ABSTRACT

The planning process of public and private companies relies on optimal project selection and scheduling and the efficient allocation of scarce resources. This process is complicated due in part to the fact that project investment must consider multiple criteria, project cash flows are uncertain, and there are several operational business and technical constraints. The proposed mixed-integer programming model assists the planning manager/analyst by choosing from a bank of projects in which projects to invest and when to invest. The model maximizes the sum of net present values of the chosen projects while minimizing their variance. The model satisfies simultaneously a set of precedence relations among projects; early and tardy project starting dates; exogenous budget limits; and endogenous project cash flow generation. Finally, by quantifying the opportunity cost, the model shows how arbitrary project selection and sequencing can reflect non-desirable solutions for the company and the society.

1 INTRODUCTION

The planning process for companies is complex due to the great amount of investment projects, the interrelation between them (Vonortas and Hertzfeld, 1998; Childs, Ott, and Triantis, 1998), the multiple criteria that can be relevant when evaluating each alternative (Benjamin, 1985; Ehie et. al 1990), and the constraints inherent to the corporate operation (resources, time, regulation, among others). The mathematical models proposed in the project selection literature facilitate the decision-making process avoiding the use of subjective criteria that may generate certain states susceptible of improvement or sub-optimal solutions with harmful consequences for the company or the society.

Since the pioneer work of Lorie & Savage (1955), the project selection problem has attracted several researchers. Many techniques have been applied to the project selection problem: linear programming (Benhard, 1969; Freeland and Rosenblatt 1978; Myers, 1972), multiobjective linear programming (Ringuet and Graves 1989; Ringuet and Graves, 1990), integer programming (Beged-Dov, 1965), goal programming (Benjamin 1985; Mukherjee and Bera, 1995), and evolutionary algorithms (Medaglia, 2003) among others.

Benli and Yavuz (2002) have addressed timing and sequencing in project selection problems using zero-one programming. Their model provides starting dates for the projects, while maximizing the net present value (NPV). Gupta, Kyparisis and Ip (1992) and Kyparisis, Gupta and Ip (1995) use the same criteria in order to select and sequence projects.

Multiple criteria have been used to quantify the projects’ performance both in selection and sequencing problems. The most frequently used criteria are NPV (Gupta, Kyparisis and Ip 1992; Kyparisis, Gupta and Ip, 1995; Weingartner, 1967) and risk (Kangari and Boyer, 1981; Orman and Duggan, 1999; Stone 1973). Several researchers have handled risk using the traditional Markovitz methodology (Markowitz, 1952).

To the best of our knowledge, the proposed model is a novel addition to the existing project selection literature. It combines the project selection and sequencing decisions, while considering risk and profitability as optimization criteria. The uncertainty present in the forecasts of the projects’ cash flows is the source of the NPV’s variance. The model provides an intertemporal risk diversification by including NPV covariance terms for all projects and all starting dates. The proposed model is an extension of our experience in one of Colombia’s largest water and sewage companies, where a deterministic model was successfully designed and implemented (Medaglia et. al, 2005).

This article is divided into four sections. Section 2 defines the portfolio’s expected return and variance when using forecasts. Section 3 contains the formulation of the proposed mixed-integer programming model. In Section 4, we provide computational experiments using a set of sample projects. In this Section, we also perform a sensitivity analysis and show how to construct an efficient frontier. We conclude in Section 5.

2 PROFITABILITY AND RISK UNDER THE FORECAST APPROACH

The portfolio’s profitability is measured by the net present value (NPV). Let \( P \) be the set of investment projects to be considered. Let \( T \) be the planning horizon (no investments
are made after period \( T \). Let \( y_t \) be a variable that takes the value of 1 if the project \( i \) (in \( P \)) begins in the year \( t \) (\( t \in \{0,1,\ldots,T\} \)); it takes the value of 0, otherwise. Let \( \text{NPV}_i \) be the net present value of the project \( i \) given that it starts in the year \( t \) of the planning horizon. Therefore, it is possible to express the portfolio’s NPV as follows:

\[
\text{NPV} = \sum_{i \in P} \sum_{t=0}^{T} \text{NPV}_i y_t
\]

When dealing with investment projects without historical information (e.g., infrastructure projects), the NPV for each project must be modeled according to the following steps:

- **Step 1 – Variable Identification:** Variables that can be related to the project’s cash flows must be identified.
- **Step 2 – Variable Forecast:** The identified variables are forecasted. The variance associated with the forecast must be captured for each period.
- **Step 3 – Variable-Cash Flow linkage:** Each project’s cash flows must be built in accordance to the forecasts found in Step 2.
- **Step 4 – Simulation:** Using each forecast and its variance we must simulate various realizations for each variable. This step provides multiple realizations for the NPV, one for each starting date and project. With these realizations, it is possible to calculate the NPV’s mean, variance, and covariance.

Lets assume that \( t(i) \) is the starting date of project \( i \in P \). Even though \( t(i) \) is a decision to be determined, we assume that it is known to illustrate a set of important definitions used from this point on. Let \( \beta \) be the discount factor. Let \( v_i \) be the length or lifespan of the project \( i \in P \), meaning the number of periods included between the first investment and the last positive cash flow (project’s return). Let \( F_i(\omega) \) be the cash flow for the project \( i \) in period \( t \in \{0,1,\ldots,T\} \), which is affected by the random component \( \omega \) present in the forecast. The NPV for the project \( i \) starting in period \( t=t(i) \), follows the formula:

\[
\text{NPV}_{i(t(i)} = \sum_{t=t(i)}^{t(i)+v_i} \beta^t F_i(\omega)
\]

We measure risk by computing the NPV’s volatility. The nature of this variability arises from the cash flows’ forecast. Let \( \sigma^2_i \) be the NPV variance for the project \( i \) starting in year \( t \). Let \( \text{cov}(\text{NPV}_{i(t)},\text{NPV}_{j(t)}) \) be the covariance between the NPV for the project \( i \) starting in \( t(i) \) and the NPV for the project \( j \) starting in \( t(j) \). The variance for the portfolio’s return can be expressed as follows:

\[
\sigma^2 = E\left[(\text{NPV} - E[\text{NPV}])^2\right]
\]

\[
= E\left[\left(\sum_{i \in P} \sum_{t=0}^{T} \text{NPV}_i y_t - E\left[\sum_{i \in P} \sum_{t=0}^{T} \text{NPV}_i y_t\right]\right)^2\right]
\]

Note that the terms in the form of \( \text{cov}(\text{NPV}_{i(t)},\text{NPV}_{j(t)}) \) for \( i=j \) and \( t(i) \neq t(j) \) are equal to zero due to the fact that a project is started only once during the planning horizon. The variance for the NPV of project \( i \) given that it starts in period \( t(j) \) can be expressed as follows:

\[
\sigma^2_{i(t(i))} = \text{Var}\left[\sum_{t=t(i)}^{t(i)+v_i} \beta^t F_i(\omega)\right]
\]

An approximation of this expression (and the covariance) can be tackled via Monte Carlo simulation using forecasts. Let \( H \) be the forecast horizon for the variables affecting the cash flows. Let \( f_i \) be the forecast value for period \( t \in \{0,1,\ldots,H\} \) for a given variable and \( \sigma^2(f_i) \) its variance. The proposed model does not strictly require an specific distribution for the forecasts, but for reasons explained in Section 4.1, we could assume that the forecast value is normally distributed with parameters \( (f_i, \sigma^2(f_i)) \). In general, a Monte simulation experiment can be conducted to estimate the average \( \text{NPV}_{i(t(i))} \), its variance \( \text{cov}(\text{NPV}_{i(t(i))},\text{NPV}_{j(t(j))}) \).

**3 MODEL**

Let \( P \) be the set of projects to be selected and sequenced. Let \( A \) be the set of precedence relations between projects, that is, if project \( i \in P \) precedes project \( j \in P \), then \((i,j) \in A\).

Let \( u_i \) be the length or lifespan of negative cash flows measured in periods of time (months, years), that is, the number of periods between the first and last investment costs. Let \( t'_i \) be the earliest starting date for project \( i \), meaning the earliest period in which the project can be started. Let \( t''_i \) be the latest starting date for project \( i \), meaning the lastest period in which the project can be started. Therefore, \( t'_i \leq t''_i \) must be guaranteed. Let \( N_i \) and \( N'_i \) be the minimum and maximum number of projects to be carried out, respectively. Let \( g_{ij} \) be the gap allowed between precedence relations, meaning that if project \( i \) precedes project \( j \), \( g_{ij} \) indicates the number of periods of separation or overlap between them. For example, if the overlap of investment periods is allowed between projects \( i \) and \( j \), then \( g_{ij}=1 \). In case, a strict precedence is required (no overlap or separation), then \( g_{ij}=0 \). Let \( r'_i \) be the amount of available resources for investment (i.e., budget) for year \( t \), \( t \in \{0,1,\ldots,T\} \). Let \( c_{ik} \) be the investment cost (negative cash flow) for project \( i \) in period \( k \), \( k \in \{0,1,\ldots,u_i-1\} \). On the other hand, let \( b_i \) be the expected financial income (positive cash flow) generated by project \( i \) in period \( t \).

Let us recall that \( y_t \) is the binary decision variable that takes the value of 1 if project \( i \) starts on year \( t \) (\( t=t'_i \ldots \min(t'_i', T-u_i+1) \)); it takes the value of 0, otherwise. Let \( x_{it} \) be the binary decision variable that takes the
value of 1 if for project $i$, the period $k (k=0,..., v_i-1)$ is assigned to year $t$ in the planning horizon; $(i=t'i', ..., \min \{ t'_i', T - u_i + 1 \})$, it takes the value of 0, otherwise. Let $r_i$ be the amount of investment resources not used at the end of period $t_i$, and carried over as budget for the next period $t+1$.

Finally, let $z_{gi}(t,j)$ be a binary variable with value of 1 if the projects $i$ and $j$ begin in periods $t(i)$ and $t(j)$, respectively; it takes value of 0, otherwise.

The two objective functions are shown in equations 2 and 3.

$$\max \sum_{i \in P} \sum_{t \in \mathbb{T}} NPV_i y_{it}$$ (2)

$$\min \sum_{i \in P} \sum_{t \in \mathbb{T}} z_{gi}(t,j) \text{cov}(NPV_i(t), NPV_{j(t)})$$ (3)

Equation (2) is the criterion that maximizes the portfolio’s NPV or sum of net present values of the chosen projects, whereas equation (3) minimizes the portfolio’s NPV variability.

The first set of constraints tells the model that every project may or may not be selected. It indicates that only some of the projects will be part of the optimal portfolio. This is expressed in Equation 4.

$$\min \{ t'_i, T - u_i + 1 \} \ y_{it} \leq 1, \ i \in P$$ (4)

The number of projects in the portfolio can be controlled imposing lower and upper bounds $K_i$ and $K_u$, respectively (see Equation 5).

$$N_i \leq \sum_{i \in P} \sum_{t \in \mathbb{T}} y_{it} \leq N_u$$ (5)

As shown in Equation 6, it is necessary to activate the corresponding periods of investment and income generation, once a project starts on a given period. Equation 7 guarantees that every period of investment is assigned (to a given year in the planning horizon) at most once for each project.

$$y_{it} \leq x_{ik}; \ k = 0, ..., v_i - 1, \ t = t'_i, ..., \min \{ t'_i, T - u_i + 1 \}$$ (6)

$$\sum_{i \in P} x_{it} \leq 1 ; i \in P, \ k = 0, ..., v_i - 1$$ (7)

Equation 8 shows how the precedence relations are modeled. Equation 9 explicitly forbids starting projects when it is not possible for them to be carried out according to the precedence relations. From a computational point of view, this variable fixing could be achieved efficiently using preprocessing and not explicit constraints as shown in Equation 9.

$$y_{it} \leq \sum_{t' \in \mathbb{T}} y_{j' t} (i, j) \in A, \ t - u_i - g_j \geq 0; \ t = t'_i, ..., t'_j$$ (8)

$$y_{it} = 0 ; (i, j) \in A, \ t - u_i - g_j < 0; \ t = t'_i, ..., t'_j$$ (9)

Equation 10 shows the budget constraint. This constraint includes the resources coming from the previous period and the financial income generated by the projects previously carried out.

$$r_{i+1} = r_i + r_i^0 - \sum_{i \in P} \sum_{k=0}^{v_i-1} c_{ik} x_{ik} + \sum_{i \in P} \sum_{k=0}^{v_i-1} b_{ik} x_{ik}; \ t = 0, ..., T$$ (10)

Finally, the family of constraints in Equation 11 shows the existing relation between variables $y$’s and $z$’s, allowing a linearization of the non-linear risk measure shown in Equation 1.

$$y_{it} + y_{j(t)} \geq 2z_{gi}(t,j)$$ (11)

The biobjective mixed-integer problem described by Equations 2 through 11, can be solved in two phases. In the first phase, one of the criteria is maximized or minimized (depending if it is Equation 2 or 3), subject to all constraints (Equations 4 through 11). In the second phase, the second criterion is optimized including an additional restriction that avoids the deterioration of the first objective, guaranteeing Pareto optimality (Steuer, 1986). Let $NPV^*$ and $\sigma^2$ be the optimal values for each criterion obtained in the first phase. Equations 12 and 13 show the constraints used in the second phase, depending on the objective selected in the first phase. Note that just one and only one will be used in the second phase.

$$\sum_{i \in P} \sum_{t \in \mathbb{T}} y_{it} \geq NPV^*$$ (12)

$$\sum_{i \in P} \sum_{t \in \mathbb{T}} y_{it} \geq \sum_{i \in P} \sum_{t \in \mathbb{T}} \text{cov}(NPV_i(t), NPV_{j(t)}) \leq \sigma^2$$ (13)

4 COMPUTATIONAL EXPERIMENTS

4.1 A Ten-project Example

Consider a company with 10 project investment options. The company wishes to carry out its investment plan for the next 13 years starting in 2004. Figure 1 shows the forecast for the income (cash) flows and their associated confidence interval (dotted lines). For instance, if the time series methodology is used for forecasting, it is possible to model the increasing uncertainty of distant periods. We assume that the forecasts are normally distributed to use the NPV variance as a measure of risk. For a thorough discussion about measures of risk the reader is referred to Szégo, 2004.
Let us assume, that the company has certainty of the investment costs for each project. This negative cash flows are shown in Table 1. With these investment costs and forecasted (positive) cash flows it is possible to obtain, via Monte Carlo simulation, the average NPV, its variance (\(\sigma_{\text{NPV}}^2\)) and the covariance matrix (\(\text{cov}(\text{NPV}_{it}(i),\text{NPV}_{jt}(j))\)). The discount rate used in this example is 14%.

Figure 2 shows the available investment budget during the planning horizon. For the first few years, the company shows a high budget that tends to stabilize by 2006. A shock that reduces the budget is foreseeable at the year 2013.

Table 2 shows the early (\(t^-\)) and tardy (\(t^+\)) starting dates, investment lengths (\(u_i\)), lifespans (\(v_i\)), and precedence gaps (\(g_{ij}\)) for the set of 10 projects. These exogenous restrictions originate from previous planning exercises, market knowledge, technical needs or political rules. The company requires to undertake all of its projects (\(N_i=N_j=10\)).

**4.2 Results**

Figures 3a and 3b show the results obtained for the optimal sequencing. We show in gray only the periods with investments. Note the compliance to the precedence relations. Figure 3a shows the results for the case when the maximization of the portfolio’s NPV is used as the criterion for the first phase. Figure 3b shows the results when the minimization of variance is used in the first phase. The portfolio’s NPV as well as its variance are shown for each case (see lower right of each figure).
These results suggest that the company should choose the project sequencing that balances the existing tradeoff between risk and profitability. Figure 4 shows the set of optimal portfolios generated through an iterative procedure. This procedure is divided into two parts. First, the portfolio’s variance is minimized subject to a fixed NPV value ($\tilde{NPV} = NPV^*$ in equation 12). Second, the portfolio’s NPV is maximized subject to the variance obtained in the previous step. This two-step procedure is repeated for each value of the discretized NPV’s domain. This iterative process, over non-decreasing values of the project portfolio $\tilde{NPV}^*$, ultimately produces the efficient frontier shown in Figure 4. Note that the second step of this process assures that the points obtained are indeed Pareto optimal. To assure that a good sample of the efficient frontier is obtained, the same process is carried out starting with the NPV maximization subject to fixed variance values. By discretizing the variance domain and iterating over it, it is possible to unveil some new points of the efficient frontier. Abrupt changes may be due to the nature of the problem itself; small changes in the NPV’s target may generate a radical change in the project sequencing.

When incorporating the company’s indifference curve, we obtain that the best portfolio is tangent to the highest indifference curve (point A). Figure 4 shows the efficient frontier formed by a discrete set of black diamonds and dots (extreme efficient points). Furthermore, the continuous line shows the hypothetical indifference curve for the company.

Figure 4: Optimal portfolios and the company’s indifference curve

4.3 Sensitivity Analysis

Suppose that the tardy starting dates for all the projects ($t_i^* = T, i \in P$) are relaxed. Figures 5a and 5b show the results for this scenario. Considering the NPV as a criterion for the first phase, we observe that project $p7$ is pushed away from its initial starting date, causing an increment of 22% in the portfolio’s NPV and a reduction of 58% in the portfolio’s variance. Taking into account the minimization of the variance in the first phase, we observe that 4 projects are postponed, increasing the portfolio’s NPV by 61% and decreasing the risk by 33%. This case shows that if the restrictions imposed on the tardy starting dates are not real or respond to subjective criteria (perhaps political), it is not possible to efficiently allocate the investment resources, resulting in states susceptible of improvement and preventing the company or the society as a whole from accessing better returns or lower uncertainty.

Figure 5: Late starting date relaxation scenario – (a) NPV in the first phase (b) Variance in the first phase.

Now assume that the company wishes to program a variable number of projects, allowing the selection of the number of projects between $N_l = 0$ and $N_u = 10$. Also assume that we allow the overlap of investment periods between projects $p9$ and $p10$, that is, $g_{9,10} = -1$. The remaining parameters are not altered (including the original restrictions for the starting dates). Figure 6 shows the results for this new scenario. We observe that project $p7$ is removed from the optimal bank of projects, resulting on almost doubling the portfolio’s NPV and an even more significant risk reduction. The allowed overlap between projects $p9$ and $p10$ also contributes to the portfolio’s improvement. This shows that by arbitrarily fixing a project, we may cause the resource allocation to be inefficient by undertaking risky projects that do not generate added value.

Figure 6: Optimal sequencing after relaxing the limits on the number of projects
The model uses the concept of endogenous resources represented by the projects’ positive cash flows and by carrying over into the next period the resources that were not used in the previous period. Note that the optimal sequencing shown in Figures 3a and 3b is feasible due in part by considering the initial budget of the company \((r, t)\) along with the endogenous resources. Figure 7 emphasizes the fact that the proposed sequencing is not feasible with just the initial allocated budget.

![Figure 7. Budget and investment](image)

5 CONCLUDING REMARKS

The proposed model can become part of a larger decision support tool for companies interested in selecting and scheduling a set of projects while simultaneously maximizing return and minimizing risk. The model is especially important for public companies who have the responsibility of efficiently allocating resources to obtain the best possible results for the society as a whole.

The model provides valuable information for a company with limited resources or with difficulties obtaining credit. The proposed model is able to program future financial requirements and uses endogenous cash generation intelligently. It is also possible to model legal, technical or business restrictions faced by most companies. A sample of these restrictions are early and tardy project starting dates, precedence relations among projects, and lower and upper bounds on the number of projects in the bank of projects.

The proposed mixed-integer program provides a linearized approach to a problem that originally is quadratic (see Equation 1). Our preliminary experiments show that the model scales well computationally, thus allowing us to include a great number of projects, business constraints, and interrelations that would have otherwise been difficult to include.

The model deals with projects that have not been previously carried out. We have outlined a methodology that uses forecasts of underlying variables linked to the projects’ cash flows. To generate the cash flows, we propose the use of Monte Carlo simulation, making the calculation of the portfolio’s variance and NPV possible.

The model carries out an intertemporal risk diversification by including the covariance between the projects’ returns given their respective starting years. Thus, the model does not only accounts for a negative correlation between returns referring to the same time period, but also through the variation of the project’s starting dates. A negative covariance reduces the portfolio’s volatility, scheduling projects in a time period where the covariance is as negative as possible.

This model also shows that arbitrary decisions concerning projects may provide non-desirable solutions. For companies in the public sector, the selection and sequencing of investment projects is subject to external pressures that can influence the selection of sub-optimal portfolios that do not guarantee the best allocation of resources. Using the proposed model, a company can quantify the sacrificed profits and the additional risk caused by arbitrary (perhaps political) decisions.

Finally, this model contributes to the promotion of a planning culture inside the companies, especially those in the public sector, by using technical criteria that will allow them to reach the best possible allocations for the company and the society as a whole.

REFERENCES


Ehie, I.C. and Benjamin, C.O. and Omurtag, I. and Clarke, L. 1990. Prioritizing development goals in low-income


**AUTHOR BIOGRAPHIES**

**JORGE A. SEFAIR C.** received a B.Sc. (2005) in Industrial Engineering from Universidad de los Andes. He is a B.A. student in Economics at Universidad de los Andes. He works at the Center for Studies in Economic Development (CEDE) since 2002, and at the Centro para la Optimización y Probabilidad Aplicada (COPA) since 2004. He can be reached by e-mail at j-sefair@uniandes.edu.co.

**ANDRÉS MEDAGLIA** received a B.Sc. (1992) in Industrial Engineering from Universidad Javeriana and a M.Sc. (1995) in Industrial Engineering from Universidad de los Andes. He holds a Ph.D. (2001) in Operations Research from North Carolina State University. In 2001-2002, he worked as a postdoctoral fellow in the Industrial Engineering Department at N.C. State, sponsored by SAS Institute Inc. (Cary, North Carolina). In 2002, he was appointed Assistant Professor of the Industrial Engineering Department at Universidad de los Andes. He is a founding member of the Centro de Optimización y Probabilidad Aplicada (COPA). He can be reached by e-mail at amedagli@uniandes.edu.co and his web address is http://wwwprof.uniandes.edu.co/~amedagli/