OPERATIONS RESEARCH
Vol. 00, No. 0, Xxxxx 0000, pp. 000–000
ISSN 0030-364X | EISSN 1526-5463 | 00 | 0000 | 0001

INFORMS
DOI 10.1287/xxxx.0000.0000

R&D Project Portfolio Analysis for the Semiconductor Industry

Banu Gemici-Ozkan, S. David Wu
Department of Industrial and Systems Engineering, Lehigh University
200 W. Packer Ave. Bethlehem, PA 18015
bag4@lehigh.edu, sdw1@lehigh.edu

Jeffrey T. Linderoth
Department of Industrial and Systems Engineering, University of Wisconsin-Madison,
3226 Mechanical Engineering Building, 1513 University Avenue, Madison, WI 53706-1572
linderoth@wisc.edu

Jeffry Moore
LSI Corporation,
1110 American Pkwy NE, Allentown, PA 18109
jeffry.moore@lsi.com

We introduce a decision support framework for the research and development (R&D) portfolio selection problem faced by a major U.S. semiconductor manufacturer. R&D portfolio selection is of critical importance to high-tech operations such as semiconductor and pharmaceutical, as it determines the blend of technological development the firm must invest in its R&D resources. This R&D investment leads to differentiating technologies that drive the firm's market position. We developed a general, three-phase decision support structure for the R&D portfolio selection problem. First is the scenario generation phase where we transform qualitative assessment and market foresight from senior executives and market analysts into quantitative data. This is combined with the company's financial data (e.g., revenue projections) to generate scenarios of potential project revenue outcomes. This is followed by the optimization phase where a multistage stochastic program (SP) is solved to maximize expected operating income (OI) subject to risk, product interdependency, capacity, and resource allocation constraints. The optimization procedure generates an efficient frontier of portfolios at different OI (return) and risk levels. The refinement phase offers managerial insights through a variety analysis tools that utilize the optimization results. For example, the robustness of the optimal portfolio with respect to the risk level, the variability of a portfolio's OI, and the resource level usage as a function of the optimal portfolio can be analyzed and compared to any qualitatively suggested portfolio of projects. The decision support structure is implemented, tested, and validated with various real world cases and managerial recommendations. We discuss our implementation experience using a case example and explain how the system is incorporated into the corporate R&D investment decisions.

Subject classifications: Research and development/ Project Selection: R&D project interdependency, multi period horizon. Programming/Stochastic: scenario generation; Secondary: Organizational Studies/ Strategy: semiconductor industry
Area of review:
History:

1. Introduction

Effective management of the Research and Development (R&D) portfolio is critical to effective market positioning in high-tech industries such as computers, semiconductors, and pharmaceutical. In these industries, a firm's market position is tied directly to its portfolio of intellectual property (IP), which must be developed, acquired, or licensed. Studies of these industries (c.f., Boisot (1998), Bekkers et al. (2002)) point to the important conclusion that a firm's ability to attain significant market share in any technology area depends on its ownership of the essential IP and its ability to leverage essential IP from other firms.

Notable examples of using portfolio selection and management concepts in strategic planning can be found in pharmaceutical companies. In the pharmaceutical industry, the R&D budget can constitute over 40% of the total operating cost. The drug development portfolio drives the R&D budget, determines the firm's market positioning, and drives the firm's brand image. Major firms try to balance their portfolio of drugs such that high-risk, high-margin specialty drug development projects are mixed with low-risk, high-volume projects. When facing R&D portfolio decisions, not only are companies subject to risks due to market uncertainty, they are also subject to risk factors such as prolonged FDA approval cycles and potential legal liabilities. Pharmaceutical product development exemplifies an emerging trend in the high-tech industries where products are characterized by short life cycle, capital intensive development cycle, and long production lead-time. While our focus in this paper is on the specialty semiconductor industry, the methodologies developed herein can be generalized for other industries such as pharmaceutical, computers and electronics.

Portfolio management for R&D projects involves a great deal of uncertainty. It is due to the highly volatile nature of this industry that the portfolio management issues become so critical. In order to assess revenue impacts of a specific project, it is necessary to consider the full life cycle of the project, from the initiation of the product development, to product launch, to the end of the project life cycle. In this paper we introduce an integrated framework for R&D project portfolio selection and management at a major U.S. semiconductor manufacturer. Our contribution is two-fold. First, we model uncertainty through a scenario structure that combines (a) milestone information during typical R&D projects' life-cycle, and (b) data from daily business transactions. This approach allows us to evaluate the risk associated with each project by evaluating its dependencies to other projects and its potential impacts to the revenue streams. The scenario structure is incorporated into a multi-stage stochastic programming model that, in turn, optimizes the project portfolio. Second, we develop a decision support system (DSS) by combining scenario generation, optimization, and solution modules into an integrated tool. In the next section, we will describe the business context of the specialty semiconductor industry and a typical R&D project's life cycle.

2. Semiconductor Product Development

Semiconductor manufacturers undertake the wafer fabrication, packaging, assembly, and test responsibilities of microelectronic chips. In the specialty semiconductor industry, a majority of the chips are custom-designed to handle special functionality of an electronic device. The brand-carrying customer and the semiconductor manufacturer typically share the responsibilities in new product development (NPD) project, which includes the following steps (see Figure 1):

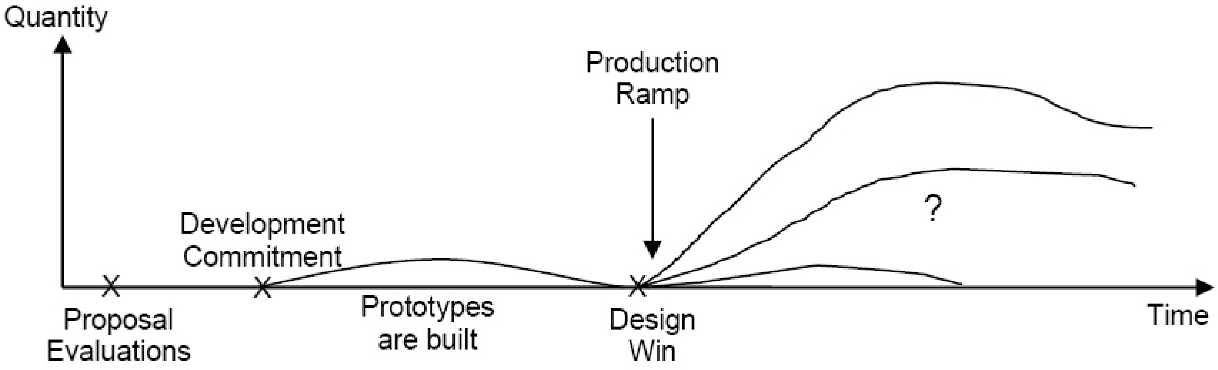
1. **Evaluation of project proposals from customers:** During this stage, the customer and the semiconductor manufacturer discuss and identify the type of technology to be used, technical specifications, expected resource requirements, time to market, and the demand potential. If the project proposal is for a renewal of a previous project, then most of this information already exists. However, if the project proposal calls for NPD, typically very limited information is available. At the end of the evaluation phase, the manufacturer must decide to accept the project or not.

2. **Development Phase:** If the semiconductor manufacturer commits to the project proposal, then development begins. The duration of the development phase depends on the complexity of the design, the skill level of the R&D staff, and the availability of relevant prior technology. If prior technology needs to be extended, or new technology needs to be invented, then investments must be made to develop new intellectual property (IP). A completed design leads to the development of prototypes. If the prototypes meet the specifications, then the development phase continues as scheduled; otherwise the date of the product release must be delayed, and more investment is needed to develop the IP. Once the product is successfully prototyped and tested, the customer evaluates the product design. In case the customer rejects the design, the project is terminated, and majority of the development costs are borne by the semiconductor company. This is known as a loss of design win and is a major source of uncertainty for the semiconductor manufacturer.

3. **Production Ramp:** When the customer accepts the product design, the production phase starts. During production, market and capacity-related risks are crucial to consider. Market risks are influenced by both external factors, such as the nature of market competition, and internal factors, such as the product quality. Capacity risks arise from the uncertain nature of available resources. These two market risk factors collectively affect the expected revenue streams for the product.

The project life cycle depicted in Figure 1 is similar to projects that follow a stage-gate project funnel (Griffin 1997), so while we have tailored our tool to support decision making in the semiconductor industry, its general applicability could be much broader.

Figure 1 Project Life Cycle



Decisions on a particular project proposal must be made at the beginning of the project life cycle, in anticipation of the uncertainties. While uncertainties in the project life cycle represent a major challenge to the portfolio decisions, project relationships and dependencies also have to be taken into account. The relationship between projects might be in different forms. For example, a new project proposal might make a set of existing projects obsolete, or competing customers might offer similar proposals. We refer to these as mutually exclusive projects. Projects may also have prerequisite relations. Portfolio decisions must be made considering all project dependencies, leading to complex trade-offs. Moreover, uncertainties associated with a project are often amplified when the project is considered together with its prerequisite projects. This leads to the conclusion that the value and its variability *as a whole* must be considered by management. In the DSS designed for the company, the scenario analysis tool takes into account all inter-project relationships, as well as possible outcomes of uncertainty that create the risks.

The remainder of the paper is organized as follows. We summarize relevant literature for project/product portfolio management, followed by the problem statement and model formulation. We then introduce the three-stage DSS for project portfolio management, and we describe the implementation and use of this DSS at a particular semiconductor firm. We discuss the robustness of the tool and key feedbacks from the management. The paper finishes with conclusions and directions for further research.

3. Related Literature

Methods for Product Portfolio Management (PPM) can be both qualitative and quantitative. The success of each method depends on the business environment, the information requirements, the users' understanding of the techniques, and the buy-in from senior management. For background on PPM methods, the reader is referred to Cooper et al. (1998).

In the last decade, the use of Decision Support Systems (DSS) has become popular in PPM. DSS are interactive, computer-based systems that help decision makers in their decision process, typically combining multiple components (both qualitative and quantitative) into an integrated system. For example, Greiner et al. (2003) developed a system for screening weapons systems that first develops qualitative project rankings. The priority rankings of the Analytic Hierarchy Process (AHP) are used to represent a measure of value of each project, and a 0-1 integer model is used to construct a "good" portfolio of projects that obeys a budget constraint. Stummer and Heidenberger (2003) use qualitative scoring techniques to rank projects and identify the projects that are worthy of further evaluation. Next, a multi-objective integer linear program is solved to find Pareto-efficient portfolios. Dickinson et al. (2001) use qualitative judgement to define a dependency matrix quantifying the interdependencies between projects. A nonlinear integer program maximizing the Net Present Value (NPV) of the portfolio is solved, subject to budget and strategic alignment constraints. Our proposed DSS also combines both qualitative and quantitative aspects. Qualitative judgments are used to help model uncertainty throughout the project's life cycle. A multi-stage stochastic program (SP) is solved that selects the best portfolio constrained by the cost of resource level adjustments, a risk measure, project dependencies, and strategic alignment criteria. The SP allows for scenario-dependent solutions via recourse actions.

The idea to generate path-dependent valuation of a portfolio's value is similar to Anderson and Joglekar (2005)'s planning framework. In their work, recourse actions are suggested to hierarchically manage the product development process considering contingent events in project life cycle, including uncertainties related to market, creative, technological and process dimensions. Our methodology can be evaluated as a specific implementation of concepts suggested in Anderson and Joglekar (2005) into a complete, integrated tool for the semiconductor industry.

An important feature of the SP developed for the semiconductor industry is that resource levels must be increased/decreased in conjunction with the go/no-go decisions for the projects. Loch and Kavadias (2002) also model resource allocation among product lines in a multi-period setting. They explicitly include carry on benefits of the previous period's investment in the current period and derive closed form solutions for the allocation policy under the cases where product lines have increasing and decreasing returns. The solutions are extended for the cases where market interaction exists and managers are risk-averse. In our case the dependency structure is too complex to directly apply these results. For example, a project's value depends on prerequisite projects IP development success, design win, market structure, customer status, product quality, resource availability and technological status. Hence, the DSS we propose resorts to simulation to compute the numeric valuation of projects in this business setting.

The DSS we propose will measure (and mitigate) risk using a mean-Gini approach, a concept introduced by Shalit and Yitzhaki (1984) in the context of equity portfolio construction. Ringuest et al. (2000) were the first to extend the use of mean-Gini risk to R&D portfolios.

As pointed out by Ringuest and Graves (2005), the estimate of a portfolio Gini considering all available projects can be conservative, overestimating the risk of the selected portfolio. Ringuest and Graves (2005) go on to extend the methodology of Ringuest et al. (2000) with a branch-and-bound method that determines a portfolio using a more precise estimate of the Gini risk. Ringuest et al. (2004) use mean Gini analysis to construct non dominated portfolios and efficient frontier using these nondominated portfolios. A significant distinction of our method from that of Ringuest et al. (2000) and Ringuest and Graves (2005) is that we explicitly model the scenario-based project valuation throughout its life cycle, so our DSS generates path-dependent solutions taking into account the Gini risk in a multi-period horizon. Next we describe the our generic model for the project portfolio selection problem.

4. Model Formulation

4.1. Justification

Prior to our study, the management had employed two different techniques to aid their PPM decisions. The first approach was to rank the projects based on NPV of the cash flows. Heuristic strategic alignment criteria and a budget constraint limited the allowable projects. This methodology was able to select projects with strategic considerations, but it was not capable of assessing and incorporating quantitative measures of risk into the decision process, nor was the method able to accurately model complex inter-project dependencies. In their second approach, the management worked on a mean-variance model. The interdependencies between projects were quantified using several factors. A disadvantage of this approach was that the required data was not captured in the company's routine business transaction data. Moreover the mean-variance model was not dynamic enough to account for the fast rate of change of the semiconductor industry.

Based on the decision makers' needs and past experience, the following design requirements for our PPM decision support tool were put in place:

1. use data available from routine business transactions,
2. dynamically account for changes in the business environment, and
3. incorporate a risk measure for projects in the portfolio.

Considering these requirements, a decision was made to use a Stochastic Programming(SP) approach to the problem. First, the SP model can easily be built from scenario data generated from routine business transactions. Second, in SP models, decision makers are able to change their decisions as they learn new information in a dynamic environment. Last, a parameterized risk constraint can easily be incorporated into the model, allowing managers to obtain different solutions depending on their view of risk. In our view, the solution to a stochastic program is neither a number nor a policy, but a distribution for an important measure or measures of interest. In this case, the main purpose of adding the risk constraint to the SP model was to allow managers to tune, or shape, the distribution of expected returns based on their initial investment decisions.

4.2. Data, Uncertainty and Risk

As part of the renewed project management effort, the company keeps basic financial figures, customer information, and technical product information in their database. At any point in time, the company is considering a set of projects P that will consume resources from a set R. Each project $p \in P$ is associated with a set of prerequisite projects $Q_p \subset P$ and with a set of mutually exclusive projects $E_p \subset P$. Forecast cost and resource information is kept at a project level for quarterly intervals. For each project $p \in P$ the following data is available for each period t in a specified time horizon $\mathcal{T} = \{1, 2, \ldots, T\}$:

- the forecast gross margin of the project p at period t: GM_{pt},
- the forecast fixed costs of the project p in period t: FC_{pt},
- the forecast usage of resource r in period t by project p: W_{rpt}, and
- the periodic unit cost of each resource r: c_r.

As new information is available, forecast figures are updated. From the analysis of the project life cycle and through discussions with management, we identified that a new business state can be modeled by adjusting the gross margins, fixed costs, and resource usage $(GM_{pt}, FC_{pt}\ W_{rpt})$ based on random events that occur during the project's life cycle. That is, these quantities are functions of some random variables. To fix notation, let $\Omega_t, t \in \mathcal{T}$ be the set of all possible random events that occur during period t. The set of all sequences of events is then $\Omega \stackrel{\text{def}}{=} \Omega_1 \times \Omega_2 \times \ldots \times \Omega_T$. We use the common notation $\Omega_{[1,t]}$ to denote the set of all sequences of events that can occur from stages 1 to t. The dependence of gross margin, fixed cost, and resource usage on random variables is denoted by referring to these quantities as $GM_{pt}(\omega_{[1,t]}), FC_{pt}(\omega_{[1,t]})$, and $W_{rpt}(\omega_{[1,t]})$, where $\omega_{[1,t]} \in \Omega_{[1,t]}$. We discuss in Section 5.3 the exact manner in which the quantities are derived, but for purpose of model discussion it suffices to know that for each project and resource we have all the necessary information to assess project's value and determine resource requirements under different sequences of realizations of random events.

The model also includes a mechanism for controlling the company's risk in choosing collections of projects to include in the portfolio. The mechanism is based on calculating "risk coefficients" for each project in each stage. The approach is not unlike assessing an equity portfolio's riskiness by its β coefficient. In our model, the risk coefficients are based on mean-Gini considerations. The description and exact calculation of Gini coefficients is detailed in Section 5.3.1. For purposes of the model discussion, it suffices to know that for each project and time period, a coefficient λ_{pt} is computed that quantifies the risk of project p in time period t. The total risk of the portfolio is taken to be a linear combination of the individual risk coefficients chosen to be included in the portfolio. As pointed by Ringuest and Graves (2005) this approach tends to overestimate the risk of the chosen portfolio, and they suggest a branch-and-bound approach that accurately measures the Gini-risk during the portfolio selection algorithm. We decided not to apply their branch-and-bound approach for the proposed DSS for a number of reasons. First, as described in Section 5, the DSS solves many instances of the problem at varying risk levels. The increased computational burden of a branch-and-bound approach is not attractive in this regard. Further, in the instances we examined, the difference between the approximate Gini risk (as measured by the Gini coefficients λ_{pt}) and the exact Gini-risk was typically quite small. In Section 6, we demonstrate that the exact Gini measure differs from the approximate Gini measure of the final selected portfolio by less than 4% on average for the case-study instance. Finally, a customized branch-and-bound implementation was not desirable to the company, who wished to exclusively use off-the-shelf commercial software in the DSS.

It is important to note that the main purpose of the risk constraint was not to accurately measure the Gini-risk of a project portfolio, but rather to give decision makers the opportunity to adjust the distributions of key output random variables (e.g. net profit) via a parameter in the model.

4.3. Decision Variables

In stochastic programs, decisions are made in stages, and in-between stages information about the state of the business becomes available. To ease the exposition, we will assume that there is one stage in the stochastic programming model for each period in the decision making process. In the actual decision support system described in Section 5, several decision making periods might be aggregated into one stage. There are two classes of decisions that the company must make: strategic and operational.

The strategic decisions are modeled with variables $x_{pt}(\omega_{[1,t]})$, indicating whether or not project p is to be included in the portfolio in time period t. An important characteristic of the model is that $x_{pt}(\cdot)$ is solely a function of the random variables that occur from periods 1 to t. The decision of whether or not to include p in the portfolio at t is non-anticipative of random events that occur after period t. At this point, we model the nonanticipativity implicitly by simply stating $x_{pt}(\cdot)$ is a function only of $\omega_{[1,t]}$. Algorithmic mechanisms for enforcing nonanticipativity will be described in Section 4.4. Associated with each project $p \in P$ is a begin time $b_p \in \{1, 2, \ldots, T\}$, which indicates the beginning time of project's life cycle. If a project is not selected at the beginning of its life cycle, then it cannot be selected later. However, if a project is selected at the beginning of its life cycle, it may be discontinued later.

Completion of a project requires resources to be allocated to the project. The main resources necessary in the model for the company are human resources. As such, there is a cost associated with increasing the level of resources and with decreasing the level of resources from period to period. There are two types of human resources that affect project completion: design team members and administrative staff. The initial available level of resource r is defined by I_r. There are three classes of operational decisions, each dependent on the random events that occur during the PPM process. The total available level of resource r at the end of stage t under scenario $\omega_{[1,t]}$ is modeled with the decision variable $y_{rt}(\omega_{[1,t]})$. Increase and decrease in the level of resource r at the end of stage t under scenario $\omega_{[1,t]}$ are modeled with the decision variables $h_{rt}(\omega_{[1,t]})$ and $f_{rt}(\omega_{[1,t]})$, respectively. The periodic cost of keeping one unit of resource r is given by the parameter c_r. There is a cost associated with the unit increase or decrease of resource level r, given by the parameters η_{rt} and ζ_{rt}, respectively.

4.4. Model Objective and Constraints

The objective is to maximize the total expected operating income,

$$\mathbb{E}_\omega \left[\sum_{t \in \mathcal{T}} \sum_{p \in P} (GM_{pt}(\omega_{[1,t]}) - FC_{pt}(\omega_{[1,t]})) x_{pt}(\omega_{[1,t]}) - \sum_{r \in R} (c_r y_{rt}(\omega_{[1,t]}) + \eta_{rt} h_{rt}(\omega_{[1,t]}) + \zeta_{rt} f_{rt}(\omega_{[1,t]})) \right] \tag{1}$$

The first summation of Equation (1) represents the total net profit associated with doing projects, and the second summation accounts for the operating expenses and adjusting the resource levels.

Most constraints of the model contain random variables, and we enforce these constraints in a probabilistic sense by saying that the constraints will hold with probability one, or almost surely, represented by the notation a.s. in the constraints of the model. The required level of each type of resource should be enough to continue selected projects at all stages and can be described by the following constraint:

$$\sum_{p\in P} W_{rpt}(\omega_{[1,t]})x_{pt}(\omega_{[1,t]}) \leq y_{rt}(\omega_{[1,t]}) \qquad \forall r\in R\ t\in \mathcal{T}, \text{a.s..} \tag{2}$$

The headcount level for a type of resource at the end of a specific stage is the sum of headcount level in the previous stage and the net adjustment of the level of resources at the current stage. For the first stage, headcount level is adjusted over the initial number of headcount level, I_r,

$$y_{r1} = I_r + h_{r1} - f_{r1} \qquad \forall r\in R, \tag{3}$$

$$y_{rt}(\omega_{[1,t]}) = y_{r,t-1}(\omega_{[1,t]}) + h_{rt}(\omega_{[1,t]}) - f_{rt}(\omega_{[1,t]}) \qquad \forall r\in R,\ t\in \mathcal{T}\setminus\{1\}, \text{a.s..} \tag{4}$$

Project p can be undertaken if and only if all projects in its prerequisite set Q_p are undertaken at stage t, a requirement enforced by Equation (5). Similarly, Equation (6) ensure that is project p is in the portfolio at stage t, then none of the projects in mutually exclusive set E_p are also in the portfolio.

$$x_{pt}(\omega_{[1,t]}) - x_{qt}(\omega_{[1,t]}) \leq 0 \qquad \forall p\in P,\ q\in Q_p,\ t\in\mathcal{T} \text{ a.s.,} \tag{5}$$

$$x_{pt}(\omega_{[1,t]}) + x_{lt}(\omega_{[1,t]}) \leq 1 \qquad \forall p\in P,\ l\in E_p,\ t\in\mathcal{T} \text{ a.s..} \tag{6}$$

We can only select a project p to perform when the stage of the decision coincides with the beginning period of that project's life cycle b_p. If a product is not selected at the beginning of its life cycle, then it cannot be selected later. We ensure these restrictions by the following constraints:

$$x_{pt}(\omega_{[1,t]}) = 0 \qquad \forall p\in P,\ \forall t\in\{1,\dots b_p-1\},\ \text{a.s.,} \tag{7}$$

$$x_{p,t+1}(\omega_{[1,t+1]}) - x_{pt}(\omega_{[1,t]}) \leq 0 \qquad \forall p\in P, \forall t\in\{1,2,..,T-1\},\ \text{a.s..} \tag{8}$$

As explained in Section 4.2, there is a risk coefficient λ_{pt} for each product p at each stage $t, (t\in\mathcal{T}\setminus\{1\})$. Total risk is limited to a level K in the following fashion:

$$\sum_{t\in\mathcal{T}\setminus\{1\}}\sum_{p\in P}\lambda_{pt}x_{pt}(\omega_{[1,t]}) \leq K \text{ a.s..} \tag{9}$$

Note that it is a simple matter to change constraint(9) to limit risk on a per-period basis. However, keeping the risk level parameterized by one constant (K) was desirable so that management could easily assess the impact of varying risk on the solution output.

Again, an important consideration in the model is that the strategic and operational decisions made at a stage t are <u>independent</u> of the random events $\omega_{[t+1,T]}$. These nonanticipativity constraints are given implicitly here by defining the $x_{pt}(\cdot)$ as functions of the proper arguments ($\omega_{[1,t]}$). Algorithmic methods for enforcing this nonanticipativity will be discussed in the next section.

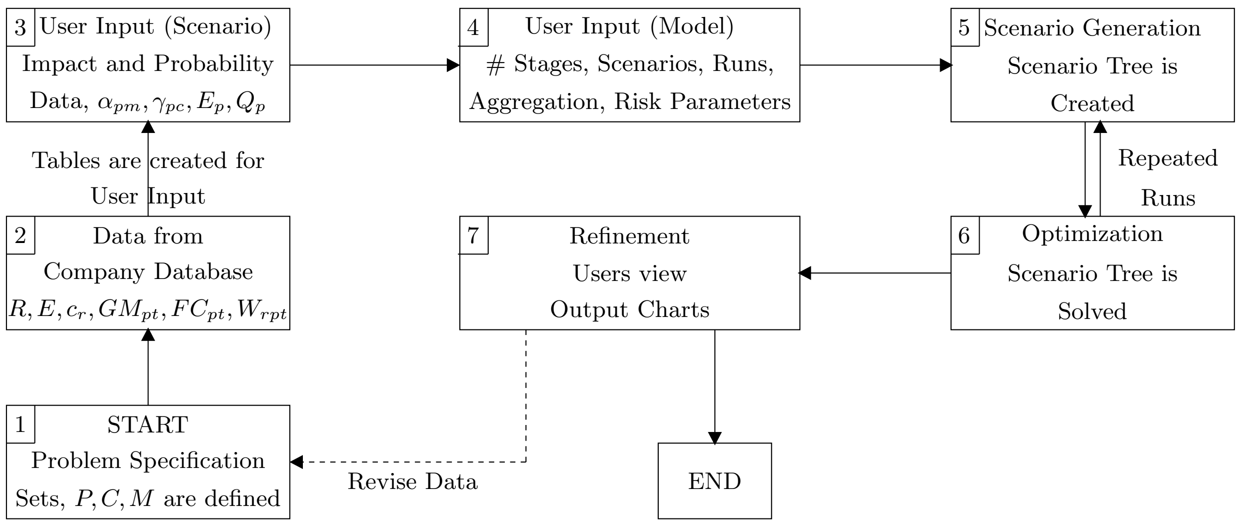


Figure 2 Event Flow Diagram of the DSS

To make the optimization model (1-9) tractable, there must be some reasonable approximation of the uncertainty Ω. We chose a sample-path, or sample-average approach, in which set Ω is replaced by a randomly sampled approximation consisting of a finite number of scenarios. This approximation, or scenario generation, forms the first phase of our three phase DSS that we describe in the Section 5. The sample average approach has been used as of late on a variety of practical planning problems, including supply chain planning by Santoso et al. (2005) and vehicle routing by Verweij et al. (2003). Further, recent theoretical and empirical evidence suggests that an accurate answer to the true problem can be obtained by approximating the uncertainty set with a surprisingly few number of scenarios (Shapiro and Homem-de-Mello 2000, Linderoth et al. 2006). The use of scenarios also enables us to enforce the nonanticipativity of decisions by only creating decision variables that can depend on the appropriate scenarios and previous decisions.

5. A Three-Phase Decision Support System

Rarely is a real-life problem simply "solved" by applying an optimization model. Rather, the road from real-life problem to working solution is an iterative process. The model we propose in Section 4 is no exception; it is part of a larger decision support system (DSS) now in place at the company. Figure 2 depicts the various components and stages of the DSS, as the user is guided through the project portfolio planning process. Note that in order to instantiate an instance, user input is required both before and after data is drawn from the company's database. In this section, we describe the components of the DSS in greater detail.

5.1. Problem Specification

Initially, decision makers specify the planning horizon (in fiscal quarters), the set of projects to be analyzed (P), the customer set (C), and the set of target markets (M) for these candidate projects. One individual project may be split over multiple customers and market segments. The decision maker defines the exposure of project p to its markets m as a fraction α_{pm} and the exposure of each project to its customers c as a fraction γ_{pc}.

Once the initial instance specification data is obtained from the users, the DSS queries the company database to retrieve the resources (R) consumed by the projects in set P, the periodic unit cost of each resource (c_r), the set of technologies used in the product development phase (E), and the forecast gross margins (GM_{pt}), fixed costs (FC_{pt}), and resource usage (W_{rpt}) for the specified projects, resources and periods. Further, all important time-line information about the project life cycle, such as design-win date, are obtained.

This information is used to help describe a scenario, an instance of all possible sequences of random events relating to projects, resources, technologies, market segments and customers from the beginning until the end of the planning horizon. The set of all possible scenarios is denoted as $\Omega_{[1,T]}$. From analysis of the project life cycle and feedback from project managers, we concluded that seven Risk Groups were sufficient to accurately describe $\Omega_{[1,T]}$. The risk groups relate to (1)Product Performance, (2)Resource Performance, (3)Technology Status, (4)Intellectual Property (IP) Development Status, (5) Design-Win Status, (6) Market Performance, and (7) Customer Performance.

The risk groups and possible outcomes of events in each risk group are the same regardless of the instance being considered. What varies from instance to instance is the impact associated with each outcome. Each outcome is associated with an impact parameter (ρ) that will be used to adjust forecast data.

The outcome of events in one risk group may be correlated through time or correlated with outcomes of events from different risk group. These event relationships, the definitions of the risk groups, and the impact factors for each group is described in Section 5.2.

5.2. User Input

A key characteristic of the DSS in place at the company is that the project managers and high-level executives are responsible for quantifying the impact of various random events that may occur during the planning horizon. This domain expertise is spread among various individuals in the company.

Table 1 Risk Groups and Possible Outcomes

RISK GROUPS	OUTCOMES				IMPACT
Product Performance	Superior Performance	Expected Nominal	Poor Performance	Product Failure	$\rho_{pt}^{PP_{gm}}, \rho_{pt}^{PP_{fc}}$
Resource Performance	Over Performance	Expected Nominal	Under Performance	Loss of Key Resources	ρ_{rt}^{R}
Technology Status	Development Goes Well	Expected Nominal	Behind Schedule	Failure	ρ_{et}^{E}
IP Development Status	Successfully completed		Were not completed		Late Time to Market or continue
Design-Win Status	Got the design-win		Did not get the design-win		Zero GM_{pt}'s or continue
Market Performance	Market Expands	Expected Nominal	Market Contracts	Market Collapses	ρ_{mt}^{M}
Customer Performance	Superior Performance	Expected Nominal	Poor Performance		ρ_{ct}^{C}

While the impact coefficients and associated probabilities certainly may seem arbitrary, it is important to note that they are used in a systematic manner and obtained from people within the company who are best-poised to provide such information. The DSS is a leap forward for the decision makers in the company who routinely did "what-if" analysis of various scenarios, but lacked the machinery necessary to properly capture the correlations between various events and to act intelligently in the face of uncertainty. The new DSS overcomes both of these obstacles: a simulation is used to ensure that all events are correlated through time, and an optimization model is used to simultaneously consider many different uncertain scenarios to choose the best action.

Typically, to create an instance, strategy meetings among executives and individuals with domain knowledge are held, and impact factors and probabilities are assigned for each event in each risk group.

The risk groups and their probable outcomes are given in Table 1 and may be summarized as follows:

- Product Performance: Events related to the quality of a product or prototype. The impact of the product performance on the gross margin GM_{pt} is denoted by $\rho_{pt}^{PP_{gm}}$ and its impact on fixed cost FC_{pt} is denoted by $\rho_{pt}^{PP_{fc}}$. The outcomes of product performance can affect the likelihood of the "Design-Win" event if the project is in the development phase.
- Resource Performance: Events related to the performance of R&D and administrative staff. There is one outcome for each resource $r \in R$ and for each stage $t \in \mathcal{T}$. The outcome has an impact ρ_{rt}^{R} on the forecast resource usage rate W_{rpt}. If resources are lost in a stage, then projects using that resource cannot be completed unless new resources are obtained.

- Technology Status: Events related to the timing of the technology and project schedule. Each project is built using one type of technology. The technology impact ρ^E_{et} is used to update the probabilities of IP development status events.
- IP Development Status: A binary and one time event indicating the IP development status for project p. Failure of the IP development delays the time to market of the project and hence all other projects for which p is a prerequisite. The delay of the market release time is handled by shifting the gross margins GM_{pt} further out in the planning horizon. The number of stages to shift the release and any increase in development cost FC_{pt} is specified by the users. The probability of successfully completing the IP development at time t, π_{IPt}, depends on the outcome of the previous period's technology status.
- Design-Win Status: A binary and one time event indicating if the customers accept the prototype design. If the customer accepts the design, the production phase starts and the product is released to the market. If the customer does not accept the design, the project is not released to the market. However, in the case of customer rejection, the dependent projects can still benefit from the project's IP (if the IP was successfully developed). The probability of a design-win at time t, π_{DWt} depends on the previous period's product performance outcome.
- Market Performance: Events related to market condition. A product might serve more than one market. In this case, the percentage exposure of product p to market m, α_{pm}, is used to obtain the overall effect of the market related outcomes. The outcomes have an impact ρ^M_{mt} on the gross margins GM_{pt}.
- Customer Performance: Events related to customer performance. A product might serve more than one customer. In this case, the percentage exposure of product p to customer c, γ_{pc}, is used to obtain the overall effect of the customer related outcomes. The customer performance outcomes have impact ρ^C_{ct} on the product gross margins GM_{pt}.

After the probability and impact data is collected, the product-market (α_{pm}) and product-customer (γ_{pc}) exposures are provided by the decision makers. Last, project dependencies are specified in terms of prerequisite set Q_p and mutually exclusive set E_p for each project p. These sets are taken from the database, but at this point the user may override the information.

Data collection for large-scale problems is always a burden. However, the user-friendly interface of the DSS greatly eases this overhead. The DSS is connected to the company database, so the forecast data for the project, technology, and resource sets over the planning horizon are obtained instantly. The DSS generates tables to gather probability and impact data, and when the data is collected, error-checking routines ensure that all the data are present and fall within reasonable nominal ranges.

Even though the space of outcomes has been discretized, so that $\Omega_{[1,T]}$ is a finite set, the cardinality of this set is too large for us to consider all possible combinations of outcomes. To create a tractable instance of our optimization model, we sample from the set $\Omega_{[1,T]}$, and an instance of the model is created when the user specify the number of stages of the stochastic programming model, the number of random event realizations at each stage, and information on how the fiscal quarters (periods) are spanned by stages (Figure 2, Step 4). If aggregation is needed, the system does this by summing GM_{pt}, FC_{pt}, and W_{rpt} over the periods in that stage. In this step, the users also specify the number of sampled instances to be created and solved and the number of different risk levels for which each sampled instance should be solved. That is, each SP is sampled multiple times, and each sampled instance of the SP is solved at varying risk levels K. After all the above information is acquired, the DSS checks whether all the data required for scenario generation is complete and logical. If the information is error free, then the DSS proceeds to the Scenario Generation Phase.

5.3. Scenario Generation

Conditional sampling is used to create a manageable sample of the discretized scenario space $\Omega_{[1,T]}$. The size of the scenario tree depends on the number of stages (T) and the number of random event realizations at each stage t (M_t). The conditional sampling procedure works by first selecting a random sample of size M_2 for the second stage.

The probability of each realization in the second stage is $1/M_2$. The second stage realizations are given by

$$\zeta_2^i = \left(GM_{p2}(\omega_2^i), FC_{p2}(\omega_2^i), W_{rp2}(\omega_2^i) \right), \ i = 1, ..., M_2.$$

Formulae for dependence of gross margin, fixed cost, and resource consumption ($GM_{pt}(\cdot), FC_{pt}(\cdot)$, and $W_{rpt}(\cdot)$ on the outcome of the random event ω^i are given subsequently in Equations (10)-(12).

Next, for every $i \in \{1, ..., M_2\}$ a random sample of size M_3 is generated. Thus, there are M_2M_3 realizations of events in the third stage, each with probability $1/(M_2M_3)$. The third stage realizations given the history of events up to this point $\omega_{[1,2]}$ are given by

$$\zeta_3^{ij} = \left(GM_{p3}(\omega_3^j|\omega_{[1,2]}^i), FC_{p3}(\omega_3^j|\omega_{[1,2]}^i), W_{p3}(\omega_3^j|\omega_{[1,2]}^i) \right), \ i = 1, ..., M_2, \ j = 1, ..., M_3.$$

The procedure is called conditional sampling due to the dependence of ζ_3^{ij} on ζ_2^i.

The procedure continues in this fashion until the T^{th} stage is reached, creating a history of events $\omega_{[1,T-1]}$. At the T^{th} stage, we have $M_2M_3\cdots M_T$ realizations of events, each with equal probability $1/(M_2M_3\cdots M_T)$. Each path from root node to leaf node in this tree is a scenario.

Conditional sampling is required to capture the correlations between random events. The events may be correlated through time and within risk groups. The relation between the outcomes is shown in Figure 3. The figure depicts that the outcome of the product performance event in period t will affect the outcome of the design-win status event in period $t+1$. The figure also shows the dependence of the gross margins, fixed costs, and resource consumption on the outcome of random events. The exact functional dependence is given in Equations (10)-(12).

$$GM_{pt}(\omega_t^i|\omega_{[1,t-1]}) = \rho_{pt}^{PP_{gm}} GM_{pt}\left[\sum_{m\in M}\sum_{c\in C}(\alpha_{pm}\rho_{mt}^M)(\rho_{DW}^C\gamma_{pc}\rho_{ct}^C)\right] \quad \forall p\in P, \forall t\in\{2,3,\dots,T\}, \quad \forall i\in\{1,2,...,M_t\}, \tag{10}$$

$$FC_{pt}(\omega_t^i|\omega_{[1,t-1]}) = \rho_{pt}^{PP_{fc}} FC_{pt} \qquad \forall p\in P,\ \forall t\in\{2,3,\dots,T\},\ \forall i\in\{1,2,...,M_t\}, \tag{11}$$

$$W_{rpt}(\omega_t^i|\omega_{[1,t-1]}) = \rho_{rt}^{R} W_{rpt} \qquad \forall p\in P,\ \forall r\in R, \forall t\in\{2,3,\dots T\},\ \forall i\in\{1,2,...,M_t\}. \tag{12}$$

The outcome of scenario generation is a scenario tree, where each node carries information regarding to the system state at that stage. We will denote the set of nodes that belong to stage t as S_t. From this point forward, we will use node notation to describe gross margins, fixed costs and required headcount levels. That is, at stage t, GM_{pn}, FC_{pn} and W_{rpn} will be the updated gross margins, fixed costs and required headcount levels at node n for each $n\in S_t$.

5.3.1. Risk Measure Once the scenario tree is created, the Gini coefficients λ_{pt} are calculated for each product p ($p\in P$) and stage t ($t\in\mathcal{T}$). Similar to variance, the Gini statistic is a measure of dispersion and is defined as the expected value of the half the absolute difference between every pair of realizations of the random variable r having cumulative distribution function (cdf), $F(\cdot)$. The Gini statistic can be mathematically written as $\mathbb{E}|r_1-r_2|/2$ where r_1 and r_2 are independent random draws from the distribution. Shalit and Yitzhaki (1989) demonstrate that the Gini statistic of a random variable r can also be written as twice the covariance of random variable r and its cdf, $F(\cdot) : 2\,\mathrm{cov}(r, F(r))$, where $F(\cdot)$ is uniformly distributed over [0,1].

To compute the Gini coefficients λ_{pt} in the context of our optimization model, the return of each available product for each realization of the uncertainty in that stage (θ_{pn}) is calculated

$$\theta_{pn} \stackrel{\text{def}}{=} GM_{pn} - FC_{pn},\ \forall p\in P_t,\ \forall n\in S_t, \tag{13}$$

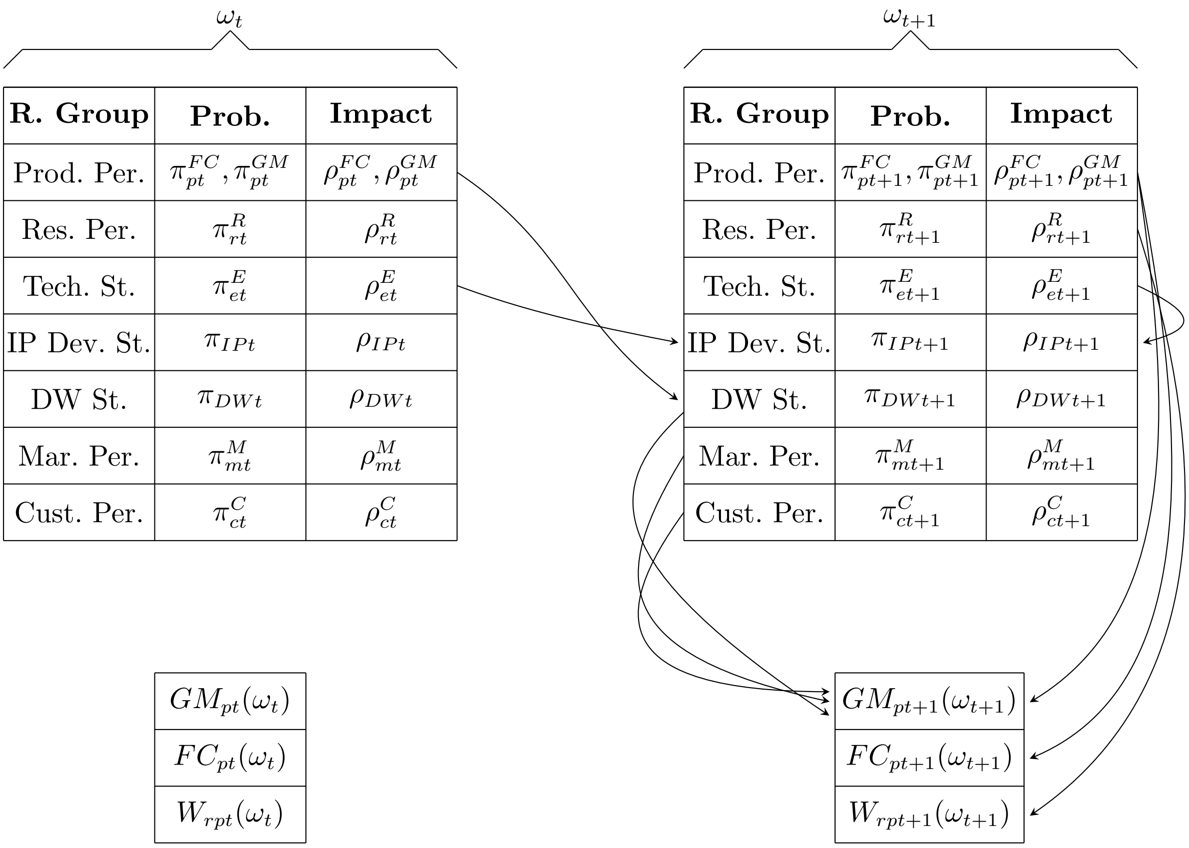


Figure 3 Outcome Relations

and the total return of all products for that realization (θ_n) is computed

$$\theta_n \stackrel{\text{def}}{=} \sum_{p \in P_t} \theta_{pn} \ \forall p \in P_t, \ \forall n \in S_t. \tag{14}$$

In Equations (13) and (14), the symbol $P_t \subset P$ is the set of all products that can be selected in stage t (i.e. $p \in P_t \Rightarrow b_p \leq t$). Next, the values θ_{pn} and θ_n are assembled into collections of length $|S_t|$ and the vector ϕ_{pt} is constructed from θ_{pn} by applying the same transformation made on vector θ_n to sort it in ascending order hence obtain ϕ_t. The Gini coefficient is then computed as

$$\lambda_{pt} = 2\,\mathrm{Cov}(\phi_{pt}, F_t(\phi_t)) \quad \forall t \in \mathcal{T} \setminus \{1\}, \ \forall p \in P_t, \tag{15}$$

where $F_t(\phi_t)$ is the cumulative probability of the total return ϕ_t. Since returns are generated via a sampling procedure, and the probability of each realization is equal, $F_t(\phi_t)$ is found by dividing the rank of each element in ϕ_t by the cardinality of set S_t. A deeper explanation of the meaning and calculation of the Gini coefficients can be found in Yitzhaki (1998).

5.3.2. Complexity of Scenario Generation Algorithm Do we need this?

The time required to generate a single scenario tree depends on the number of stages, T, number of events at each stages, $(M_t\ t \in \mathcal{T} \setminus \{1\})$, total number of nodes, N, and the cardinality of project, customer, market, resource and technology sets. Assuming the size of latter is small and one has on average same number of events at each stage,M, the time complexity is $O(M^{T-1})$. As seen obviously, the choice of T and M could dramatically effect the time required to generate one scenario tree. For example $T = 2$ and $M = 10^5$ is equivalent to $T = 6$ and $M = 10$ The choice of M and T should be made considering the problem context and the solution time of the optimization model in described in next section.

5.4. Optimization

Once the scenario tree is created, we can build a tractable approximate optimization model. The solution of the model produces a "best" portfolio and required headcount levels for each resource. Decision variables are the counterparts of the decision variables described in Section 4.3. The variables x_{pn} indicate whether or not project p is included in the portfolio at node n. The variables y_{rn}, h_{rn}, f_{rn} respectively indicate the number of available resource level, the increase, and decrease in resource levels for each resource r at node n. For any node n of the scenario tree, π_n is the path probability of node n, or probability that the sequence of events leading to node n will occur. The sampling methodology employed implies that if $n \in S_t$, then $\pi_n = 1/(M_1 \cdots M_t)$. With these definitions, the product portfolio management problem can be modeled as follows:

$$\text{Maximize:} \sum_{t\in\mathcal{T}} \sum_{n\in S_t} \pi_n \left[\sum_{p\in P} (GM_{pn} - FC_{pn}) x_{pn} - \sum_{r\in R} (\eta_r h_{rn} + \zeta_r f_{rn} + c_r y_{rn}) \right] \quad \textbf{(1)} \tag{16}$$

subject to

$$\sum_{p\in P} W_{rpn} x_{pn} \le y_{rn} \qquad \forall r \in R,\ \forall t \in \mathcal{T},\ \forall n \in S_t, \quad \textbf{(2)} \tag{17}$$

$$y_{r1} = I_r + h_{r1} - f_{r1} \qquad \forall r \in R, \quad \textbf{(3)} \tag{18}$$

$$y_{rn} - y_{r\rho(n)} = h_{rn} - f_{rn} \qquad \forall r \in R,\ \forall t \in \mathcal{T} \setminus \{1\},\ \forall n \in S_t, \quad \textbf{(4)} \tag{19}$$

$$x_{pn} \le x_{qn} \qquad \forall p \in P,\ \forall q \in Q_p,\ \forall t \in \mathcal{T},\ \forall n \in S_t, \quad \textbf{(5)} \tag{20}$$

$$x_{pn} + x_{un} \le 1 \qquad \forall p \in P,\ \forall u \in E_p,\ \forall t \in \mathcal{T},\ \forall n \in S_t, \quad \textbf{(6)} \tag{21}$$

$$x_{pn} = 0 \qquad \forall p \in P,\ \forall t \in \{1, 2, ..., b_p - 1\},\ \forall n \in S_t, \quad \textbf{(7)} \tag{22}$$

$$x_{p\rho(n)} \ge x_{pn} \qquad \forall p \in P,\ \forall t \in \mathcal{T} \setminus \{1\},\ \forall n \in S_t, \quad \textbf{(8)} \tag{23}$$

$$\sum_{t\in\mathcal{T}\setminus\{1\}} \sum_{p\in P} \lambda_{pt} \sum_{n\in S_t} \pi_n x_{pn} \le K, \quad \textbf{(9)} \tag{24}$$

$$x_{pn} = 1 \qquad \forall p \in F_1,\ \forall t \in \mathcal{T},\ \forall n \in S_t, \tag{25}$$

$$x_{pn} = 0 \qquad \forall p \in F_0,\ \forall t \in \mathcal{T},\ \forall n \in S_t. \tag{26}$$

The bold equations numbers in parenthesis represent the counterpart equations in the model present in Section 4.4. Additionally, we include Equations (25) and (26)) to include or exclude a specific project or set of projects for the whole planning horizon. The sets F_1 and F_0 are given by the user, should he or she wish to specify a partial solution.

All of the steps in Figure 2, up to the creation and solution of the optimization model are done using Excel VBA. The optimization models are created by the AMPL modeling language (Fourer et al. 2002), reading and writing data from and to Excel via ODprimary market. The optimization instances are solved using CPLEX v9.1. The DSS is responsible for ensuring that the entire batch of optimization instances (coming from different sampled scenario trees solved at varying risk levels) is solved, as depicted in Figure 2, Steps 5 and 6. When all the runs are complete, the DSS proceeds with the Refinement Phase.

Do we need this?

As mentioned in Section 5.3.2 time required to reach the Refinement Phase depends on the choice of M and T. Although time required to generate one scenario tree grows polynomially (DSS does not allow T more than 6), required time to solve the above problem grows exponentially. For example if one doubles M, it takes 2^{T-1} more time to generate the tree, however, the number of decision variables and the constraints would grow by 2^{T-1}. Hence, the choice of model parameters should be carefully made. The trade off between the quality of the optimal solution and versus time required to solve a single instance should be determined considering problem context.

5.5. Refinement

A key component of the DSS is the feedback given to decision makers about solutions obtained from the optimization model. Once an instance is created and solved, the sampled outcomes of uncertainty and the decision made for each outcome of the uncertainty are written to a database. Next, charts and other visual aides are constructed from the database to help the decision maker see the cause and effect relationships of model inputs. From our experience, the management found two types of charts useful in the decision making process. (Specific examples of each type of chart are given along with the case study presented in Section 6).

1. Efficient Frontier: The optimization problem (16)-(26) is solved for increasing risk levels (K), and the optimal expected portfolio profit is plotted against K. Also, using the results from multiple runs, quantiles of the expected profit are constructed and displayed.

2. Portfolio Composition: Since the multi-stage stochastic program is solved at different risk levels (K) and for different sampled scenario trees, the optimization offers many different suggestions as to what projects should be included in the portfolio. The portfolio composition charts give an indication of how often a project is included in the portfolio for the different realizations of uncertainty. The charts show for each project, the percentage of cases in which the project was included in the portfolio. One type of this chart is created on a per-year basis, in which case the solutions are averaged over all risk levels and over all sampled scenario trees. Another type of chart is created on a per-risk level basis, averaging first stage solutions over different sampled scenario trees. Typically, this chart is created to demonstrate the projects that are done in the initial portfolio.

The three-phase DSS has been tested and validated with real case studies. In the next section, we describe one of these case studies, in which a semiconductor manufacturer's management is on the verge of constructing project portfolio for one of their major business units. Each of the different charts providing solution feedback will be demonstrated via this case study example.

6. Case Study

The company was faced with a strategic portfolio decision charting the course of one of their major divisions over a four year time period. The strategic decision involved creating a portfolio of projects to undertake from a candidate set of size 21. For this particular business decision, inferring reasonable expected project returns from forecast values was turning out to be very difficult for management. The various "information points", such as IP development completion and the design-win date, during a project's life-cycle had a very significant impact on the project returns. Further, there was a significant amount of interdependence (pre-requisite and exclusivity) information for the projects that could be undertaken. For these reasons, this particular strategic portfolio decision seemed like an ideal case study for the DSS in place at the company. The time line, cash flows, and total resource cost of the various candidate projects is depicted in Table 2. The numbers in parenthesis represent negative values.

In this case study, to more accurately represent business decisions that could be feasibly taken in subsequent years, the management decided to put additional projects as "business exits", indicating total (P1) and partial (P2) business exit from this business area. Note the large lists of mutually exclusive projects with these business exit projects. In order to depict the dependency of one project's cash flow on another project's IP development status, an alternative cash flow line is given under the original one. For example, the cash flow for project P7 depends on the IP development status of project P6. If project P6 does not complete IP development on time, then the second line

Table 2 Data from Company Database

					Forecast Fixed Costs and Gross Margins									
					Year 1		Year 2		Year 3		Year 4			
No.	**Proj.**	**IP Depend.**	**Preq. Proj.**	**M. Ex. Proj.**	**FC**	**GM**	**FC**	**GM**	**FC**	**GM**	**FC**	**GM**	$\sum_t \sum_r W_{rpt}$	**Tot. Value**
1	P1	-	-	2,4,7,8 9,10,11	0.00	0.00	0.00	15.97	0.00	0.00	0.00	0.00	(35.30)	(19.33)
2	P2	-	-	1,4,7 8,9,10	0.00	0.00	0.00	2.45	0.00	0.00	0.00	0.00	(14.70)	(12.25)
3	P3	-	-	-	(0.30)	0.00	0.00	0.00	0.00	0.00	0.00	0.00	(14.70)	(15.00)
4	P4	-	3	1,2,5,11	(5.00)	3.50	(3.30)	4.80	0.00	0.00	0.00	0.00	(1.30)	(1.30)
5	P5	-	2,3	4	(5.00)	3.50	(3.30)	4.80	0.00	0.00	0.00	0.00	(1.30)	(1.30)
6	P6	-	3	-	(0.50)	0.00	(0.30)	7.00	0.00	0.00	0.00	0.00	(1.10)	(5.10)
7	P7	P6 Pass	3	1,2,11	(2.80)	(0.10)	(6.30)	14.50	(6.50)	23.00	(0.90)	0.90	(17.90)	3.90
		P6 Fail			(2.80)	(0.10)	(8.80)	12.00	(6.50)	23.00	(0.90)	0.90		(1.10)
8	P8	-	-	1,2,11	(1.00)	0.00	(6.80)	0.60	(11.30)	46.60	(9.30)	44.90	(32.50)	31.20
9	P9	-	8	1,2,11	0.00	0.00	0.00	8.00	0.00	0.00	0.00	0.00	0.00	8.00
10	P10	P8 Pass	-	1,2,11	0.00	0.00	(0.60)	0.00	(5.70)	5.20	(13.10)	60.40	(23.30)	22.90
		P8 Fail			0.00	0.00	(0.60)	0.00	(12.10)	5.20	(24.60)	60.40		5.00
11	P11	-	2	1,4,7 8,9,10	0.00	0.00	(1.10)	0.00	(1.00)	0.00	(1.90)	0.00	(22.40)	(26.40)
12	P12	-	-	-	(1.10)	0.90	(0.20)	0.50	0.00	0.00	0.00	0.00	(1.90)	(1.80)
13	P13	-	12	-	(1.20)	1.00	(0.80)	1.40	(0.40)	0.20	0.00	0.00	(1.50)	(1.30)
14	P14	-	13	-	0.00	0.00	(1.20)	0.30	(0.90)	2.30	(1.10)	1.90	(1.40)	(0.10)
15	P15	-	3,12	-	(0.60)	0.00	(0.19)	3.50	(2.30)	6.30	(0.80)	2.00	(3.20)	4.71
16	P16	-	15	-	(0.30)	0.00	(1.60)	5.90	(1.80)	5.80	(0.60)	0.90	(0.30)	8.00
17	P17	-	11,15	-	0.00	0.00	(2.80)	(0.10)	(4.10)	13.60	(4.80)	24.00	(5.90)	19.90
18	P18	-	16	-	0.00	0.00	(0.10)	0.00	(1.40)	5.80	(3.80)	24.50	(3.00)	22.00)
19	P19	-	3	-	(8.00)	(0.50)	(15.40)	7.30	(18.10)	17.00	(21.60)	46.00	0.00	6.70
20	P20	-	8,19	-	0.00	0.00	0.00	0.00	0.00	0.00	(13.00)	25.00	0.00	12.00
21	P21	-	11	-	0.00	0.00	0.00	0.00	(27.90)	48.90	(31.00)	60.30	0.00	50.30
22	P22	-	11	-	0.00	0.00	0.00	0.00	0.00	0.00	(3.30)	15.20	0.00	11.90
23	P23	-	11	-	0.00	0.00	0.00	0.00	0.00	0.00	(3.62)	18.40	0.00	14.78

Table 3 Resource Related Figures (in $ millions)

	Overhead cost (c_r)	Hiring cost (η_r)	Firing cost (ζ_r)	Initial level (I_r)
R&D	$0.256	$0.200	$0.300	80
SGA	$0.150	$0.100	$0.200	20

of P7 gives its future projected cash flows. A similar relationship holds for projects P8 and P10. For this business area, the IP development status does not affect the time to market. However, in other business areas, where successful IP development does impact time to market, this would by represented by shifting all cash flows on the second line of cash flow data to the right in the figure.

A noteworthy point about Table 2 is that a project's begin time b_p is the first period when a non-zero fixed cost or gross margin is observed for the project, and the project's design-win event should be considered in the period when the first nonzero gross margin is observed for the project. If a project passes the design-win then the forecast gross margins GM_{pt} in the table would be used as a base to calculate the scenario dependent gross margins via Equation (10).

There are two types of resources: R&D and Administrative. Table 3 provides cost and initial count data related to both types of resources in the case study instance.

Management considers three projects vital for the company strategy. Independent of the remaining portfolio selections, P3, P12 and P15 are "must do" projects. These projects are technology development projects, and have very little or zero gross margins expected throughout their life cycle. Thus, $F_1 = \{\text{P3, P12, P15}\}$ in Equation (25).

Before undertaking a careful, quantitative analysis using the DSS in place at the company, the managers had concluded that there were likely two different "business lines" that were attractive options to follow. The general consensus was that the first business line (BL1) was a conservative option, that would yield reasonable returns at low risk levels, and that the second business line (BL2) was a riskier, but potentially more profitable, undertaking. A key portfolio characteristic that distinguishes between the business lines is the commencement of project P11. If P11 is done, this implies that the second business line BL2 is being followed. Management was particularly interested in gaining insight from the DSS in support of one of these business lines, or in learning if there were portfolio options from the set of candidate projects with more desirable risk/return characteristics than either BL1 or BL2.

For the analysis, in addition to the project specific information from the company database shown in Table 2, management defined the key markets (M), key customers (C), and probabilities and impact data for each risk group outcome defined in Table 1. For this instance, there are 4 markets, 5 customers, 3 technologies, 2 resource types and 23 projects (including the 2 business exit projects). The data collection and scenario generation process are handled with Excel's VBA module. For this instance, it was reasonable to consider a four-stage stochastic program, as the company can decide to start or end projects roughly at the beginning of its Fiscal Year.

Even with the relatively coarse discretization of the uncertainty space described in Table 1, there are still on the order of a quadrillion (10^{15}) outcomes for each stage. Sampled optimization instances containing 10,000 scenarios were created by considering $M_2 = 50$ realizations in the second stage, $M_3 = 20$ realizations in the third stage, and $M_4 = 10$ realizations in the final stage. Twenty-five different sampled instances of this size were created, and each instance was solved for twelve different risk levels K, varying from $K = 10$ to $K = 65$ in increments of five units. Solving all 300 instances took roughly 6 1/2 hours, and there was not significant variation in solution time from instance to instance. The instance size of 10,000 scenarios was chosen so that the entire family of 300 instances could be solved overnight on a single PC. While 10,000 scenarios is a very small fraction of the total number of scenarios, our observation has been that both the optimal solution and its value obtained from different sampled instances tends to be very "stable", not changing much from instance to instance. Theoretical evidence showing that there is a high probability of

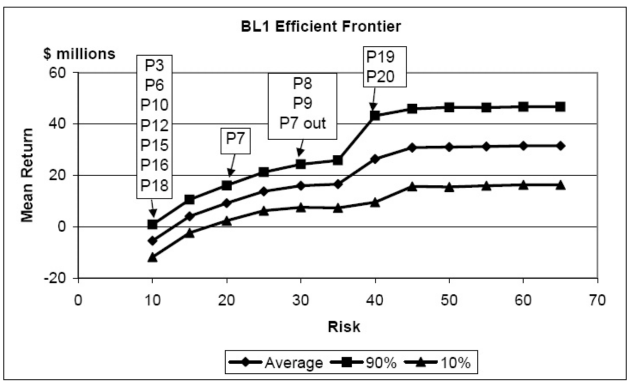
Table 4 Number of BL1 and BL2 instances for varying risk levels

	Risk Level											
	10	15	20	25	30	35	40	45	50	55	60	65
# BL1 instances	25	25	25	20	9	4	10	11	11	13	13	13
# BL2 instances	0	0	0	5	16	21	15	14	14	12	12	12

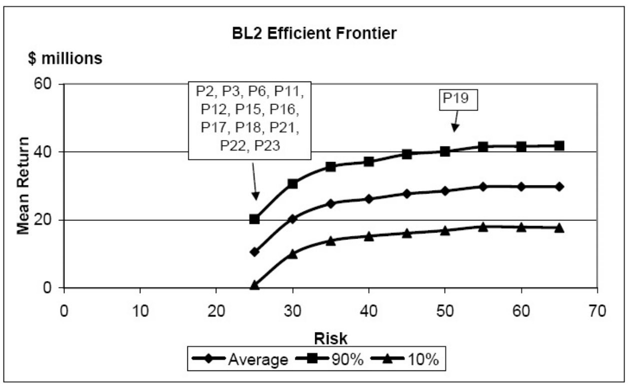
obtaining the true optimal solution from a small sample size in the case of two-stage stochastic recourse problems is given by Shapiro and Homem-de-Mello (2000) and Kleywegt et al. (2001).

When the suite of 300 instances was run, the results indicated that nearly all portfolios from the optimization fell roughly into one of two sets. Further, these two sets corresponded quite accurately to the company's notions of the two business lines BL1 and BL2. In subsequent analysis, each of the 300 optimal solutions was categorized as to belong to one of BL1 or BL2.

Figure 4 shows the efficient frontier plot, described in Section 5.5, for optimal portfolios corresponding to Business Line 1 and Business Line 2. The middle line in each figure is the average expected return for each risk level, and the upper and lower lines are the 90% and 10% (resp.) quantiles of the optimal return. One can note from the figures that at low risk levels, an optimal portfolio always follows BL1, but at higher risk levels, the optimal portfolio sometimes follows BL1 and sometimes follows BL2. Table 4 shows for each risk level how many of the 25 solutions followed BL1 and BL2. In general, as more risk is allowed in the portfolio selection, more projects are added. Figure 4 also displays the first risk level at which a project begins to appear in portfolio solutions. With the aid of Figure 4, management was able to draw conclusions about the specific risk of each project based on the risk level at which projects first entered an optimal portfolio. For example P6, P10, P16 and P18 are low risk projects, P7, P8,P9 are moderate risk, and P19, P20 are high risk projects for BL1. Almost every project suggested by BL2 is a moderate risk project,



(a)Efficient Frontier for BL1

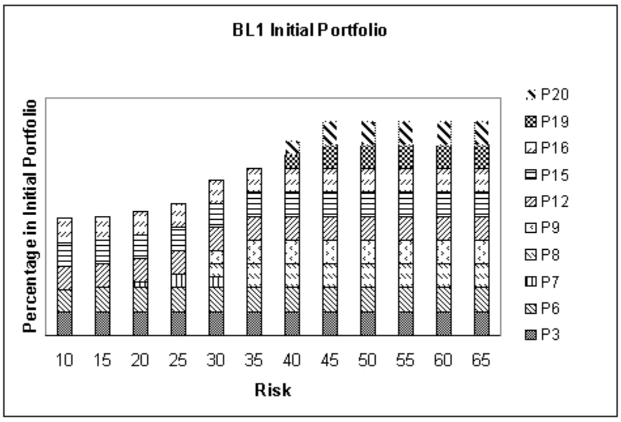


(b)Efficient Frontier for BL2

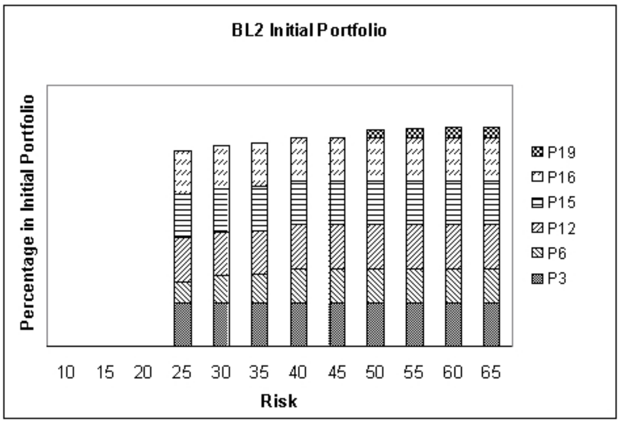
Figure 4 Efficient Frontiers for Business Lines

Table 5 Comparison of Average Expected Return in Business Lines

	Risk Level											
	10	15	20	25	30	35	40	45	50	55	60	65
BL1	(5.49)	4.05	9.21	13.74	15.94	16.59	16.39	30.80	31.02	31.23	31.51	31.52
BL2	0	0	0	10.54	20.27	24.74	26.18	27.70	28.51	29.77	29.77	29.77



(a)BL1 Initial Portfolio Composition



(b)BL2 Initial Portfolio Composition

Figure 5 Portfolio Composition for Business Lines

save for P19.

The average optimal expected return for the business lines is compared in Table 5. Management was very surprised to see that for the highest risk levels, BL1's average expected return was higher than that of BL2. Another key insight gained from the solutions was that even at the highest risk levels, both BL1 and BL2 reveal lower expected returns than the forecast figures in Table 2. Further analysis revealed that this shortfall is due to the fact that forecast figures do not adequately take into account the possibility of zero gross margins when a project fails. Another driver of the difference in forecast figures and the expected optimal figures is the probability and impact data specified by the management. It seems that for this particular instance, the management consensus was to be slightly "pessimistic" about a few key impact events.

Figure 5 depicts the initial portfolio composition for each business line. Specifically, for each project and risk level, the height of the bar for P3 for the project illustrates the percentage of instances in which the project appears in BL1's (BL2's) optimal portfolio for the first time period. Since the project P3 is fixed to be included in <u>all</u> portfolios, the height of the bar shows the scale of 100% in the figure. For example, Figure 5a graphically shows that at risk level $K = 40$, the project P20 appears in an optimal portfolio for BL1 60% of the time. Similarly, P19 at risk level $K = 50$ appears in an optimal portfolio 19% of the time. (Figure 5b)

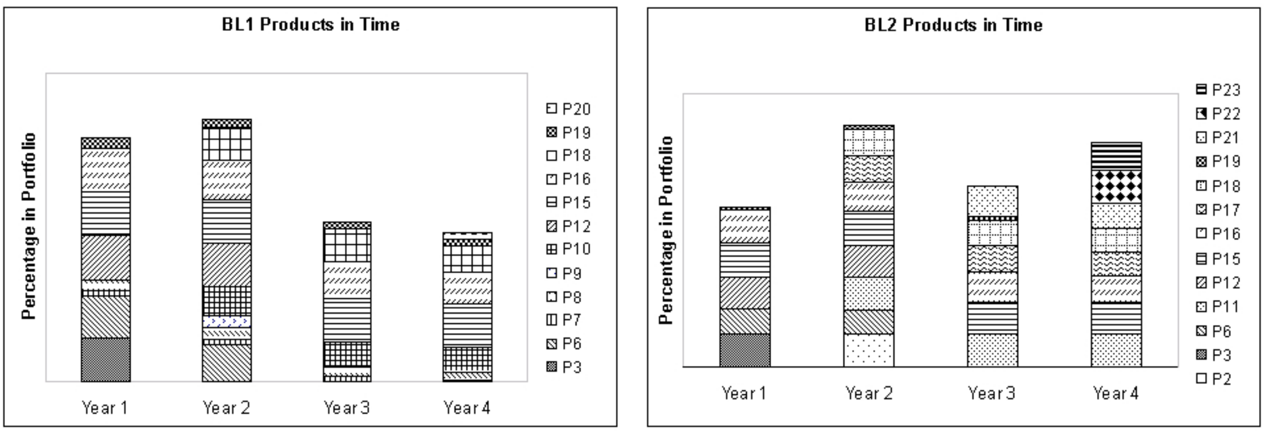
Table 6 Headcount and Profit Comparison in Time: Low Risk levels (10-25)

	Year 1		Year 2		Year 3		Year 4		Total	
	BL1	BL2	BL1	BL2	BL1	BL2	BL1	BL2	BL1	BL2
R&D	80	142	35	54	27	36	27	36	Mean:**42**	Mean:**67**
SGA	18	31	9	14	7	9	7	9	Mean:**10**	Mean:**16**
Return	(28.17)	(14.21)	(10.18)	(8.23)	5.26	5.99	37.92	19.09	**4.83**	**2.63**

Figure 6 is another illustration of the portfolio composition for BL1 and BL2. By the height of the bar for a specific project, the figure shows for each year, the percentage of optimal portfolios that contain that project in that year. In this figure, the information is averaged over all risk levels. Management found these aggregate charts useful to get an overview of the business line and the "robustness" of projects appearing in a particular business line.

Tables 6-8 display the expected yearly required headcount levels of each type of resources for both BL1 and BL2, averaged over low ($K = 10, 15, 20, 25$), medium ($K = 30, 35, 40, 45$), and high ($K = 50, 55, 60, 65$) risk levels. An important observation made by the company's management is that the required level of resources for BL2 dramatically reduces after the first year, yielding a high turnover rate. The required level of resources for BL1 is relatively more stable. The last line in the Tables 6-8 lists the average expected yearly profit for the business lines at the various risk levels. Note that neither BL1 nor BL2 are expected to return profits for the first two years of the planning horizon.

The choice of the portfolio depends on the degree of risk averseness. The optimization solutions suggest to in general follow BL1 at all the risk levels except $K = 30$ and $K = 35$. After viewing the results at various aggregations, management decided to invest in the BL1 projects. This portfolio



(a)BL1 Yearly Portfolio Composition (b)BL2 Yearly Portfolio Composition

Figure 6 Yearly Portfolio Compositions for Business Lines

Table 7 Headcount and Profit Comparison in Time: Medium Risk levels (30-45)

	Year 1		Year 2		Year 3		Year 4		Total	
	BL1	BL2	BL1	BL2	BL1	BL2	BL1	BL2	BL1	BL2
R&D	111	142	69	55	58	37	58	37	Mean:**74**	Mean:**67**
SGA	24	31	18	14	15	9	15	9	Mean:**18**	Mean:**16**
Return	(26.82)	(49.32)	(13.89)	(27.21)	9.82	23.98	45.10	74.55	**14.21**	**22.00**

Table 8 Headcount and Profit Comparison in Time: High Risk levels (50-65)

	Year 1		Year 2		Year 3		Year 4		Total	
	BL1	BL2	BL1	BL2	BL1	BL2	BL1	BL2	BL1	BL2
R&D	114	137	72	52	61	36	61	36	Mean:**77**	Mean:**65**
SGA	25	30	19	13	16	9	16	9	Mean:**19**	Mean:**15**
Return	(54.48)	(55.87)	(30.06)	(31.08)	13.14	27.5	103.21	87.58	**31.81**	**28.12**

family was in general more attractive for the following reasons; First, on the average it yields higher profit over all risk levels. Second, the portfolio and the expected profit is more responsive to the degree of risk averseness, which, according to management would give them extra flexibility. Last, the required level of resources to maintain this portfolio of projects is less variable.

After solving the case study problem, we conducted an analysis to quantify the impact of using an approximate expected risk measure (Equation (20)) to limit risk as opposed to the exact Gini risk of the portfolio. Recall that the risk measure employed for the optimization is approximate since it assumes that all projects are included in the portfolio when computing the CDF of the portfolio return in Equation (15). This approximation is necessary to keep the optimization problem solved by the DSS a linear model. To perform the analysis, fifteen additional sampled instances were created using the case study data. Each instance was solved, yielding a solution x_{pn}. For each instance, the Gini-statistic of the portfolio return was computed both by assuming all products were available in the portfolio (the approximate method) and by assuming that only the products selected by the solution x_{pn} were available in the portfolio (the true method). For the fifteen re-sampled instances, the true Gini-statistic for portfolio return differed from the approximate Gini-statistic by less than 4% on average, with a maximum difference of 11%.

A final analysis was done to attempt to quantify the value of using a stochastic programming approach against a deterministic (or mean-value) model. As such, the mean-value problem was created by fixing all the random quantities to their mean values. The long-run average profit of operating using the optimal suggested by the stochastic program versus the mean-value portfolio is known as the value of the stochastic solution (VSS) (Birge and Louveaux 1997). In our case, this analysis was performed for all levels of risk, and the long-run average cost of operating using a given portfolio was estimated using new sampled scenario sets of size 25. Table 9 shows that

Table 9 Average Value of Stochastic Solution (VSS) under different risk levels

	Risk Level											
	10	15	20	25	30	35	40	45	50	55	60	65
VSS ($ millions)	10.11	4.50	0.49	1.85	7.41	12.37	12.81	12.87	11.46	10.29	9.55	9.35

for nearly all levels of risk, the VSS was significantly greater than 0, indicating that by explicitly considering the cost of reacting to uncertainty in the model, better decisions are being made.

7. Implementation Experience

The project portfolio selection DSS described here has been constructed, revised, and remodeled over a two-year period via continuous interaction with the firm's senior management. Through our interactions with the decision makers expected to use the tools, we observed that they do not feel comfortable tinkering with the DSS during the optimization phase, but it is intuitive for them to comprehend the system's recommendations through the tables and charts. Thus, we designed the DSS such that after the initial data input, the process flows automatically, and the requisite charts and the tables are displayed at the end. The management discusses the results using these charts and tables. If more "what if" analysis or in-depth statistics are needed, the DSS allow them to delve directly into the solution through a spreadsheet interface. The system also provides utilities for decision makers to perform additional analysis by changing portions of the input. The DSS will then automatically re-run the entire process.

During our implementation, the management decided to connect the DSS with the company database so that information that is already available can flow in directly. A main concern of the management is to make sure that the DSS generates the type of charts and tables that the whole decision team can easily interpret and understand. They were heavily involved in the design of the output format such that charts and tables generated were very similar, if not identical, to what they have already being using in the process. Overall, the senior management has been very satisfied with the way the DSS tool helped in their decision process. Since it is primarily data driven while taking into consideration some human judgements, the tool creates a level of formality and credibility to the process. The firm uses the DSS for the portfolio selection and management process, and they are planning to use the tool for higher-level strategic decisions such as the alignment of market potentials and business units within the company.

8. Limitations

The company has found the DSS produced on their behalf a flexible and adaptive decision-making aid, but it is not appropriate for all R&D project portfolio problems. One limitation of the mechanism is that the stochastic program (16)-(26) can quickly grow too large to be tractable for available software. Therefore, the approach is best applied on decision problems that have few stages–either when making a short-term decision or where a coarse time-aggregation of the decision-making stages is appropriate. It is important to weight the trade off between quality of the solutions obtained and time required to solve a single instance. The decision maker can decide the problem parameters considering the problem context and by trial and error. The recourse actions modeled in the stochastic program were suitable for our business cases of interest, but they were rather limited. All product decisions were of the "go/no-go" variety, and resource decisions available were to kill a project and to adjust headcount resources appropriately. In real life the action space to select and manage a portfolio may be much more diverse. For example, the decision maker may wish to delay the start of a project or slow down the rate at which the project is completed. The model also assumes that whatever additional resources are necessary to ensure project completion can always be obtained in a timely manner, an assumption that may not be true in the business environment. In theory, each of these complications could be accounted for by increasing the complexity of the underlying stochastic program, adding additional decision variables and constraints to model the situation. However, in practice, the increased complexity of the resulting model may again pose difficulties for available software to solve.

9. Conclusions

As is typical in the high-tech environments, the semiconductor industry faces a dynamic and volatile market. The development of intellectual property (IP) via R&D projects is among the most important decisions that high-tech companies make, and the selection of a well balanced R&D portfolio requires significant effort and analysis. In this paper, we introduced a three-phase decision support structure for the project portfolio selection process at a major U.S. semiconductor company. The key features of the decision support system developed for the company are the following:

- Flexible risk modeling via scenarios, allowing the incorporation of quantitative information from the company database with qualitative information distributed among the decision makers, capturing human assessment of different aspects of the business environment.
- A multi-stage stochastic program that provides an effective way to synthesize operational information such as business constraints and project interdependence, while systematically considering and evaluating the various sources of uncertainties present in the business environment.
- An effective interface for decision makers that provides access to all components of the DSS. The interface includes detailed information gathering from company databases, surveys for key

decision makers, a wizard-like user interface to assist in model building, error-checking routines that notify the users in case of missing or illogical data, interaction with sampling and optimization tools (for advanced users), and automatically generated charts, tables, and figures.

- Sensitivity analysis tools that allow decision makers to resolve particular instances with different parameters and evaluate the robustness of the outcome.

We demonstrated the applicability and flexibility of the DSS with a real-work R&D portfolio selection case study. The case study demonstrated the typical output generated by the tools and the different ways that the decision maker may use the DSS to analyze the trade-off in different portfolio selection alternatives, balancing risk with expected return. The optimal portfolio constructed using the DSS has been implemented for a particular business unit within the company.

Acknowledgments

We would like to thank Jim Bennewitz and S. Thayer Van Winkle for dedicating their valuable time, effort, and ideas throughout the construction and the finalization of the DSS. The careful and detailed comments of two referees and the Associate Editor greatly improved the final presentation of this work.

References

Anderson, E. G., N. R. Joglekar. 2005. A Hierarchical Product Development Planning Framework. *Production and Operations Management* **14** 344–361.

Bekkers, R., G. Duysters, B. Verspagen. 2002. Intellectual Property Rights, Strategic Technology Agreements and Market Share Structure: The Case of GSM. *Reseach Policy* **31** 1141–1161.

Birge, J. R., F. Louveaux. 1997. *Introduction to Stochastic Programming*. Springer, New York.

Boisot, M. H. 1998. *Knowledge Assets, Securing Competitive Advantage in the Information Economy*. Oxford University Press, Oxford.

Cooper, R. G., S. J. Edgett, E. J. Kleinschmidt. 1998. *Portfolio Management for New Products*. Addison-Wesley, Reading, MA.

Dickinson, M. W., A. C. Thornton, S. Graves. 2001. Technology Portfolio Management: Optimizing Interdependent Projects Over Multiple Time Periods. *IEEE Transactions on Engineering Management* **48** 518–527.

Fourer, Robert, David M. Gay, Brian W. Kernighan. 2002. *AMPL: A Modeling Language for Mathematical Programming*. Duxbury Press.

Greiner, M. A., J. W. Fowler, D. L. Shunk, W. M. Carlyle, R. T. McNutt. 2003. A Hybrid Approach Using the Analytic Hierarchy Process and Integer Programming to Screen Weapon Systems Projects. *IEEE Transactions on Engineering Management* **50** 192–203.

Griffin, A. 1997. Pdma research on new product development practices: Updating trends and benchmarking best practices. *Journal of Product Innovation Management* **14** 429–458.

Kleywegt, A., A. Shapiro, T. Homem-de-Mello. 2001. The sample average approximation method for stochastic discrete optimization. *SIAM Journal on Optimization* **2** 479–502.

Linderoth, J. T., A. Shapiro, S. J. Wright. 2006. The empirical behavior of sampling methods for stochastic programming. *Annals of Operations Research* **142** 219–245.

Loch, C. H., S. Kavadias. 2002. Dynamic Portfolio Selection of NPD Programs Using Marginal Returns. *Management Science* **48** 1227–1241.

Ringuest, J. L., S. B. Graves. 2005. Formulating Optimal R&D Portfolios. *Reseach Technology Management* **48** 42–47.

Ringuest, J. L., S. B. Graves, R. H. Case. 2000. Conditional Stochastic Dominance in R&D Portfolio Selection. *IEEE Transactions on Engineering Management* **47**(4) 478–484.

Ringuest, J. L., S. B. Graves, R. H. Case. 2004. Mean-Gini analysis in r&d portfolio selection. *European Journal of Operational Research* **154** 157–169.

Santoso, T., S. Ahmed, M. Goetschalckx, A. Shapiro. 2005. A Stochastic Programming Approach for Supply Chain Network Design under Uncertainty. *European Journal of Operational Research* **167** 96–115.

Shalit, H., S. Yitzhaki. 1984. Mean-gini, portfolio theory, and the pricing of risky assets. *Journal of Finance* **39**(5) 1449–1468.

Shalit, H., S. Yitzhaki. 1989. Evaluating the mean-Gini approach to portfolio selection. *The International Journal of Finance* **1** 16–31.

Shapiro, A., T. Homem-de-Mello. 2000. On the rate of convergence of optimal solutions of Monte Carlo approximations of stochastic programs. *SIAM Journal on Optimization* **11** 70–86.

Stummer, C., K. Heidenberger. 2003. Interactive R&D Portfolio Analysis with Project Interdependencies and Time Profiles of Multiple Objectives. *IEEE Transactions on Engineering Management* **50** 478–84.

Verweij, B., S. Ahmed, A.J. Kleywegt, G. Nemhauser, A. Shapiro. 2003. The Sample Average Approximation Method Applied to Stochastic Routing Problems: A Computational Study. *Computational Optimization and Applications* **24** 289–333.

Yitzhaki, S. 1998. More Than a Dozen Alternative Ways of Spelling Gini. *Reseach on Economic Inequality* **8** 13–30.