ACME recently acquired three long-term projects to develop software for three globally relevant banks. All projects are designed to run over a three-year horizon with constant division of work over the runtime:

- Project 1: A-Bank plans to introduce a brokerage and financial asset administration system to provide for better and faster services to its customers and a better administration of the customer data for the employees. The challenge is to integrate the functionalities for customers and the back office. The mission of ACME is to implement a new system using J2EE technology and to migrate the highly sensible customer data. Based on the tasks the system shall perform, the implementation effort is

---

[1] For reasons of confidentiality, all names are undisclosed and the figures are slightly modified.

estimated to 600,000 lines of code (about 200,000 lines of code per year). The project can be cut in five modules that have roughly the same size.[2]

- Project 2: B-Bank changed its IT strategy with the objective to replace the old application-based architecture with a modern, more flexible Service-Oriented Architecture. B-Bank expects better data integration, faster communication and a stronger linkage among its different divisions from the IT reorganization. The first step is the implementation of a company-wide system for business intelligence. ACME's challenge is to implement a single enterprise architecture across all divisions facing the challenge of a complex technological solution and the need for a great reuse of the software. The experts estimate the effort to amount to 240,000 lines of code (about 80,000 lines of code per year). The project can be cut in eight modules that have roughly the same size.
- Project 3: C-Bank pursues a strategy similar to B-Bank. Due to a different organization of the company and a more complex integration of the current databases accompanied by a more laborious migration, the effort is estimated to amount to 360,000 lines of code (about 120,000 lines of code per year). The project can be cut in six modules that have roughly the same size.

ACME negotiated contractually fixed prices for the three projects and had to decide on how to allocate the projects on its sites. Since the responsible executives already had experience in nearshore/offshore sourcing, they took into account their German site and two delivery centers they had already cooperated with in former projects for their decision:

- Site 1 – Germany: ACME usually implements the software development projects for its German clients at its site in Frankfurt/Main. Thus, no additional transaction costs arise in this case. Germany has a labor pool with matured IT specialists and a low fluctuation, but comparably high loaded costs[3] because

---

[2] To make the description of the decision situation and the application of our model easier to follow, we do not use the slightly different module sizes that ACME was confronted with. In so far we refer to stylized facts in the following.

[3] Loaded costs include labor costs, costs for benefits, space as well as overheads (Everest Research 2005).

of the German wage levels. A main advantage of developing onshore for a German client is the fact that all parties involved speak German. Any relocation would necessitate the (co-)usage of English as project language and as a primary language in the documentation.

- Site 2 – Czech Republic: ACME runs a delivery center in Prague. This nearshore site has become a very popular sourcing location for German clients as the skills of the employees are nearly as mature as in the German branch, both cultures are similar, and many of the employees speak or have at least basic knowledge of German, thus fostering communication. The fully loaded costs are about half as high as in Germany but with a higher expected annual increase. Sourcing work to this nearshore site causes moderate extra nearshore costs (transaction costs).
- Site 3 – India: ACME also operates a large delivery center in Bangalore. The skills of the Indian IT specialists are nowadays comparable to that of German or American specialists. A disadvantage for India lies in the cultural differences to Europe and the specialists' limited proficiency in German. The loaded costs are significantly lower (about a quarter of the loaded costs at the German site) than in the other countries that were considered.

After the identification of the potential locations, the executives of the German division had to decide on the allocation of the projects to the available sites. Therefore, they estimated the onshore effort in person months per year using the established estimation model COCOMO II (Boehm et al. 2000) for the three projects. Subsequently, they multiplied the estimated effort with the average loaded costs for an employee per month of each site to get the periodical development costs for each project at each possible site. To consider extra offshore risks, they defined different risk surcharges on the estimated development costs for each site. For the onshore development (Site 1 – Germany) they calculated with no risk surcharge, for the nearshore development (Site 2 – Czech Republic) with 5%, and for the offshore development (Site 3 – India) with 10%. Another site-specific surcharge was made for upcoming extra offshore costs at the different sites based on information gained from setting up previous projects at the delivery centers. Using the discounted cash flow (DCF) approach, ACME discounted the resulting cost values per period using a project-specific discount rate, representing the site-independent project risks. The resulting net present values were called “risk adjusted costs”.

Based on this approach, ACME concluded that for project 1, development in the Czech Republic and for project 2 and project 3, development in India leads to the lowest risk-adjusted costs. Taking into account both that the projects could be modularized to a certain extent (see brief project descriptions) and that ACME was especially worried about unexpected resource conflicts, which affected the project success at the different sites, the executives of ACME decided to develop substantial shares of each project onshore. In order to diversify the risk resulting from the unexpected resource conflicts of each project, they cut each project in two parts. One part was conducted onshore, the other one nearshore or offshore. ACME’s project and site allocation is shown in Table 1.

**Table 1. Project and Site Allocation of ACME**

| | Project 1 | Project 2 | Project 3 |
|---|---|---|---|
| Site 1 – Germany | 80% | 75% | 67% |
| Site 2 – Czech Republic | 20% | 0% | 0% |
| Site 3 – India | 0% | 25% | 33% |

The described approach - partly based on “gut feeling” rather than a rationale and objective decision model - led in a couple of times to unsatisfying results. As ACME is faced with intensified competition due to the rise of several competitors from low-wage countries, the executives of the German division of ACME asked themselves what they could have done better in this decision process. In particular, site and project interdependencies puzzled the decision makers of ACME. One of the executives stated: “Last year, a software developer specialized in customer data management at our German site was absent for six weeks due to a severe illness. This negatively affected all projects running at our German site at that time that had to do with customer data management. Now, two projects that we have to allocate require specific knowledge in implementing Web Services for Private Wealth Management. How can we incorporate such interdependencies in our decision model?” Another one added: “And we are a bit worried about the political stability in the Far East. Given a negative development there, would this also affect our projects running at other sites?"

The first type of interdependency – further on called *project dependency* – occurs among resource-dependent software development projects, which compete for similar resources (e.g. skills of a software developer). In case a scarce resource unexpectedly becomes unavailable, all projects that share this resource will suffer likewise. The second type of interdependency – further on called *site dependency* – arises from project-external risk, such as political or economical risk. These risks were just considered in terms of specific risk surcharges on the estimated development costs and standalone per site. However, if two sites are located in countries or regions that are strongly integrated, the occurrence of these risks may be heavily interdependent.

Both types of dependencies have thus far not been adequately considered in ACME's DCF approach. This real world business situation sets the stage for the development of an integrated sourcing decision model incorporating these aspects. In the following, we suggest MPT as the theoretical basis for such a sourcing model and deduce important modeling issues.

### ***Modeling Issues***

MPT was specifically designed to build a portfolio of financial assets while considering stochastic dependencies among the financial assets' returns. This notion can be transferred to model stochastic dependencies between the costs of projects that are conducted on the same or at different sites. In this subsection we analyze how MPT can be applied for the described sourcing decision problem and deduce modeling issues resulting from this analysis. To do this, we list and describe assumptions of MPT and discuss them in the context of a sourcing decision situation described above.

MPT assumes a market where market imperfections are ignored (Markowitz 1952). The following conditions are considered necessary for such a perfect market (Copeland et al. 2005):

1. *Homogeneous assets*

In perfect markets the traded assets are homogenous, i.e., there are no quality differences. For financial markets this for instance means that there are no differences between two shares of IBM common stock. In our sourcing decision problem we have selected a portfolio of sites to develop a given number of lines of code which belong to different software development projects. Applying the standard MPT to our problem would imply that each site is characterized by the same expected development costs and risks per line of code and the same pair-wise interdependencies with other sites. Realistically, this cannot be assumed because usually different experiences, know-how and specializations in certain kinds of software development projects exist at each site. In addition, projects also differ substantially in terms of required technology, knowledge and other factors, thus resulting in different effort estimations.[4] This leads to different costs, risks and interdependencies for different projects at one site and at the same time to different costs, risks and interdependencies for the same project at different sites. Thus, we have to model site/project combinations (SPCs) – instead of just sites – as the available "assets in the market" (*Modeling Issue 1*) and we have to determine the expected development costs, the risks, and the interdependencies to other SPCs for each SPC. From this point of view, a SPC in the sourcing model corresponds to a financial asset in MPT.

2. *No transaction costs or taxes*

In a perfect market there are no transaction costs for searching and contracting a transaction partner as well as for coordinating and managing a transaction. Moreover, there is no tax regime in place.

Globally distributed software development causes transaction costs, also labeled as extra offshore costs (Carmel and Tjia 2005). They may increase the production costs for an offshore development

---

[4] This is also reflected by a number of scale factors and effort multipliers in COCOMO that have an impact on the overall effort needed to run a project (Boehm et al. 2000).

by more than half (Davidson 2003). Thus, neglecting transaction costs would systematically distort the results of an optimization. For a more detailed examination, transaction costs may be differentiated into variable and fixed transaction costs. Variable transaction costs generally increase with the size of the project and include costs for traveling, management, communication and controlling (Dibbern et al. 2008). Fixed transaction costs occur during the planning and set up of a software development project. They are (largely) independent of the project size. Typical kinds are legal, negotiation or initiation costs, such as costs for infrastructure set up and initial training costs for offshore staff. Since transaction costs also arise on financial markets, MPT was accordingly further developed to include variable transaction costs and fixed transaction costs.[5] Pogue (1970) includes variable transaction costs by a specific surcharge on the expected returns of the single assets included in the portfolio. Patel and Subrahmanyam (1982) model fixed transaction costs in such a way that they only occur when an asset is part of the portfolio, independently of the amount of strictly positive portfolio shares. Drawing from these extensions of MPT, we have to consider variable and fixed transaction costs in our model (*Modeling Issue 2*).

3. *Steady marketability and divisibility*

In a perfect market the assets can be traded any time in an arbitrary volume at the market price. Moreover, the traded assets are characterized by infinite divisibility. Financial assets are often characterized both by steady marketability – at least when the markets are open and not too thin – and fine granularity.

Verhoef (2002) states that MPT is not adequate to support IT decision making because IT investments become illiquid by their conversion into software functionality. Verhoef’s critique is appropriate for the class of problems where *project selection* is the issue of interest. It does not

---

[5] In this context, taxes can also be modeled as either fixed or variable transaction costs. To keep things simple, we will not consider taxes in this paper.

comprise the *problem of allocating* already acquired, i.e., selected, software development projects to different available sites. They may be shifted from one delivery center to another at any time.[6] Thus, this aspect leads to no further modeling issue.

The condition of infinite divisibility does not hold for software development projects. Usually, they can just be cut in coarse modules as described in the real world decision problem of ACME. Thus, the proportion of a project that can be conducted at a specific site has to be modeled discretely (*Modeling Issue 3*) instead of continuously as assumed in MPT.

4. *No controls*

In a perfect market there are no barriers to entry and exit, respectively, or any further regulatory constraints. The ITSP owns the sites (corresponding to "free entry") and can allocate the software development projects to the available sites at its own discretion (no regulatory constraints). Thus, this condition leads to no further modeling issue.

5. *Market participants aim to maximize their utility*

In perfect markets the participants aim to maximize their expected utility. A utility-maximizing market participant in MPT is a risk-averse decision maker. Markowitz (1952) states "that the investor does (or should) consider expected return a desirable thing and variance of return an undesirable thing." Consequently, in MPT so-called efficient portfolios are identified by minimizing risk for a given expected return or maximizing return for a given level of risk. Applying this to our decision problem we should be able to identify efficient SPC portfolios where

1.) for given expected costs, no other SPC portfolio can be found with a lower risk,

2.) for given risk, no other SPC portfolio can be found with lower expected costs, and

---

[6] This shifting of projects between sites will lead to transaction costs. Their consideration has already been called for in *Modeling Issue 2*.

3.) no other SPC portfolio can be found with lower expected costs as well as a lower risk.

However, an ITSP like ACME wants to identify *one optimal portfolio* that maximizes the value contribution rather than a *set of efficient portfolios*. One specific function to select an optimal portfolio, which assigns to each possible risk/cost combination a specific value, is not explicitly proposed in MPT, but different functions, all integrating risk and return in different ways, are discussed (e.g. Markowitz 1959, Elton et al. 2006). These functions are often denoted as *preference functions*. To ensure that the decisions based on a preference function are rational in terms of utility maximization, it has to be compatible with the Bernoulli principle[7]. The selection of such a preference function that is rooted in the theory of expected utility and that is applicable as well as intuitive denotes another modeling issue that has to be addressed (*Modeling Issue 4*).

6. *Perfect information*

“The process of selecting a portfolio can be divided into two stages: (1) observation resulting in beliefs about the future performance of assets and (2) based on these beliefs a portfolio of assets is selected” (Markowitz 1952). Markowitz’s paper is only concerned with the second stage because in a perfect market for every available asset, each investor has perfect information about the discounted expected return and the risk represented by the variance of the return. Furthermore, each investor has perfect information about the pair-wise correlation between different assets. The existence of numerous financial analysts continuously estimating the necessary parameters and disagreeing with each other may be taken as evidence that this condition does not hold in financial markets. Still,

---

[7] The Bernoulli principle states that a utility function exists that assigns each possible outcome of a random variable to an unequivocal utility value (in decision theory also called “Bernoulli utility” (Bernoulli 1738) or “von Neumann-Morgenstern-utility” (Neumann and Morgenstern 1944)) and that the decision maker decides with the objective to maximize his/her expected utility.

historical data about price movements is available in abundance and estimations – especially of the variance and correlation coefficients – are usually based on these data.

Expected costs, risks and interdependencies that are associated with specific SPCs are even harder to estimate. Thus, we have to address stage (1) of the portfolio selection process and illustrate, how these input data can be estimated in a real world business situation (*Modeling Issue 5*).[8]

MPT can come in different versions that vary with respect to the allowance of short sales and the opportunity for riskless lending and borrowing (Elton et al. 2006). The situation where short sales are not allowed and no opportunity for riskless lending and borrowing exists matches the best with our decision problem. The correspondence to short sales in financial markets would be SPCs with a negative share. Apparently less than no implementation of a project module at a specific site is not possible. Riskless lending and borrowing – among other things – provides the opportunity to invest some funds of the total amount that shall be invested in riskless assets. In our setting, an ITSP has already made contracts with clients. Thus, a decision for conducting only a part of a project is not a feasible solution. The necessary lines of code have to be implemented and there is no "riskless investment opportunity". This *full allocation* of each project has to be reflected in our model (*Modeling Issue 6*).

To summarize, we identified six modeling issues, which have to be considered in developing our Site Selection Model for ITSPs:

- *Modeling Issue 1*: Site/project combinations (SPCs) have to be the available assets.
- *Modeling Issue 2*: Variable and fixed transaction costs have to be considered.
- *Modeling Issue 3:* Portfolio shares have to be discrete.
- *Modeling Issue 4:* A preference function has to be chosen to decide on the optimal portfolio.

---

[8] In the strict sense, this does not represent a *modeling issue* but an issue that has to be taken into account in the *application*. But since we have to keep in mind the applicability when designing the model, we still call it a modeling issue here.

- *Modeling Issue 5*: It has to be illustrated how the input data of the model can be determined in a real world business situation.
- *Modeling Issue 6:* A full allocation of each project to the available sites has to be ensured.

As we will illustrate in the following, these modeling issues do not impose serious limitations on the computability and the interpretability of results achieved with an MPT-based sourcing model.

## Development of a Site Selection Model for IT Services Providers

Based on the modeling issues discussed in the last section, we develop a decision model for global sourcing of software development projects based on MPT in the following. First, we introduce necessary notations and assumptions. Second, we describe the model design. Finally, we suggest an adequate optimization technique to solve the resulting optimization problem.

### *Notations and Assumptions*

An ITSP simultaneously conducts $M \in IN$ software development projects. A project is running over multiple periods.[9] For project $m \in \{1,2,\ldots M\}$ the ITSP receives a contractually fixed income. $Size_m \in IN$ denotes the number of lines of code that has to be produced per period for project *m*. The production of one line of code comprises not only its implementation but also its definition, design and test.

To conduct the projects, the ITSP has $N \in IN$ available sites. Site $n \in \{1,2,\ldots N\}$ can handle each project. A site is not limited by capacity constraints.[10] Each site is specialized in different project types because of different competencies in branches, technologies etc. This results in different productivities for each SPC. Thus, in combination with the site specific loaded costs, each SPC is characterized by different

---

[9] As we use net present values in our model we do not need to introduce a parameter for the number of periods.

[10] Note the difference between *capacity constraints* and *resource conflicts*. Capacity constraints refer to a situation where a site is limited in the overall number of lines of code that can be produced at a specific point in time. Resource conflicts occur unexpected and have a negative impact on the LOC-costs of more than one project.

development costs per line of code (LOC-costs). To ensure an unequivocal assignment of project *m* and site *n* to SPC *g* we use equation 1).

(1) $g = (n-1) \cdot M + m$

By defining SPCs as the homogenous assets for constructing an optimal portfolio, we address *Modeling Issue 1*.

Unexpected external effects ranging from temporary unavailability of required experts to political crises at a site may occur during the planning horizon of the projects. For that reason the LOC-costs for each SPC cannot be determined deterministically. To consider risk we make the following assumption.

> **Assumption 1:** *The LOC-costs are uncertain, given net present values, which are represented by a normally distributed* ($N(\mu,\sigma)$) *random variable* $\tilde{C}_g$. *Risk is understood as possible negative or positive deviation from the given expected value* $E(\tilde{C}_g) = \mu_g$ *and is quantified by the given standard deviation* $\sigma(\tilde{C}_g) = \sigma_g$.

Note that the optimization results would not change if we considered just the downside risk of $\mu_g$ using the semi-standard deviation, since we assume a symmetric distribution here. We follow standard MPT and use the standard deviation.

Since the occurrence of a risk, like the unexpected illness of a specialist, may not only affect the LOC-costs of one SPC, we also have to consider stochastic interdependencies among the SPCs. Such interdependencies can be represented by a correlation coefficient $\rho_{g,h} \in [-1,1]$, where $g$ and $h$ are representing one SPC each.

Besides development costs, globally distributed work causes transaction costs which have to be considered in the model (cf. *Modeling Issue 2*). As variable transaction costs generally increase with the size of the project, we include them in analogy to Pogue (1970) in the LOC-costs for each SPC. Fixed transaction costs $C^f_g$ are independent of the project size. They occur once during the planning and set up of a software development project *m* at site *n*, and their height is known with certainty.

To determine the set of efficient SPC portfolios in a first step, we have to calculate the expected portfolio LOC-costs $E(\tilde{C}_{PF}) = \mu_{PF}$ and the portfolio risk $\sigma(\tilde{C}_{PF}) = \sigma_{PF}$. Let $w_g$ represent the percentage share of the lines of code to be developed for project *m* at site *n*, i.e. for SPC *g*, relative to the number of lines of code of the whole project portfolio. These portfolio shares $w_g$ constitute the decision variables of the optimization. The following constraint defines the portfolio shares $w_g$ with *K* possible discrete values $r_k$ between 0 and 1 (*Modeling Issue 3*) and has to hold in the optimization.

(2) $w_g \in \{r_1, r_2, ..., r_K\}$ with $r_k \in [0,1],\ K \in IN$

To ensure the full allocation mentioned in *Modeling Issue 6* we define another constraint represented in equation (3). We equate the sum of the portfolio shares over all sites for project *m* with the number of lines of code of project *m* divided by the number of lines of code of the whole project portfolio. Thereby, it is ensured that the necessary and predefined number of lines of code for project *m* is fully allocated to the available sites. As this constraint has to hold for each project *m*, we ensure that each possible portfolio comprises the predefined total number of lines of code that have to be developed and the sum of all calculated portfolio shares equals 1.

(3) $\sum_{n=1}^{N} w_g = \frac{Size_m}{\sum_{l=1}^{M} Size_l} \quad \forall m \quad \text{with } g = (n-1) \cdot M + m$

To select the optimal portfolio from the efficient portfolios we need a preference function that is rooted in the theory of expected utility (*Modeling Issue 4*). The preference function shall integrate expected costs and risk of a portfolio while considering the risk preferences of the decision maker. Therefore, a further assumption is required.

**Assumption 2:** *A utility function* $u(\tilde{C}_{PF})$ *exists which assigns a specific utility to every realization of the random variable* $\tilde{C}_{PF}$. *We assume a risk-averse decision maker who maximizes utility by taking into account uncertain costs.*

The parameters of the model are summarized in a Table in the Appendix.

***Model Description***

To determine an optimal portfolio of SPCs, we have to calculate the expected LOC-costs per line of code for each specific portfolio. Therefore, we sum up the expected LOC-costs per line of code of each SPC weighted with the adherent portfolio shares and the arising fixed costs. As the fixed costs are absolute values, we have to normalize them to one line of code per SPC to get a correct value for the expected LOC-costs per line of code of a SPC portfolio. Fixed costs of a SPC arise only if a SPC is part of the considered portfolio. Building on Patel and Subrahmanyam (1982), we model them to be independent of the magnitude of strictly positive portfolio shares. We use the signum function (Courant and John 1965) to include the fixed costs into the model. It returns either 0 if $w_g = 0$ or 1 if $w_g > 0$.

$$(4)\ E(\tilde{C}_{PF}) = \sum_{g=1}^{N \cdot M} (w_g \cdot \mu_g + \operatorname{sgn}(w_g) \cdot \frac{C_g^f}{\sum_{l=1}^{M} Size_l}) = \mu_{PF}$$

Consistent with MPT, the portfolio risk is calculated based on the individual risk per SPC, the pair-wise correlation among all SPCs and the portfolio shares of the single SPCs as follows:

$$(5)\ \sigma(\tilde{C}_{PF}) = \sqrt{\sum_{g=1}^{N \cdot M} \sum_{h=1}^{N \cdot M} w_g \sigma_g w_h \sigma_h \rho_{gh}} = \sigma_{PF}$$

To select an optimal portfolio we have to choose a suitable preference function (*Modeling Issue 4*). Daniel Bernoulli mathematically analyzed decisions in chance situations using the famous St. Petersburg paradox (Bernoulli 1738). Based on his work and its axiomatic foundation in (Neumann and Morgenstern 1944), decision makers base their decisions on the expected utility of the returns rather than the utility of the expected returns. Besides these formal mathematical treatments of chance situations, the notion of risk-averse decision makers is also incorporated in the widely applied and intuitive "$\mu$-$\sigma$-rule", also denoted as a "classical decision rule" (Schneeweiss 1967). The $\mu$-$\sigma$-rule limits the analysis of chance situations to two influencing factors: mean and variance. It combines these factors in an intuitive way: Given the same expected return, the higher the risk of an action alternative, the less attractive this alternative shall appear to a decision maker. Let $\Phi_{PF}$ denote the preference function of a decision maker. A very popular $\mu$-$\sigma$-rule as a preference function (Hanink 1985, Elton et al. 2006, Copeland et al. 2005) is given by:

$$(6)\quad \Phi_{PF}(\mu_{PF}, \sigma_{PF}) = -\mu_{PF} - \frac{a}{2}\sigma_{PF}^{2} \rightarrow \text{max!}$$

This $\mu$-$\sigma$-rule is the only one producing consistently rationale decisions in terms of utility maximization subject to the constraint of normally distributed random variables (Schneeweiss 1967, Freund 1956).[11] Based on our Assumption 1 and 2, this $\mu$-$\sigma$-rule is adequate for our decision problem.

Note that the expected portfolio LOC-costs enter negatively into the preference function to get an objective function to be maximized. The Arrow-Pratt parameter *a* represents the individual level of risk aversion (Arrow 1970) and is the measure of absolute risk aversion.[12] It is called absolute risk aversion because it measures risk aversion for a given level of wealth (Copeland et al. 2005). That is, with changing wealth, the attitude towards risk may change as well. A positive value for *a* is consistent with a risk-averse decision maker. The higher the absolute value of *a*, the less attractive the riskier alternatives, given the same expected return.

---

[11] To be precise, we should note that this preference function is not a utility function in a strict sense, but it is consistent with the Bernoulli-compatible utility function $u(x)=1-e^{-ax}$ (Schneeweiss 1967).

[12] More formally, let $u$ denote the utility function and $x$ a random variable; then the Arrow-Pratt parameter is defined as $-u''(x)/u'(x)$.

The objective function (7) is based on the preference function (6) and is composed of two parts, the expected portfolio LOC-costs from (4) and the standard deviation of the portfolio representing the portfolio risk from (5). The objective function displays the average risk adjusted costs per line of code of the whole sourcing portfolio.[13]

$$(7)\max\ \Phi_{PF} = -\sum_{g=1}^{N\cdot M}(w_g\cdot\mu_g + \operatorname{sgn}(w_g)\cdot\frac{C_g^f}{\sum_{l=1}^{M} Size_l}) - \frac{a}{2}\cdot\sum_{g=1}^{N\cdot M}\sum_{h=1}^{N\cdot M}\sigma_g\cdot\sigma_h\cdot w_g\cdot w_h\cdot\rho_{gh}$$

s.t.: (I) $w_g \in \{r_1, r_2, ..., r_K\}$ with $r_k \in [0,1], K \in IN$

(II) $\sum_{n=1}^{N} w_g = \frac{Size_m}{\sum_{l=1}^{M} Size_l} \quad \forall m \quad \text{with}\ g = (n-1)\cdot M + m$

***Optimization Technique***

As we modeled the portfolio shares as discrete values and the behavior of the fixed costs with the signum function, the objective function is discontinuous and cannot be solved by a closed-form expression. Such an objective function can be solved by a permutation of all feasible portfolio shares for any SPC. For each permutation we calculate $\mu_{PF}$ and $\sigma_{PF}$ and by applying equation (6) on each $\mu_{PF}/\sigma_{PF}$ combination, we are able to determine the risk adjusted costs for each feasible site project portfolio. By selecting the portfolio with the lowest risk-adjusted costs per line of code we have determined the optimal portfolio. This optimization technique is illustrated within the application of our approach in the following section.[14]

---

[13] Note that it suffices to minimize the average risk adjusted costs per line of code of the whole portfolio since we demand a full allocation of all lines of code to be implemented (*Modeling Issue 6*).

[14] Having a large number of feasible portfolio shares, this approach may require time consuming calculations. If this calculation effort is considered to be too high, it can be reduced by a heuristic like the *subtract-add-approach* described in Buhl and Heinrich (2008).

## Application: Real World Business Case of ACME

To address *Modeling Issue 5* we illustrate how the model presented in the previous section may be applied by continuing the real world business case of ACME answering the following major questions:

1. How may the expected LOC-costs and the risks of the SPCs be estimated?
2. How may the interdependencies among the SPCs be parameterized using correlations?
3. How may the model be applied to demonstrate the advantages of our model compared to the current approach of ACME?

### *Estimating LOC-Costs and Risks*

To estimate the development costs of the three projects, ACME used the cost estimation model COCOMO II. To have a better basis for a comparison, we based our estimations on this approach, too.

The COCOMO equation (8) specifies how the effort (*PM)* measured in person months depends on the size of the project (*Size*) measured in lines of code. This dependency is influenced by scale factors *SF*s, effort multipliers *EM*s and the calibration constants *A* and *B*.[15] The exponent including the *SF*s represents the (dis-)economies of scale in software development projects and the *EM*s represent the quality of the capabilities of sites and projects, respectively. The calibration constants *A* and *B* have to be individually set according to the user's general ability to implement software. Low numerical values for *SF*s, *EM*s and the calibration constants represent a high productivity in producing software and vice versa.

(8) $PM = A \cdot \prod EM \cdot Size^{\,B + 0.01 \cdot \sum SF}$

However, using COCOMO II, ACME did not cover differing productivities properly among the different sites because they only assessed the project effort *PM* onshore and multiplied this effort with the loaded costs of the three sites to get the development costs for each site. In addition, some *EM*s like ACAP

---

[15] The different *EM*s and *SF*s with the associated domains of definition are listed in Boehm et al. (2000).

(Analyst Capability) and PCAP (Programmer Capability) are differing among the sites resulting in productivity differences. To consider this, the effort *PM* has to be estimated for each SPC separately.

COCOMO II also proposes a site specific *EM* called SITE (Multi Site development) to take additional effort for international distributed work and the resulting variable transaction costs into account.

To estimate the development costs considering productivity differences and variable transaction costs for each SPC, we proceeded as follows: First, we took the already estimated COCOMO input data of the three projects for the onshore site Germany and adjusted the site specific *EM*s (including the *EM* SITE) for the near- and offshore delivery centers based on experiences and a consultation of the executives of ACME. After these adaptations, we had a set of COCOMO input data for each SPC.

Second, we did the following calculations for each of the three sites: Based on the COCOMO input data, we calculated the effort of each possible project module[16] using equation (8). To obtain the effort of each project, we summed up the calculated efforts of the modules which belong to one project. Thus, we obtained the effort of each SPC, assuming that the whole project is conducted at the respective site.

Third, we calculated the loaded costs. We took the average labor costs per person and month as arithmetic mean of the implemented team structure, which is globally homogenous at ACME. The teams are composed of a manager, two experienced team leaders and six software engineers. Additional costs for benefits, space, and overheads were deduced from historical data of ACME. Furthermore, the loaded costs were estimated to grow in Germany by 3%, in the Czech Republic by 6%, and in India by 10% per year. Using these data we were able to calculate the loaded costs at each site for the next 3 years.

---

[16] We did not calculate the effort based on the number of lines of code of each project, but just for each project module. The diseconomies of scale represented by the exponent of the COCOMO equation 8) describe an overproportional increase of the complexity that is related to the project size. As the projects can be cut in several project modules, which can be implemented relatively isolated from each other, the increasing complexity can only be justified within the modules but not for the overall project. This was particularly true for the project 2 and 3, where bundles of loosely coupled services had to be developed.

Multiplying the effort of each SPC resulting from the second step and the periodical site specific loaded costs resulting from the third step, we obtained the development costs for each SPC and each period.

Finally, we divided the periodical development costs for each SPC by the associated project size $Size_m$ and discounted the resulting average periodical development costs per line of code by the risk free rate.[17] This results in the net present value of the development costs per line of code $DC_g$ for each SPC $g$.

To obtain the LOC-costs $\tilde{C}_g$ as a random variable and the corresponding first two moments $\mu_g$ and $\sigma_g$, we estimated the risks for each SPC. Based on the idea of Aubert et al. (1998), we defined independent scenarios for each SPC by identifying project and site specific risk influence factors. For each scenario we estimated the expected change of $DC_g$ ($\Delta DC_g$) and the (subjective) probability of occurrence ($p_g$) based on experienced data and a consultation of the executives of ACME.

How we estimated the expected LOC-costs ($\mu_g$) and the risk ($\sigma_g$) is illustrated by the following example. We consider project $n$=1 (“Introducing a brokerage and financial asset administration system”) conducted at site $m$=3 (“India”). Using equation (1) this SPC is represented by $g$=7. Resulting from the previously described approach, we calculated $DC_7$ to \$46.58. Table 2 summarizes the selected risk scenarios with their probabilities of occurrence and the $\Delta DC_7$ (estimated by executives of ACME).

**Table 2. Risk scenarios (Extract)**

| # | Description of the scenario | $p_7$ | $\Delta DC_7$ |
|---|---|---|---|
| 1 | The software shall be developed on a new platform. This was classified as a technology risk which occurs if the development does not work as planned on the new platform. | 20% | + \$8.00 |
| 2 | The IT infrastructure in Bangalore has not been running stable in the last few months. Further breakdowns during the runtime of the project will lead to a decreasing productivity. | 30% | + \$10.00 |
| 3 | Two new specialized managers are currently being sought on the very tight labor market in India for the site in Bangalore. If they can be hired the productivity increases. | 20% | - \$8.00 |
| 4 | To conduct the project, an intensive collaboration of three brokerage specialists is required. Their availability is not secured. Therefore external specialists may have to be acquired. | 40% | + \$10.00 |

---

[17] As we consider risk separately by the standard deviation, a risk adjustment of the discount rate as proposed by the DCF approaches is not appropriate. Using a risk-adjusted discount rate and the standard deviation would mean we consider risk twice.

Based on these independent scenarios, the associated probabilities and the resulting development costs, we constructed the decision tree represented in Figure 1.

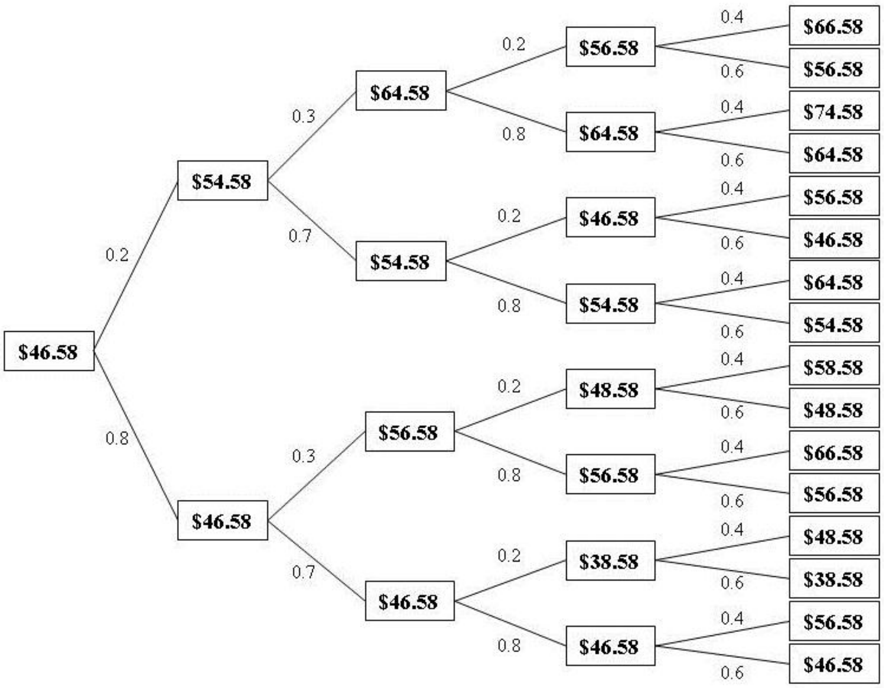

**Figure 1. Decision tree for LOC-costs for *n*=3 (IND) and *m*=1**

We used the maximum likelihood method (Eliason 1993) to estimate the two moments of the normally distributed LOC-costs $\tilde{C}_7$. Based on the resulting instances of the development costs per line of code (values of the end nodes) and the according conditional probabilities, we were able to calculate the expected LOC-costs for project 1 being carried out in India at $\mu_7$ = \$53.58 and the respective standard deviation at $\sigma_7$ = \$8.09.

The fixed transaction costs for each SPC could be taken from the business cases. They consisted of negotiation, planning and traveling costs for the setup of the project. The greater cultural and geographical distance was the reason for the higher fixed costs for the Indian site in comparison to the Czech site.

With this procedure we estimated all expected LOC-costs and their standard deviations per line of code as well as the absolute fixed costs as listed in Table 3.

**Table 3. Input parameters**

| | $n$=1 (GER), $m$=1 → $g$=1 | $n$=1 (GER), $m$=2 → $g$=2 | $n$=1 (GER), $m$=3 → $g$=3 | $n$=2 (CZ), $m$=1 → $g$=4 | $n$=2 (CZ), $m$=2 → $g$=5 | $n$=2 (CZ), $m$=3 → $g$=6 | $n$=3 (IND), $m$=1 → $g$=7 | $n$=3 (IND), $m$=2 → $g$=8 | $n$= 3 (IND), $m$=3 → $g$=9 |
|---|---|---|---|---|---|---|---|---|---|
| $\mu_g$[$] | 88.50 | 116.72 | 117.24 | 70.95 | 93.57 | 93.99 | 53.58 | 70.66 | 70.98 |
| $\sigma_g$[$] | 4.43 | 5.84 | 5.86 | 6.39 | 8.42 | 8.46 | 8.09 | 10.56 | 10.65 |
| $C^f_g$[$] | 0 | 0 | 0 | 110,000 | 100,000 | 100,000 | 160,000 | 200,000 | 200,000 |

### *Estimating Interdependencies*

The correlation coefficients were also determined together with experts of ACME. The starting point was the recognition of two different correlation situations:

(1) Correlation of projects that are conducted on the same site (project dependencies).

(2) Correlation of projects that are conducted on different sites (site dependencies).

Project dependencies typically result from *resource dependencies*. Resource dependent projects compete for similar resources. In consequence, projects were rated as being positively correlated. If projects did not compete on any resources, they were rated as being independent from each other (correlation of zero). Since it was very hard to come up with exact figures for the positive correlations of all pairs of projects, we defined instead three categories of potential resource conflicts and four classes of magnitudes of resource dependencies. Resource conflicts could refer to an unexpected competition for skilled IT workers with

- specific industry/functional knowledge,
- specific technical knowledge, or
- knowledge of client specific processes and problems.

The four classes of dependencies ranged from “high” to “none”. For each category of resource conflict every possible pair of projects was analyzed. To keep our approach simple and to avoid feigned accuracy, we decided for a binary instead of a fuzzy approach when rating two projects as being dependent with respect to a specific kind of resource conflict. That is, for each category of resource conflict a pair of projects could either have a conflict or not. The overall dependency of two projects was rated “high”, if they competed for specifically trained or skilled human resources in all three knowledge domains. If two out of three categories were affected, the dependency was rated “mediocre”. If just one category was affected, the dependency was rated “low”. In our case all projects needed the same industry/functional knowledge and the projects 2 and 3 both required the same technical knowledge. Table 4 summarizes this procedure and lists the assigned correlation coefficients associated with each rating as well as the project combinations.

**Table 4. Project dependency classes based on potential resource conflicts**

| Class | High dependency | Mediocre dependency | Low dependency | No dependency |
|---|---|---|---|---|
| Number of resource conflicts | 3 | 2 | 1 | 0 |
| Used correlation | 1 | 0.67 | 0.33 | 0 |
| Project combination | - | Project 1 and 2 | Project 1 and 3<br>Project 2 and 3 | - |

For projects that are conducted on different sites, the correlation coefficients have to capture site dependencies. These dependencies are concerned with the unexpected occurrence of political, legal and economic risks as well as cultural differences. For example, projects on sites in adjacent and intensely integrated countries – like Germany and France – are in general strongly positively correlated because of spillover effects between these countries. For instance, an economic crisis in one country will also affect the labor market in the other country. Consequently, conducting two projects on sites lying in different and less integrated regions may help to reduce the risk in the site/project portfolio. Note, that at the same time, transaction costs may increase. This constitutes a trade-off, which is accounted for in the model.

To determine the site dependencies between two projects that are conducted on different sites, we refer to financial theory. The integration of two markets (or regions) can be described by the co-movement of stock market returns. Risks and spillover effects between markets (or regions) are reflected in the overall market prices. In order to separate these broader effects from firm-specific effects, we make use of cross-industry indices consisting of large and midcap companies. We used the gross total return Morgan Stanley Capital International (MSCI) index in local currency[18] including a broad selection of listed large and midcap companies for each country where a site is situated, i.e. MSCI Germany (Bloomberg ticker symbol: GDDLGR), MSCI Czech Republic (GDLESCZ), and MSCI India (GDLESIA). Based on the discrete daily returns of the respective indices from 2007 until 2009, we determined the correlation coefficients. In order to check for robustness, we calculated the correlation coefficients also for the two year period 2008 and 2009 as well as just for 2009. All coefficients were significant at the 5%-level and largely robust over time. Table 5 summarizes the correlation coefficients between the different sites.

**Table 5. Site Correlation**

| | n=1 (GER) | n=2 (CZ) | n=3 (IND) |
|---|---|---|---|
| n=1 (GER) | 1 | 0.551 | 0.455 |
| n=2 (CZ) | 0.551 | 1 | 0.445 |
| n=3 (IND) | 0.455 | 0.445 | 1 |

The figures were consistent with our expectations that the correlation between the neighboring countries Germany and Czech Republic is higher compared to the correlation between the European countries and India. Moreover, due to the higher economical integration we expected a higher correlation between Germany and India compared to the Czech Republic and India, which is also consistent with our results. In addition, the magnitude of the correlation coefficients is consistent with results reported e.g. by Elton et al. (2006), who find an average correlation of 0.48 among 15 countries in the period 1991 – 2000.

---

[18] Exchange rate risk is typically fully hedged by ACME and therefore not considered here.

Likewise, in their study for the period 1960 – 1990 Login and Solnik (1995) found that the lowest correlation coefficient is 0.24 (between Germany and Japan) and the highest is 0.71 (between Canada and the USA). The average coefficient in their study is around 0.5.

A special case denotes the potential situation when two or more delivery centers are located in the same country but different regions. For example, many ITSP maintain several locations in India. If the physical distance of two centers is large and if there are substantial cultural differences, the correlation situation (2) applies. In such a case regional instead of country-wide stock market indices should be identified and used to capture the political and economical integration of the two sites. If the physical distance is small and the cost structures and productivities are (mostly) the same, the resources could be shifted between these sites. Thus, the two sites could be included just as one combined site in the decision process.

With these considerations and several validation checks with business case experts of ACME, we ended up with the following correlations for all SPCs as shown in Table 6.

**Table 6. Correlation coefficients among the SPCs**

| | $n$=1 (GER), $m$=1 → $g$=1 | $n$=1 (GER), $m$=2 → $g$=2 | $n$=1 (GER), $m$=3 → $g$=3 | $n$=2 (CZ), $m$=1 → $g$=4 | $n$=2 (CZ), $m$=2 → $g$=5 | $n$=2 (CZ), $m$=3 → $g$=6 | $n$=3 (IND), $m$=1 → $g$=7 | $n$=3 (IND), $m$=2 → $g$=8 | $n$= 3 (IND), $m$=3 → $g$=9 |
|---|---|---|---|---|---|---|---|---|---|
| $n$=1 (GER), $m$=1 → $g$=1 | 1 | 0.33 | 0.33 | 0.551 | 0.33 | 0.33 | 0.455 | 0.33 | 0.33 |
| $n$=1 (GER), $m$=2 → $g$=2 | 0.33 | 1 | 0.67 | 0.33 | 0.551 | 0.67 | 0.33 | 0.455 | 0.67 |
| $n$=1 (GER), $m$=3 → $g$=3 | 0.33 | 0.67 | 1 | 0.33 | 0.67 | 0.551 | 0.33 | 0.67 | 0.455 |
| $n$=2 (CZ), $m$=1 → $g$=4 | 0.551 | 0.33 | 0.33 | 1 | 0.33 | 0.33 | 0.445 | 0.33 | 0.33 |
| $n$=2 (CZ), $m$=2 → $g$=5 | 0.33 | 0.551 | 0.67 | 0.33 | 1 | 0.67 | 0.33 | 0.445 | 0.67 |
| $n$=2 (CZ), $m$=3 → $g$=6 | 0.33 | 0.67 | 0.551 | 0.33 | 0.67 | 1 | 0.33 | 0.67 | 0.445 |
| $n$=3 (IND), $m$=1 → $g$=7 | 0.455 | 0.33 | 0.33 | 0.445 | 0.33 | 0.33 | 1 | 0.33 | 0.33 |
| $n$=3 (IND), $m$=2 → $g$=8 | 0.33 | 0.455 | 0.67 | 0.33 | 0.445 | 0.67 | 0.33 | 1 | 0.67 |
| $n$= 3 (IND), $m$=3 → $g$=9 | 0.33 | 0.67 | 0.455 | 0.33 | 0.67 | 0.445 | 0.33 | 0.67 | 1 |

Besides the estimations for the expected LOC-costs, risks and correlation coefficients for each SPC, the estimation of the risk aversion parameter is a major challenge. For this we compared the risk surcharges ACME chose in its former approach with the estimated variances of our approach. We found out that the values are linearly dependent on each other, implying that a constant risk aversion parameter *a* with the value of 2.0 is reasonable.

### *Applying the Model*

To illustrate the enhancements of our model compared to ACME's industry-typical state-of-the-art DCF approach, we apply our model in two ways.

(1) ITSPs generally make their sourcing decisions based on isolated calculations for each project. To clarify the enhancement of our model in terms of the determination of costs - including differing productivities and transaction costs – as well as risks and site dependencies, we calculate an optimal allocation over all sites using an isolated project valuation. This allocation does not include project dependencies. Further on, this is called the "isolated approach".

(2) In a second step, we consider projects in a project portfolio context. We generate an optimized site-project-allocation, additionally taking into account project dependencies. In the following, this is called the "integrated approach".

To exemplify the enhancement of our model, we calculate the optimized site allocation for the isolated approach and the integrated approach and compare the associated risk-adjusted costs with the risk-adjusted costs of the site allocation originally chosen by ACME.

After the input data were estimated, we applied the previously described optimization technique to determine the optimal portfolio using a Java application that we had implemented for this purpose. First, we generated all feasible SPC permutations and the associated values for the expected LOC-costs and risks. Second, we applied the chosen preference function (equation (6)) to evaluate each permutation.

Third, we chose the SPC permutation that generated the highest value of the preference function. This value denotes the lowest risk adjusted portfolio costs.

In Table 7 we report the results – SPC allocations, the expected portfolio costs $\mu_{PF}$ per line of code, the portfolio standard deviation $\sigma_{PF}$ per line of code as well as the risk-adjusted portfolio costs $\Phi_{PF}$ per line of code – of our calculations. We distinguish between ACME's initial allocation (upper part of the table) and the allocation after the optimization (lower part of the table). Furthermore, we distinguish between the isolated and the integrated approach.

**Table 7 – Results of the different approaches**

| | Isolated approach | | | Integrated approach | | |
|---|---|---|---|---|---|---|
| ACME | *m*=1 | *m*=2 | *m*=3 | *m*=1 | *m*=2 | *m*=3 |
| *n*=1 (GER) / n=2 (CZ) / n=3 (IND) [%] | 80 / 20 / 0 | 75 / 0 / 25 | 67 / 0 / 33 | 80 / 20 / 0 | 75 / 0 / 25 | 67 / 0 / 33 |
| $\mu_{PF}$ [$] | 95.36 | | | 95.36 | | |
| $\sigma_{PF}$ [$] | 5.31 | | | 4.47 | | |
| $\Phi_{PF}$ [$] | 123,60 | | | 115.00 | | |
| OPTIMIZED | *m*=1 | *m*=2 | *m*=3 | *m*=1 | *m*=2 | *m*=3 |
| *n*=1 (GER) / n=2 (CZ) / n=3 (IND) [%][19] | 40 / 20 / 40 | 50 / 25 / 25 | 50 / 17 / 33 | 20 /40 / 40 | 38 / 0 / 62 | 17 / 33 / 50 |
| $\mu_{PF}$ [$] | 85.50 | | | 79.18 | | |
| $\sigma_{PF}$ [$] | 5.79 | | | 5.62 | | |
| $\Phi_{PF}$ [$] | 119.01 | | | 110.81 | | |

With respect to the risk-adjusted costs, two effects are important. First, what is the effect of the optimization? Second, what is the effect of the inclusion of project interdependencies? For the first effect,

---

[19] In our model, the portfolio shares $w_g$ represent the percentage share of the lines of code to be developed for SPC $g$, relative to the number of lines of code of the whole project portfolio. To make the optimized allocation of the integrated approach intuitively comparable to the allocations of the isolated approach we normalized the portfolio shares relative to the number of lines of code of each project.

we compare the ACME allocation with the optimized allocation, both generated with the integrated approach. We find that the risk-adjusted costs are 3.6% lower after the optimization. For the second effect, we compare the risk-adjusted costs of the isolated approach with the integrated approach, both optimized. We find a potential reduction of risk-adjusted costs of 6.9% that are attributable to the additional consideration of project dependencies. Apparently, ACME has not capitalized on all risk-adjusted cost reduction potentials by their initially chosen allocation.

Examining the expected LOC-costs, both optimized portfolios show substantially lower costs by exploiting the risk-carrying capacity of ACME. Taking the expected costs per line of code of the ACME allocation of $95.36 as a basis, the costs decrease after an optimization by 10.3% for the isolated and 17.0% for the integrated approach, respectively. Assuming that ACME has a net profit ratio of 10.0%, the net profit ratio would increase to 20.3% for the portfolio optimized with the isolated approach and to 27.0% for the portfolio optimized with the integrated approach. Thus, an optimal site allocation may help to realize cost reduction potential by an acceptable level of increase in risk.

To have a better basis for a comparison, we also plotted all portfolios, i.e. 26,460 feasible SPC permutations, in a $\mu$-$\sigma$-diagram (see Figure 2, each dot represents a specific portfolio). The portfolios of the integrated and isolated approach are shown in the black scatter plot on the left and the gray scatter plot on the right, respectively. All portfolios at the outer left side of each scatter plot are not dominated by other portfolios and form the efficient frontiers. The risk/cost positions of the initially chosen allocation by ACME are illustrated with the blank circle (isolated approach) and blank square (integrated approach). The optimized allocations are represented by the circle marked with an x (isolated approach) and the square marked with an x (optimized approach).

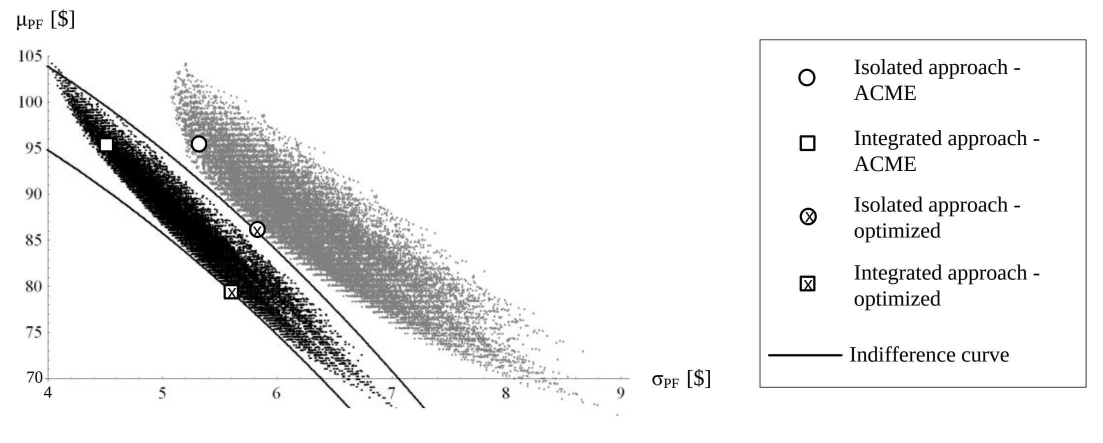

**Figure 2. Risk/cost positions, efficient frontiers and indifference curves**

The chosen site distribution of ACME for both approaches does not lie on the respective efficient frontiers nor does it fit ACME's level of risk aversion because the risk/cost positions of the original portfolio are not situated on the best indifference curve possible.

### *Sensitivity Analysis*

Parameter estimations obviously contain the risk of error of judgment: executives may misjudge potential outcomes or probabilities, there may be a lack of sufficient information to estimate the parameters, historical data may turn out not to be a good estimator for future developments, and so on. To get a feeling for the robustness of the results applying our model, we accomplished a sensitivity analysis changing one input parameter each by plus and minus 5% at a time (c.p.). For each change in an input parameter, we applied our model in two different ways to enrich the expressiveness of our sensitivity analysis.

(1) We performed the typical sensitivity analysis by applying our model to the new parameter set, i.e. the old parameter set with just one parameter changed at a time, and determined the optimal portfolio and accompanying result figures ($\mu_{PF,new}$, $\sigma_{PF,new}$, $\Phi_{PF,new}$, new allocation)

(2) We applied our model to obtain an optimized site distribution using the initial parameter set but afterwards calculated three result figures ($\mu_{PF,old}$, $\sigma_{PF,old}$, $\Phi_{PF,old}$,) using the parameter set with one

changed parameter. This illustrates the case where the optimization is performed using a flawed input parameter. Hence, this approach gives us an idea of how severely incorrect estimations affect the overall outcome.

Table 8 reports the results of the sensitivity analysis. In the first column the values for the initial parameters are reported. Each row consists of two sub-rows. The upper (lower) sub-row contains the results if the parameter value is decreased (increased) by 5% relative to the initial value.

**Table 8. Sensitivity analysis for the optimized integrated approach**

| Parameter Original value | Modified values | $\mu_{PF,old}$ | $\sigma_{PF,old}$ | $\Phi_{PF,old}$ | $\mu_{PF,new}$ | $\sigma_{PF,new}$ | $\Phi_{PF,new}$ | $m$=1 [%] GER/CZ/IND | $m$=2 [%] GER/CZ/IND | $m$=3 [%] GER/CZ/IND |
|---|---|---|---|---|---|---|---|---|---|---|
| Original optimization | | 79.18 | 5.62 | 110.81 | - | - | - | 20 / 40 / 40 | 38 / 0 / 62 | 17 / 33 / 50 |
| $LOC_2$ | 76,000 | 79.09 | 5.61 | 110.56 | 79.09 | 5.61 | 110.56 | 20 / 40 / 40 | 38 / 0 / 62 | 17 / 33 / 50 |
| 80,000/year | 84,000 | 79.26 | 5.63 | 110.96 | 79.26 | 5.63 | 110.96 | 20 / 40 / 40 | 38 / 0 / 62 | 17 / 33 / 50 |
| $E(\tilde{C}_{IND,3})$ | 67.60 | 78.67 | 5.62 | 110.30 | 78.67 | 5.62 | 110.30 | 20 / 40 / 40 | 38 / 0 / 62 | 17 / 33 / 50 |
| 70.98 | 74.53 | 79.71 | 5.62 | 111.34 | 81.84 | 5.42 | 111.25 | 20 / 40 / 40 | 38 / 0 / 62 | 33 / 33 / 34 |
| $\sigma_{CZ,3}$ | 8.056 | 79.18 | 5.60 | 110.48 | 80.33 | 5.49 | 110.43 | 20 / 40 / 40 | 38 / 0 / 62 | 17 / 50 / 33 |
| 8.46 | 8.882 | 79.18 | 5.65 | 111.14 | 81.25 | 5.46 | 111.07 | 20 / 40 / 40 | 38 / 0 / 62 | 50 / 0 / 50 |
| $C^f_{CZ,1}$ | 99,000 | 79.18 | 5.62 | 110.81 | 79.18 | 5.62 | 110.81 | 20 / 40 / 40 | 38 / 0 / 62 | 17 / 33 / 50 |
| 110,000 | 121,000 | 79.18 | 5.62 | 110.81 | 79.18 | 5.62 | 110.81 | 20 / 40 / 40 | 38 / 0 / 62 | 17 / 33 / 50 |
| $\rho_{(GER,m),(CZ,m)}$ | 0.523 | 79.18 | 5.62 | 110.76 | 79.18 | 5.62 | 110.76 | 20 / 40 / 40 | 38 / 0 / 62 | 17 / 33 / 50 |
| 0.551 | 0.579 | 79.18 | 5.79 | 110.85 | 79.18 | 5.79 | 110.85 | 20 / 40 / 40 | 38 / 0 / 62 | 17 / 33 / 50 |
| $a$ | 1.9 | 79.18 | 5.62 | 109.23 | 79.18 | 5.62 | 109.23 | 20 / 40 / 40 | 38 / 0 / 62 | 17 / 33 / 50 |
| 2.0 | 2.1 | 79.18 | 5.62 | 112.39 | 80.23 | 5.52 | 112.35 | 20 / 40 / 40 | 50 / 0 / 50 | 17 / 33 / 50 |

Overall, the allocation is broadly stable but shows project specific-differences. The allocation for project 1 is stable regardless of which parameter is changed. A different optimal allocation for project 2 emerges just for an increase in the Arrow-Pratt parameter. A different optimal allocation for project 3 emerges for an increase in the LOC-costs and an increase or decrease in the standard deviation. However, the result must be put into perspective. We report changes in costs and risk of project 3 and not for the other two projects here. Thus, we should expect a changing allocation, in particular with respect to this project.

More important, comparing the figures $\mu_{PF,old}$, $\sigma_{PF,old}$, $\Phi_{PF,old}$ with $\mu_{PF,new}$, $\sigma_{PF,new}$, $\Phi_{PF,new}$ we can observe the following: As long as the allocation of the initial optimal portfolio and the new optimal portfolio remain the same, there are no differences in these figures. This holds in our case for a change in the lines of code to be implemented, the fixed costs, the correlation coefficient, and a decrease in the expected costs. A slight misestimation of these input parameters has no effect. Higher expected costs, a changing risk assessment as well as a higher Arrow-Pratt parameter lead to a changing allocation and thus to changing result parameters. Still, comparing the risk-adjusted costs per line of code, they remain broadly stable. The changing parameters cause a difference in the risk-adjusted costs of less than 0.1%.

It is obvious that due to coarse project modules, the potential for a misallocation in the presence of small misestimations is very limited. Still, these results are site and project specific and cannot be generalized. Rather, it was our objective to illustrate how a sensitivity analysis can add value and point to estimation-critical parameters. These parameters should get special attention in a re-evaluation of the estimation process.

## Discussion

For a long time ITSP have relied on scoring models based on qualitative ratings (e.g. Pandey and Bansal 2004) and DCF approaches (like ACME) to perform the allocation of software development projects to their globally distributed delivery centers. The resulting site allocations based on these approaches frequently lead to dissatisfying results (like described by the ACME executives). While, scoring models may suggest that one project is better suited to be conducted offshore than another one, they are unable to determine whether or not such a decision contributes to an overall positive net value for the ITSP. Moreover, DCF approaches, like the one applied at ACME, assess sourcing decisions by a monetary value, but they do not take into account interdependencies among projects and sites.

To alleviate this drawback, we adapted MPT to our site selection problem by considering site/project combinations as risky assets, assuming discrete portfolio shares, and factoring in transaction costs as well as dependencies between both projects and sites. To the best of our knowledge, this model represents the first risk/cost integrating approach for a value-based Sourcing Portfolio Management. Furthermore, we applied the model to a real decision situation of our case company. For estimating costs our approach builds on prior work on software cost estimation (Boehm et al. 2000), while our risk assessment is rooted in the literature on IS outsourcing risks (Aubert 1998). The application of our approach lead to a substantially different project allocation with significantly lower costs compared to the initial allocation of our case company. One reason for this is attributed to the fact that previous approaches did not allow for the simultaneous recognition of expected costs (including transaction costs), risks and interdependencies among both projects and sites.

Although we have illustrated the strength of our MPT-based Site Selection Model in a realistic and practical case, we have to point to some critical issues that can threaten the applicability and validity of our analysis. We analyzed the fit of the assumptions of MPT to the decision problem of global sourcing decision making. This resulted in the identification of modeling issues that had to be dealt with when transferring MPT to our problem domain. The issues were addressed in the modeling process, e.g. by relaxing assumptions of a perfect market within our model (for instance the assumption of infinitesimal divisibility). Thus, the application of the model may appear straightforward to the reader based on the discussion of the case and the modeling issues. However, a number of design choices were made resulting in limitations that have to be discussed to understand their implications to management and research.

- **Parameterization Risk**: The input data for the Site Selection Model are based on estimates and experience-based data of the decision makers. Thus, in practice, errors will occur regardless of the accuracy in the estimation procedure. Practitioners should use a sensitivity analysis to identify parameters that are critical for the portfolio allocation.

- **Model Risk**: Choosing MPT and a specific preference function as the basis for our model implies some model risks. For instance, it is not ensured that the estimated LOC-costs are always best described by a normal distribution. In addition, a correlation coefficient measures the *linear* dependency of two random variables (here: LOC-costs). It is not ensured that the dependency between the LOC-costs of two projects is always best described using a linear relationship. Both design choices were necessary to have an applicable model that is still rooted in the theory.

- **Functional project modules**: To properly apply our model, we can divide the software development projects only into *functional project modules* – e.g. encapsulated in Web services – and allocate them to different sites. Cutting projects along the workflows – e.g. conducting design onshore, implementation nearshore, and test offshore – is not possible, because the LOC-costs are measured per line of code including design, implementation and test. By cutting projects along their workflows, the SPCs were no homogenous assets anymore. Thus, if the typical division of labor at an ITSP is structured along the workflows, our approach is not applicable. This denotes an issue for further research. As the *design* of a line of code of a SPC has other expected costs and risks than the *implementation* of a line of code of the same SPC, we would have to design *site/project/workflow combinations* as homogeneous assets to be optimized.
- **Capacity constraints**: In practice, the capacity of sites will typically be limited due to already running projects and limited space, funds and human resources. However, in the current version of the model potentially existing capacity constraints of sites are not considered. It is quite easy, though, to implement further constraints in the optimization. Such constraints may also have the positive effect that potential misestimations of input parameters will have less severe effects. In the context of correlation estimations in MPT, Jagannathan and Ma (2003) examine the effect of placing upper and lower limits on the amount invested in any asset in the global minimum variance portfolio. They found that this produces the same solution obtained by shrinking extreme covariance estimates towards the mean.
- **Income Uncertainty:** By just focusing on costs and risks and assuming that the income of each project is contractually fixed, the decision problem becomes manageable. In practice, apart from fixed price contracts, other contract types such as time-and-material contracts where the income is uncertain depending on progress and success (Gopal et al. 2003) are used as well. In addition

> there is uncertainty about the income caused by the potential insolvency of the client, which is not reflected in the model so far. On the managerial level this limitation is not severe once a fixed price contract has been concluded. Even in the presence of income uncertainty, the projects should be allocated in a risk/cost-optimal way. Still, the integration of income and its uncertainty in the model denotes an interesting perspective for further research. The current approach is only applicable for management *after the conclusion* of (fixed price) contracts. By integrating income uncertainty the approach could be applicable *before the conclusion* of a contract and different contract types could be evaluated endogenously in the model.

Having discussed important limitations of our model, we want to conclude the discussion with a more general perspective on the generalizability and the breadth of the application of our approach. We have illustrated that our proposed model is suitable for the allocation of predefined software development projects on existing, available sites. Beside software development projects, *IT operations* and *business processes* are globally sourced in business practice. The question is whether our approach can be transferred to these problem domains as well. Constructing Site/IT operation combinations and Site/Business Process combinations as *homogeneous assets*, they seem to have similar characteristics compared to software development projects (existence of transaction costs, steady marketability concerning the allocation, no infinitesimal divisibility etc.). Thus, we expect that the approach is also transferable to the problem of site allocation of IT operations and business processes.[20] This analysis concerning the extension of our model to comprise also the optimal allocation of IT operations and business processes is an issue for further research.

This potential for extensions is limited by the class of problems that can be addressed with our approach. The proposed model is built for a periodical allocation of software development projects to available sites.

---

[20] The COCOMO II approach to estimate the costs we used for software development projects cannot be applied there, because it is specific to software development projects.

We can change the location where a software development project is conducted at any time.[21] Thus, the liquidity of the regarded assets is ensured and the model is adequate to solve comparable *allocation problems*. But MPT is not an adequate tool, though, to support periodical *IT investment decision making*, i.e., which IT project should be made, which should be extended, which should be abandoned, and which should not be started at all. IT investments become illiquid after the first selection by their conversion into software functionality (Verhoef 2002). Thus, *investment decision problems* with real and mostly illiquid assets require different solution techniques.

## Conclusion

"A major challenge for IS research lies in making models and theories that were developed in other academic disciplines usable in IS research and practice" (Benaroch and Kauffman 1999). In this paper, we adapted MPT rooted in financial theory to the decision problem of allocating software development projects to available delivery centers of an ITSP and applied it on a real decision situation of an ITSP. Practitioners (i.e., ITSPs) are welcome to use our approach for their site selection problems. We illustrated that the use of our approach may lead to substantial expected overall cost savings. Specifically, our approach supports decision makers to better balance the trade-off between conducting many projects at one site and distributing the projects among different sites. On the one hand, conducting many projects at one site incurs low transaction costs but little risk diversification. Distributing the projects among different sites, on the other hand, incurs higher transaction costs but higher opportunities for risk diversification. So far, this trade-off had to be balanced based on experience and "gut-feeling"; now, a conceptually founded model may be applied that contributes to a more transparent and objective decision making process. The model is the first allowing a simultaneous consideration of expected costs, risks, and interdependencies among both projects and involved sites when making sourcing decisions. The model is

---

[21] Of course, switching a project module from one site to a different site will cause transaction costs, but they are explicitly considered in our model.

directly applicable in practice. Even if practitioners lack the required input data or governance structures, the model still provides them with the key aspects (e.g. project and site dependencies) to be considered when faced with similar sourcing decisions.

Yet, much remains to be done if we are to make sense of MPT in the way finance professionals do: as a means to design risk/return-efficient portfolios of risky financial assets utilizing the effect of risk diversification. Risk diversification is probably the only "free lunch" available in the financial markets. Perhaps the two most important next steps to leverage this insight for allocation problems in IS are (1) to further validate the extent to which the project and site risks are represented by the assumed relationships and distributions (e.g. through extended empirical means) and (2) to broaden the basis for the application of the model. At a time when global sourcing of software development, IT operations, and business processes has become the standard, this may lead us to a new value-based decision framework that is rooted in theory.

## References

Apte, U.M., Sobol, M.G., Hanaola, S., Shimada, T., Saarinen, T., Salmela, T., and Vepsalainen, A.P.J. "IS outsourcing practices in the USA, Japan and Finland: a comparative study," *Journal of Information Technology* (12:4), 1997, pp. 289-304.

Arrow, K.J. "The Theory of Risk Aversion", in Arrow, K.J. (Ed.): *Essays in the Theory of Risk-Bearing*, North-Holland, Amsterdam, Netherlands, 1970, pp. 90-120.

Asundi, J., and Kazman, R. "A Foundation for the Economic Analysis of Software Architectures", in *Proceedings of the 3rd Workshop on Economics-Driven Software Engineering Research (EDSER-3)*, Toronto, Canada, 2001.

Aubert, B.A., Patry, M., and Rivard, S. "Assessing the Risk of IT Outsourcing", in *Proceedings of the 31st Annual Hawaii International Conference on System Sciences*, 6, Hawaii, USA, 1998, pp. 685-692.

Benaroch, M., Kauffman, R. "A Case for Using Real Options Pricing Analysis to Evaluate Information Technology Project Investments," *Information Systems Research* (10:1), 1999, pp.70-86.

Bernoulli, D. “Specimen Theoriae Novae de Mensura Sortis,” *Commentarii Academiae Scientiarum Imperialis Petropolitanae*, Tomus V, 1738, pp. 175-192. (English translation: “Exposition of a New Theory on the Measurement of Risk,” *Econometrica* (22:1), 1954, pp. 23-36).

Boehm, B. *Software engineering economics*. Prentice-Hall, Englewood Cliffs, USA, 1981.

Boehm, B.W., Abts, C., Windsor Brown, A., Chulani, S., Clark, B.K., Horowitz, E., Madachy, R., Reifer, D., and Steece, B. *Software Cost Estimation with COCOMO II*, Upper Saddle River, New Jersey, USA, 2000.

Buhl, H.U., and Heinrich, B. “Valuing Customer Portfolios under Risk-Return-Aspects: A Model-based Approach and its Application in the Financial Services Industry,” *Academy of Marketing Science Review* (12:5), 2008.

Carmel, E., and Agarwal, R. “The Maturation of Offshore Sourcing of Information Technology Work," *MIS Quarterly Executive* (1:2), 2002, pp 65-78.

Carmel, E., and Tjia, P. *Offshoring Information Technology - Sourcing and Outsourcing to a Global Workforce*, Cambridge University Press, Cambridge, UK, 2005.

Copeland, T.E., Weston, J.F., and Shastri, K. *Financial Theory and Corporate Policy*, 4th Edition, Addison-Wesely, Boston, USA, 2005.

Courant, R., and John, F. *Introduction to Calculus and Analysis - Volume One*, Wiley, New York, USA, 1965.

Cusumano, M., MacCormack, A., Kemerer, C.F., and Crandall, B. “Software development worldwide: the state of the practice,” *IEEE Software* (20:6), 2003, pp. 28-34.

Davidson, D. “Offshore outsourcing subtleties”, *Meta Group Report*, 2003.

Dibbern, J., Goles, T., Hirschheim, R.A., and Jayatilaka, B. "Information Systems Outsourcing: A Survey and Analysis of the Literature. *The DATA BASE for Advances in Information Systems*, (35:4) 2004, pp. 6-102.

Dibbern, J., Winkler, J., and Heinzl, A. "Explaining Variations in Client Extra Costs between Software Projects Offshored to India. *MIS Quarterly*, (32:2), 2008, pp. 333-366.

Eliason, S.R. *Maximum Likelihood Estimation: Logic and Practice (Quantitative Applications in the Social Sciences)* Sage Publications, Newbury Park, USA, 1993.

Elton, E.J., Gruber, M.J., Brown, S.J., and Goetzmann, W.N. *Modern Portfolio Theory and Investment Analysis* 7th Edition, Wiley, Hoboken, USA, 2007.

Everest Research Institute (Ed.) *Global Sourcing Annual Report 2005*, Dallas, Texas, USA, 2005.

Freund, R.J. "The Introduction of Risk into a Programming Model," *Econometrica* (24:3), 1956, pp. 253-263.

Gopal, A., Sivarama, K., Krishnan, M.S., and Mukhopadhyay, T. "Contracts in Offshore Software Development: An Empirical Analysis," *Management Science* (49:12), 2003, pp. 1671-1683.

Hanink, D.M. "A Mean-Variance Model for MNF Location Strategy," *Journal of International Business Studies* (16:1), 1985, pp. 165-170.

Hodder, J.E., and Dincer, M.C. "A multifactor model for international plant location and financing under uncertainty," *Computers & Operations Research* (13:5), 1986, pp. 601-609.

Hodder, J.E., and Jucker, J.V. "A Simple Plant-Location Model for Quantity-Setting Firms Subject to Price Uncertainty," *European Journal of Operational Research* (21:1), 1985, pp. 39-46.

Jagannathan, R., and Ma, T. "Risk Reduction in Large Portfolios: Why Imposing the Wrong Constraints Helps," *Journal of Finance* (58:4), 2003, pp. 1651-1684.

Kersten, B., and Verhoef, C. "IT Portfolio Management: A Banker's Perspective on IT," *Cutter IT Journal* (16:4), 2003, pp. 34-40.

Kliem, R. "Managing the Risks of Offshore IT Development Projects," *Information Systems Management* (21:3), 2004, pp. 22-27.

Longin, F., and Solnik, B. "Is the correlation in international equity returns constant: 1960-1990?" *Journal of International Money and Finance* (14:1), 1995, pp. 3-26.

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

Markowitz, H.M. *Portfolio Selection*, Yale University Press, New Haven, USA, 1959.

Neumann, J. v., and Morgenstern, O. *Theory of Games and Economic Behavior*, Wiley, New York, USA, 1944.

Pandey, V., and Bansal, V. "A Decision-Making Framework for IT Outsourcing using the Analytic Hierarchy Process," *Proceedings of the International Conference on Systemics, Cybernetics and Informatics*, Hyderabad, India, 2004, pp. 528-533.

Patel, N.R., and Subrahmanyam, M.G. "A Simple Algorithm for Portfolio Selection with Fixed Transaction Costs," *Management Science* (28:3), 1982, pp. 303-314.

Pogue, G.A. "An Extension of the Markowitz Portfolio Selection Model to Include Variable Transactions' Costs, Short Sales, Leverage Policies and Taxes," *The Journal of Finance* (25:5), 1970, pp. 1005-1027.

Rappaport, A. *Creating Shareholder Value: A Guide for Managers and Investors: The New Standard for Business Performance*, Free Press, New York, USA, 1998.

Santhanam, R., and Kyparisis, G.J. "A decision model for interdependent information system project selection", *European Journal of Operational Research* (89:2), March 1996, pp. 380-399.

Schneeweiss, H. *Entscheidungskriterien bei Risiko*, Springer, Berlin, Germany, 1967.

Valerdi, R. *The Constructive Systems Engineering Cost Model (COSYSMO): Quantifying the Costs of Systems Engineering Effort in Complex Systems*, VDM Verlag, Saarbrücken, Germany, 2008.

Verhoef, C. "Quantitative IT portfolio management," *Science of Computer Programming* (45:1), 2002, pp. 1-96.

Winkler, J.K., Dibbern, J., and Heinzl, A.H. "Success in Offshoring of Application Development – Does Culture matter?", *Proceedings of the 14th European Conference on Information Systems (ECIS)*, Ljunberg, J., and Andersson, M. (Ed.), Göteborg, Sweden, 2006, pp. l640-1650.

# Appendix:

| Input and Output of the Site Selection Model | | |
|---|---|---|
| Symbol | Definition | Parameterization in the real world business case |
| *Input* | | |
| $Size_m$ | Amount of software that has to be produced per period for project *m* measured in lines of code | Given as input data by ACME |
| $E(\tilde{C}_g) = \mu_g$ | Expected costs per line of code of SPC *g* | Estimated using the cost estimation model COCOMO and with the introduced approach based on risk scenarios |
| $\sigma(\tilde{C}_g) = \sigma_g$ | Risk of SPC g | Estimated using a decision tree based on risk scenarios |
| $C^f_g$ | Fixed costs for SPC g | Given as input data by ACME (based on business cases) |
| $\rho_{gh}$ | Correlation coefficient capturing the linear correlation between SPC *g* and *h*. | Estimated using three categories of potential resource conflicts and the MSCI index for each country where a site is located |
| $a$ | Risk aversion of the decision maker | Estimated based on risk surcharges of ACME |
| *Output* | | |
| $w_g$ | Portfolio share of SPC g – *Decision variable* of the optimization | Resulting shares per SPC as optimization result of the described approaches in the real world business case |
| $E(\tilde{C}_{PF}) = \mu_{PF}$ | Expected costs per line of code of the portfolio | Calculated with our Site Selection Model |
| $\sigma(\tilde{C}_{PF}) = \sigma_{PF}$ | Portfolio risk per line of code | Calculated with our Site Selection Model |
| $\Phi_{PF}$ | Risk adjusted costs per line of code of the portfolio | Calculated as result of the preference function |

## Presentations

Earlier versions of this contribution were presented at the INFORMS International Conference Puerto Rico 2007, the SISTM Seminar 2008 at the University of New South Wales and at the International Conference on Information Systems ICIS 2008.

## Acknowledgments

We would like to thank the participants at the INFORMS International Conference Puerto Rico 2007, the SISTM Seminar 2008 at the University of New South Wales and the International Conference on Information Systems ICIS 2008 for their helpful feedback. Moreover, we would like to thank three anonymous Reviewers, the Associate Editor and the Track Chair of the Track "Global IS Management" at ICIS 2008. In addition, we are indebted to three anonymous reviewers and Jens Dibbern, the Senior Editor of The DATA BASE for Advances in Information Systems. Their guidance helped us tremendously to improve the paper. Finally, we would like to thank our case company that provided the data for our calculations and took a lot of time in discussing their valuation approach as well as the associated challenges with us.